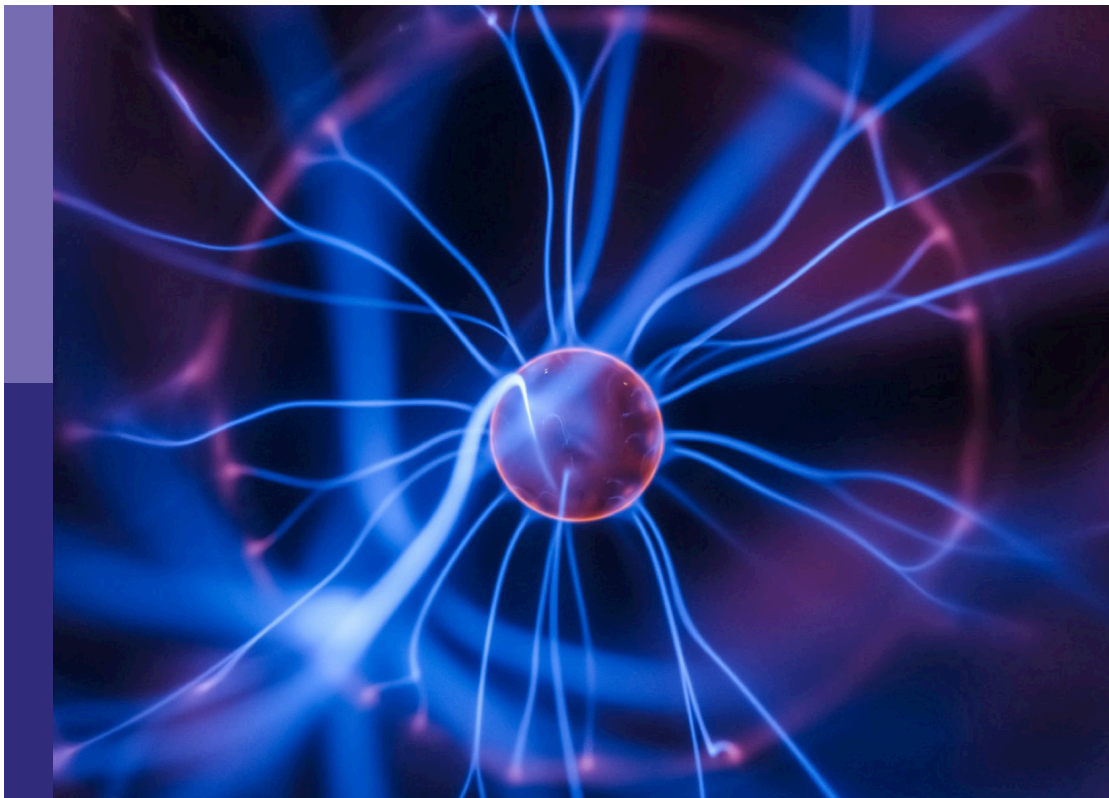
frontiers | Research Topics

Novel ideas for accelerators, particle detection and data challenges at future colliders

Edited by
Alessandro Tricoli, Patrizia Azzi, Petra Merkel and Vladimir Shiltsev

Published in
Frontiers in Physics



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-2202-8
DOI 10.3389/978-2-8325-2202-8

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Novel ideas for accelerators, particle detection and data challenges at future colliders

Topic editors

Alessandro Tricoli — Brookhaven National Laboratory (DOE), United States
Patrizia Azzi — National Institute of Nuclear Physics of Padova, Italy
Petra Merkel — Fermi National Accelerator Laboratory (DOE), United States
Vladimir Shiltsev — Fermi National Accelerator Laboratory (DOE), United States

Citation

Tricoli, A., Azzi, P., Merkel, P., Shiltsev, V., eds. (2023). *Novel ideas for accelerators, particle detection and data challenges at future colliders*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2202-8

Table of contents

05 Editorial: Novel ideas for accelerators, particle detection and data challenges at future colliders
Alessandro Tricoli, Patrizia Azzi, Petra Merkel and Vladimir Shiltsev

08 Design Concept for a Future Super Proton-Proton Collider
Jingyu Tang

26 Characterization of Low Gain Avalanche Detector Prototypes' Response to Gamma Radiation
Martin Hoeferkamp, Alissa Howard, Gregor Kramberger, Sally Seidel, Josef Sorenson and Adam Yanez

35 Electron-Hadron Colliders: EIC, LHeC and FCC-eh
Oliver Brüning, Andrei Seryi and Silvia Verdú-Andrés

46 Accelerator Technology and Beam Physics of Future Colliders
Frank Zimmermann

56 Quantum Computing Applications in Future Colliders
Heather M. Gray and Koji Terashi

66 Future Circular Collider: Integrated Programme and Feasibility Study
Michael Benedikt and Frank Zimmermann

72 Detector Simulation Challenges for Future Accelerator Experiments
John Apostolakis, Marilena Bandieramonte, Sunanda Banerjee, Nazar Bartosik, Gloria Corti, Gabriele Cosmo, V. Daniel Elvira, Thomas Evans, Andrei Gheata, Simone Pagan Griso, Vladimir Ivantchenko, Christopher Jones, Markus Klute, Charles Leggett, Ben Morgan, Tadej Novak, Kevin Pedro and Harald Paganetti

92 Multi Gigabit Wireless Data Transfer in Detectors at Future Colliders
R. Brenner, C. Dehos and E. Locci

103 Jets and Jet Substructure at Future Colliders
Johan Bonilla, Grigorios Chachamis, Barry M. Dillon, Sergei V. Chekanov, Robin Erbacher, Loukas Gouskos, Andreas Hinzmann, Stefan Höche, B. Todd Huffman, Ashutosh. V. Kotwal, Deepak Kar, Roman Kogler, Clemens Lange, Matt LeBlanc, Roy Lemmon, Christine McLean, Benjamin Nachman, Mark S. Neubauer, Tilman Plehn, Salvatore Rappoccio, Debarati Roy, Jennifer Roloff, Giordon Stark, Nhan Tran, Marcel Vos, Chih-Hsiang Yeh and Shin-Shan Yu

120 Quantum Systems for Enhanced High Energy Particle Physics Detectors
M. Doser, E. Auffray, F.M. Brunbauer, I. Frank, H. Hillemanns, G. Orlandini and G. Kornakov

136 Challenges of Future Accelerators for Particle Physics Research
Stephen Gourlay, Tor Raubenheimer and Vladimir Shiltsev

146 Tau Lepton Identification With Graph Neural Networks at Future Electron–Positron Colliders
Stefano Giagu, Luca Torresi and Matteo Di Filippo

157 RF Technologies for Future Colliders
Sergey Belomestnykh

169 Searches for long-lived particles at the future FCC-ee
C. B. Verhaaren, J. Alimena, M. Bauer, P. Azzi, R. Ruiz, M. Neubert, O. Mikulenko, M. Ovchynnikov, M. Drewes, J. Klaric, A. Blondel, C. Rizzi, A. Sfyrla, T. Sharma, S. Kulkarni, A. Thamm, A. Blondel, R. Gonzalez Suarez and L. Rygaard

194 Superconducting magnets and technologies for future colliders
Luca Bottura, Soren Prestemon, Lucio Rossi and Alexander V. Zlobin

220 Future accelerator projects: new physics at the energy frontier
Anadi Canepa and Monica D'Onofrio

frontiers | Frontiers in Physics

TYPE Editorial
PUBLISHED 04 April 2023
DOI 10.3389/fphy.2023.1177534

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Cinzia Da Via,
The University of Manchester,
United Kingdom

*CORRESPONDENCE
Alessandro Tricoli,
atricoli@bnl.gov

SPECIALTY SECTION
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

RECEIVED 01 March 2023
ACCEPTED 24 March 2023
PUBLISHED 04 April 2023

CITATION
Tricoli A, Azzi P, Merkel P and Shiltsev V (2023), Editorial: Novel ideas for accelerators, particle detection and data challenges at future colliders.
Front. Phys. 11:1177534.
doi: 10.3389/fphy.2023.1177534

COPYRIGHT
© 2023 Tricoli, Azzi, Merkel and Shiltsev. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Novel ideas for accelerators, particle detection and data challenges at future colliders

Alessandro Tricoli[1]*, Patrizia Azzi[2], Petra Merkel[3] and Vladimir Shiltsev[3]

[1]Brookhaven National Laboratory (DOE), Upton, NY, United States, [2]National Institute of Nuclear Physics of Padova, Padova, Italy, [3]Fermi National Accelerator Laboratory (DOE), Batavia, IL, United States

KEYWORDS

future colliders, detectors, instrumentation, computing, simulation, Higgs factories, event reconstruction

Editorial on the Research Topic
Novel ideas for accelerators, particle detection and data challenges at future colliders

1 Introduction

Since the observation of the Higgs boson by the ATLAS and CMS Collaborations at the Large Hadron Collider (LHC) in 2012 at CERN laboratory, which led to the 2013 Nobel Prize to Prof. Peter Higgs and Prof. François Englert, a strategic plan is being developed by the international communities of particle physicists. This foresees new collider experiments that study the properties of the Higgs boson, i.e., Higgs Factories, look for anomalies in the wider realm of the Standard Model of particle physics by performing precision measurements, for example, in the electroweak sector, and push the energy reach of particle beams to explore the unknown in the multi-TeV energy range. These new accelerator projects will follow the High Luminosity run of the LHC (HL-LHC) that is planned to start in 2029 and is expected to conclude its data-taking period about 10 years later. The most prominent proposals for Higgs Factories are electron-positron colliders, either circular or linear. In recent years, the possibility of a muon collider has attracted new and revived interest both as a Higgs Factory and as an energy-frontier machine. Proton colliders are proposed as multi-TeV machines, while also allowing a copious production of Higgs bosons. Such proton colliders will generate unprecedented radiation levels in regions close to the collision points, making even more challenging the application of new technologies that have to be highly radiation tolerant. Several accelerators are proposed to collide electrons against hadrons, which will probe the inner structure of hadrons, and may also be sensitive to new physics. Other interesting projects have been proposed, for example, the photon-photon collider, the very high-energy electron collider, *etc.*

These accelerator projects set very stringent requirements, for example, in accelerator, particle detector technologies, readout and event triggering electronics, data transmission, and computing, as well as in the algorithms for the reconstruction of collision events, particle identification, and

simulations. Advances in such technologies will not only make possible the realization of future particle physics experiments, but will also greatly benefit other areas of scientific research, industrial applications and, in the longer term, will have a significant societal impact.

This Research Topic provides a small window into the prospects for future accelerator-based experiments and focuses on novel technologies that go beyond the state-of-the-art of their fields, and push the boundaries of innovation with the goal to make giant leaps forward in the understanding of what is currently unknown. The topics present in this review include proposals for future accelerators, physics prospects at such accelerators, and research and development for the detectors.

2 Future colliders

Energy frontier particle colliders–arguably, among the largest, most complex and advanced scientific instruments of modern times–for many decades have been at the forefront of scientific discoveries in high energy physics. Due to technology advances and beam physics breakthroughs, the colliding beam facilities have progressed immensely and now operate at energies and luminosities many orders of magnitude greater than the pioneering instruments of the early 1960s. While the LHC and the Super-KEKB factory represent the frontier hadron and lepton colliders of today, respectively, future colliders are an essential component of a strategic vision for particle physics. Conceptual studies and technical developments for several exciting near- and medium-term future collider options are underway internationally.

In this Research Topic of articles, presented in depth are various aspects of the challenges, facing the future colliders. First of all there are critical beam physics considerations, such as luminosity reach of the different types of colliders, energy efficiency, particle sources and acceleration techniques, limitations due to beam instabilities, *etc.* Integrated machine design proposal of several most matured colliding beam facilities are presented for several e^+e^- Higgs Factories (high luminosity for detail explorations of the Higgs and electroweak physics), for electron-ion colliders, and energy frontier muon and large hadron (proton-proton) superconducting super-colliders.

Critical for the largest scale machines are core accelerator technologies as they directly affect facility's energy reach, size, energy consumption, cost and performance. This Research Topic includes reviews of the technologies of the accelerators' magnets and RF (radio-frequency) acceleration systems–these are linchpins in the progress toward future colliding beam facilities.

3 Physics and performance prospects

This is a unique and exciting time for the particle physics community, as the field is at a critical cross road. Despite the great success of the Standard Model, several fundamental properties of the Universe remain unexplained. Among them there is the mechanism behind the electroweak symmetry breaking, the nature and origin of dark matter, flavour and neutrino masses, the tantalizing strong CP problem, the origin of baryon asymmetry, quantum-gravity *etc.* Several future experiments are proposed to address some of these questions. These experiments include e^+e^- Higgs Factories and multi-TeV colliders. Their goals are to search for new physics beyond the SM, either directly, i.e., by producing new particles, or indirectly, i.e., by measuring precisely deviations from the expected properties of SM particles. In the absence of experimental evidence for new physics, in the past decades multiple theories and models have been proposed. Several of them predict the existence of long-lived particles, such as heavy neutral leptons, axion-like particles, new Higgs boson states, exotic decays of the Higgs boson, weakly interacting massive particles, *etc.* For the success of any future collider experiment, it is critical to efficiently and accurately reconstruct particles in collision events. As an example, the reconstruction of tau leptons is challenging at the LHC and will be important at future colliders to more precisely constrain the Higgs coupling to tau leptons and to increase the sensitivity to new physics. Towards this goal it will be beneficial to implement machine learning algorithms as well as improve the granularity of future detectors. As another example, the study of the substructure of jets of hadrons has allowed LHC experiments to exceed expectations in accessing new physics parameter space. Recent advances in experimental and theoretical techniques in jet substructure will play an important role at future lepton and hadron colliders. However, such techniques will rely on optimized detector designs, such as higher granularity, timing information, *etc.* This Research Topic presents an overview of the physics goals of future colliders at the energy frontier, in an international context, and some examples of prospects for the reconstruction and identification of particles in collision events to improve the potentials for future discoveries.

4 Detector and computational technologies

Detectors at future colliders will face several new challenges. At lepton colliders they must be made of light materials, i.e., low material budget, and highly granular, while at hadron colliders they must withstand high radiation levels. Furthermore, particle identification *via* the simultaneous measurement of several quantities, such as particle hit position, time and energy deposited, is expected to become even more important at future collider experiments, given the needed precision and thanks to recent advances in technologies. Among them, novel types of fast-timing detectors are being investigated and their radiation tolerance properties are being studied. The data acquisition and data transmission will also play a critical role. Recent studies show that multi gigabit wireless data transmission is possible and can become an alternative to the use of cables and optical links with the potential of improving the detector performance. The main advantage of such a new technology will be the transmission of large amounts of data with low power in a high radiation environment, such as at hadron collider experiments. Quantum sensing and quantum computing are expected to revolutionize particle detection and computational techniques. More specifically, quantum systems of low-dimensions as well as the ensembles of quantum systems might lead to improved calorimeter, tracking and timing detectors. In parallel, quantum computing might help solve the growing problem of the large data sets that are expected at future colliders, especially at hadron colliders, together with innovations in machine learning techniques.

Modern collider experiment rely on the accurate calculation and modeling of particle collisions and the simulation of the particle interactions with active and passive detector components. Particle simulations are important in the design stage of an experiment as well as in the optimization of the analysts strategy and ultimately in data analysis. In the past decades Monte Carlo generators and detector simulations have made great progress and have become indispensable tools for collider physics. It is expected that such tools will play an equally important role in the future, adapting to new computing platforms and computational techniques.

Author contributions

AT, VS, PM, and PA wrote this editorial contribution.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Physics

REVIEW
published: 02 March 2022
doi: 10.3389/fphy.2022.828878

Design Concept for a Future Super Proton-Proton Collider

Jingyu Tang[1,2]*

[1]Institute of High Energy Physics, CAS, Beijing, China, [2]School of Nuclear Science and Technology, University of Chinese Academy of Sciences, Beijing, China

Following the discovery of the Higgs boson at the LHC in 2012, new large colliders are being considered and studied by the international high-energy community to explore the Higgs boson in details and to probe new physics beyond the Standard Model. In China, a two-stage circular collider project, CEPC-SPPC was proposed and is under study. The first stage, CEPC (Circular Electron Positron Collier, a so-called Higgs factory) is focused on the Higgs physics, and the second stage, SPPC (Super Proton-Proton Collider) will be an energy frontier collider and a discovery machine beyond the LHC. The two colliders will share a same tunnel of 100 km in circumference, with a goal of 250 GeV in center-of-mass for CEPC and 75 TeV for SPPC Phase-I and 125–150 TeV for the SPPC ultimate goal. This article presents the design concept of the SPPC and some study results about the key accelerator physics problems and technical issues, which include luminosity optimization, beam collimation, beam-beam effects, longitudinal beam dynamics, high-field magnets and beam screen.

Keywords: proton-proton collider, physics beyond standard model, center-of-mass energy, luminosity, iron-based superconducting magnets

OPEN ACCESS

Edited by:
Vladimir Shiltsev,
Fermilab (DOE), United States

Reviewed by:
Alessandro Tricoli,
Brookhaven National Laboratory (DOE), United States
Yunhai Cai,
Stanford University, United States

***Correspondence:**
Jingyu Tang
tangjy@ihep.ac.cn

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *04 December 2021*
Accepted: *08 February 2022*
Published: *02 March 2022*

Citation:
Tang J (2022) Design Concept for a Future Super Proton-Proton Collider. Front. Phys. 10:828878. doi: 10.3389/fphy.2022.828878

1 SCIENCE REACH AT THE SPPC

SPPC (Super Proton-Proton Collider) is envisioned to be an extremely powerful machine, far beyond the scope of the Large Hadron Collider (abbr. as LHC), with a center-of-mass energy of 75 TeV, a nominal luminosity of $1.0 \times 10^{35}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ per interactive point (abbr. as IP) at the collision start, and an integrated luminosity of $30\ \mathrm{ab}^{-1}$ assuming 2 interaction points and 10–15 years of operation. A later upgrade to even higher luminosities is also possible. While the luminosity has a more modest effect on energy reach, in comparison with higher beam energy [1], raising the luminosity will likely be much cheaper than increasing the energy. The ultimate upgrading phase for SPPC is to explore physics at the center-of-mass energy of 125–150 TeV.

The CEPC (Circular Electron-Positron Collider, or a Higgs Factory) and the SPPC together will have the capability to precisely probe Higgs physics [2]. However, what the community expects more eagerly is that SPPC will explore directly a much larger region of the landscape of new physics models, and make a huge leap in our understanding of the physical world. There are many issues in energy-frontier physics that SPPC will explore, including the mechanism of Electroweak Symmetry Breaking (EWSB) and the nature of the electroweak phase transition, the naturalness problem, and the understanding of dark matter. While these three questions can be correlated, they also point to different exploration directions leading to more fundamental physics principles. SPPC will explore new ground and have great potential for making profound breakthroughs in answering all of these questions.

Extending the CEPC Higgs factory program, billions of Higgs bosons will be produced at the SPPC. This huge yield will provide important physics opportunities, especially for the rare but relatively clean physics processes.

As an energy frontier machine, the SPPC could discover an entirely new set of particles in the *O* (10 TeV) regime, and unveil new fundamental physics principles. One of the most exciting opportunities is to address the naturalness problem. This problem stems from the vast difference between two energy scales: the currently probed electroweak scale and a new fundamental scale, such as the Planck scale. Solutions to the naturalness problem almost inevitably predict the existence of a plethora of new fundamental particles not far from the electroweak scale. Such new particles will shed light on the underlying physics principles that link the low energy scale of the electroweak processes, including the light Higgs boson mass, with respect to the extremely high value of the Planck scale that sets the upper energy limit of applicability of quantum physics as we know it. Searching for these possible new particles at the LHC can probe the level of fine-tuning down to 10^{-2}, while SPPC would push this down to the unprecedented level of 10^{-4}, beyond the common concept of the naturalness principle.

Dark matter remains one of the most puzzling issues in particle physics and cosmology. Weakly interacting massive particles (WIMPs) are still the most plausible dark matter candidates. If dark matter interacts with Standard Model (SM) particles with coupling strength similar to that of the weak interaction, the mass of a WIMP particle could easily be in the TeV range, and likely to be covered at SPPC energy. Combining the relevant bounds on the mass and coupling from the direct (underground) and the indirect (astroparticle) dark matter searches, SPPC would allow us to substantially extend the coverage of the WIMP parameter space for large classes of models.

At the SPPC energy regime, all the SM particles are essentially "massless", and electroweak symmetry and flavor symmetry will be restored. The top quark and electroweak gauge bosons should behave like partons in the initial state, and like narrow jets in the final state. Understanding SM processes in such an unprecedented environment poses new challenges and offers unique opportunities for sharpening our tools in the search for new physics at higher energy scales.

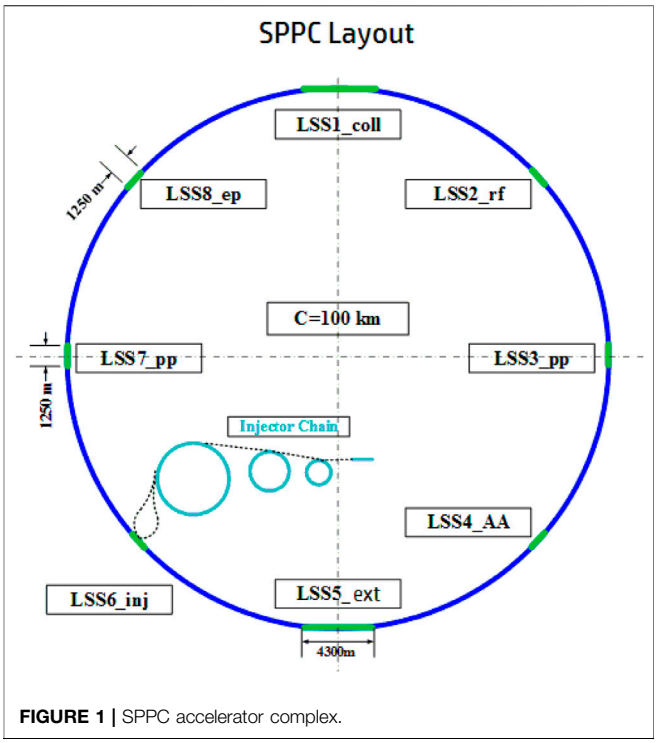


FIGURE 1 | SPPC accelerator complex.

2 THE SPPC COMPLEX

2.1 General

Since the CEPC-SPPC project is planned to be developed in two major phases: Phase One with the construction of CEPC has a timeline of 2026–2034; Phase Two with the construction of SPPC has a timeline of 2042–2050. Longer construction period for SPPC might be needed for building up the complex injector chain before the collider rings.

SPPC is a complex accelerator facility and will be able to support research in different fields of physics, similar to the multi-use accelerator complex at CERN. Besides the energy frontier physics program in the collider, the beams from each of the four accelerators in the injector chain can also support their own physics programs. The four stages, shown in **Figure 1** and with more details in **Figure 10**, are a proton linac (p-Linac), a rapid cycling synchrotron (p-RCS), a medium-stage synchrotron (MSS) and the final stage synchrotron (SS). This research can occur during periods when beam is not required by the next-stage accelerator. For example, the high-power proton beam of about 0.8 MW from the p-Linac can be used for production of intense beams of neutrons, muons and rare isotopes for a wide range of research. The high-power beams of 10 GeV from the p-RCS and 180 GeV from the MSS can be used to produce very powerful neutrino beams for neutrino oscillation experiments and the high energy beam from the SS can be used for hadron physics research.

In summary, SPPC will play a central role in experimental particle physics in the post-Higgs-boson discovery era. It is the natural next stage of the circular collider physics program after CEPC. Combining these two world-class machines will be a significant milestone in our pursuit of the fundamental laws of nature.

2.2 Design Goals

SPPC is a proton-proton collider, a discovery machine at the energy frontier. Given the 100 km circumference tunnel which is jointly defined by CEPC and SPPC, we will try to achieve the modest center-of-mass energy in proton-proton collisions with the anticipated accelerator technology and modest cost in late 2030s, but a more ambitious goal to go to higher energy is also preserved. This, of course, depends on the magnetic field that bends the protons around the collider rings: 12 T using magnets with iron-based high-temperature superconductors (iron-based

TABLE 1 | Key parameters of the SPPC baseline design.

Parameter	Value		Unit
	Phase-I	Ultimate	
Center-of-mass energy	75	125–150	TeV
Nominal luminosity	1.0×10^{35}	—	$cm^{-2}s^{-1}$
Number of IPs	2	2	—
Circumference	100	100	km
Injection energy	2.1	4.2	TeV
Overall cycle time	9–14	—	hours
Dipole field	12	20–24	T

HTS) for SPPC Phase-I and 20–24 T also using iron-based HTS magnets for Phase-II. Taking into account the expected evolution in detector technology, we can expect that the nominal luminosity of $1.0 \times 10^{35}\ cm^{-2}s^{-1}$ will be usable at the early phase but high-luminosity upgrade is also considered. At least two IPs will be made available.

This article describes the SPPC layout, basic design parameters, and its major challenges in accelerator physics and technology. It also explores the compatibility in the same tunnel with the previously built CEPC and different operating modes such as electron-proton, proton-ion, and electron-ion. Some key parameters are shown in **Table 1**.

2.3 Overview of the SPPC Design

The collider will share the same tunnel with the previously built CEPC, of circumference 100 km. A bypass scheme to avoid confliction is possible to keep the CEPC operating when adding the SPPC. The shape and symmetry of the tunnel is a compromise between the two colliders. The SPPC requires relatively longer straight sections which will be described later. This means eight identical arcs and eight long straight sections for two large detectors, injection and extraction, RF stations and a complicated collimator system. Based on expected progress in HTS technology, especially the iron-based HTS technology and also high-field magnet technology in the next 15–20 years, we expect that a field of 12 T will be attainable for the main dipole magnets with very reasonable cost or much cheaper than that based on Nb_3Sn superconductors. Twin-aperture magnets will be used for the two-ring collider. A filling factor of 78% in the arcs (similar to LHC) is assumed. The SPPC will provide beams at a center-of-mass energy of about 75 TeV.

With a circulating beam current of about 0.73 A and small beta functions (β^*) of 0.75 m at the collision points, the nominal or initial luminosity can reach $1.0 \times 10^{35}\ cm^{-2}s^{-1}$ per IP. The high beam energy, high beam current and high magnetic field will produce very strong synchrotron radiation which will impose critical requirements on the vacuum system that is based on cryogenic pumping. We expect that this technical challenge will be solved in the next 2 decades by developing efficient beam screens to extract the heavy heat load from the synchrotron radiation and reduce the electron cloud density within the beam apertures. If forced to lower the synchrotron radiation power, we would have to reduce the bunch population or the number of bunches and try to achieve a smaller β^* in order to maintain high luminosity.

As in other proton colliders using superconducting magnets, the injection energy is mainly defined by the field quality of the magnets at the bottom of their range. Persistent currents in the coils (magnetization) distort the field distribution at injection energy. Other factors favoring relatively higher injection energy are the coupling impedance, which is important to collective beam instabilities, the smaller emittance required to reduce apertures of beam screen and magnet, and the requirement on the good-field-region of the magnets. At SPPC, we have adopted a compromised injection energy of 2.1 TeV. The injector chain pre-accelerates the beam to injection energy with the required bunch current, bunch structure, and emittance. The injection chain determines the beam fill period. To reach 2.1 TeV, a four-stage injector chain is proposed: the p-Linac to 1.2 GeV, the p-RCS to 10 GeV, the MSS to 180 GeV and the SS to 2.1 TeV.

If not controlled, synchrotron cooling in SPPC would rapidly reduce the beam emittances and cause excessive beam-beam tune shifts. Noise in transverse deflecting cavities must be used to limit the minimum transverse emittances (emittance heating), and thus tune shifts. Without luminosity leveling, and with constant beam-beam tune shift, the luminosity decays exponentially from its initial peak with a lifetime of approximately 14.2 h. A shorter turnaround time (defined as the period in a machine cycle excluding the collision period), 0.8 h as base-case and 2.4 h as average-case is preferable to maximize the integrated luminosity.

There are many technical challenges in designing and building such a collider, including its injector chain. The two most difficult ones are the development and production of 12-T magnets with iron-based HTS, and the beam screen associated with very strong synchrotron radiation. Significant R&D efforts in the coming decade are needed to solve these problems.

3 KEY ACCELERATOR PHYSICS ISSUES

3.1 Main Design Parameters

3.1.1 Collision Energy and Layout

To reach the design goal for the 75-TeV center-of-mass energy with a circumference of 100 km, a modest magnetic field of 12 T is required, which is not far from the state-of-the-art magnet technology using Nb_3Sn superconductors. However, iron-based HTS technology has a bright expectation to be available and much cheaper in 10–15 years, and to generate a field higher than 20 T in the far future. Thus iron-based HTS magnet technology is chosen for SPPC. Even with the long circumference, the arc sections should be designed to be as compact as possible to provide necessary long straight sections. Although different lattice schemes are under study, it is assumed that a traditional FODO focusing structure (using uniformly distributed focusing and defocusing elements) is everywhere, except at the IPs where triplets are used to produce the very small β^*. The arcs represent most of the circumference, and the arc filling factor is taken as 0.78, similar to the LHC [3]. A key issue here is to define the number of long straight sections and their lengths. They are needed to produce those very small beta functions where the large physics detectors are to be placed, and for hosting the beam injection and extraction (abort) systems, as

TABLE 2 | Main SPPC parameters.

Parameter	Value	Unit
General design parameters		
Circumference	100	km
Beam energy	37.5	TeV
Lorentz gamma	39979	
Dipole field	12	T
Dipole curvature radius	10415.4	m
Arc filling factor	0.78	
Total dipole magnet length	65.442	km
Arc length	83.9	km
Number of long straight sections	8	
Total straight section length	16.1	km
Energy gain factor in collider rings	17.86	
Injection energy	2.1	TeV
Number of IPs	2	
Revolution frequency	3.00	kHz
Physics performance and beam parameters		
Nominal luminosity per IP	1.0×10^{35}	$cm^{-2}s^{-1}$
Beta function at collision	0.75	m
Circulating beam current	0.73	A
Nominal beam-beam tune shift limit per IP	0.0075	
Bunch separation	25	ns
Number of bunches	10080	
Bunch population	1.5×10^{11}	
Accumulated particles per beam	1.5×10^{15}	
Normalized rms transverse emittance	2.4	μm
Beam life time due to burn-off	14.2	h
Total inelastic cross section	147	mb
Reduction factor in luminosity	0.85	
Full crossing angle	110	μrad
rms bunch length	75.5	mm
rms IP spot size	6.8	μm
Beta at the first parasitic encounter	19.5	m
rms spot size at the first parasitic encounter	34.5	μm
Stored energy per beam	9.1	GJ
SR power per beam	1.1	MW
SR heat load at arc per aperture	12.8	W/m
Energy loss per turn	1.48	MeV

well as collimation systems and RF stations. Some compromises have to be made to have a relatively compact design of the long straight sections, and should be compatible with the CEPC layout. A total length of about 16.1 km is reserved for the long straight sections, with eight long straight sections of which two are 4.3 km long and the six others are 1.25 km long. With this configuration, the top beam energy is 37.5 TeV corresponding to 75 TeV in center-of-mass energy. The main parameters are listed in **Table 2**.

3.1.2 Luminosity

The initial luminosity (or nominal luminosity) of 1.0×10^{35} $cm^{-2}s^{-1}$ is much higher than in previously built machines such as the Tevatron [4] and LHC [3] and in designs such as SSC [5], VLHC [6], HE-LHC [7], and comparable to FCC-hh (Future Circular Collider–the Hadron Collider) [8, 9], though perhaps lower than in the HL-LHC [10]. In order to achieve this high luminosity, a large number of bunches and high bunch population are needed. These will be supplied by a powerful injector chain. Instant luminosity can be expressed by

$$L = \frac{n_b N_b^2 f}{4\pi \sigma_x^* \sigma_y^*} F, \tag{1}$$

where L, N_b, n_b, f, $\sigma_{x,y}^*$ and F are the instant luminosity, bunch population, number of bunches**,** revolution frequency, rms horizontal/vertical beam size at the IP, and correction factor. The SPPC initial luminosity being approximately two times the baseline design (or initial stage) of the FCC-hh [8, 11], while using the same bunch spacing, the number of interactions per bunch crossing is higher than present-day detectors could handle. It is believed, however, that ongoing R&D efforts on detectors and general technical evolution will be able to solve this problem. It also requires 1.5 times the number of protons per bunch of the FCC-hh, same for the current, and a somewhat smaller β^*. Both colliders consider even higher peak luminosity with luminosity leveling.

Another important parameter is the average, and thus integrated luminosity. One must consider the loss of stored protons from collisions, the cycle turnaround time, and the shrinking of the transverse emittance due to synchrotron radiation. The luminosity optimization and the so-called leveling are addressed in **Section 3.2.3**.

3.1.3 Bunch Structure and Population

Many bunches with relatively small bunch spacing are desirable for achieving high luminosity operation. However, the bunch spacing can be limited both by parasitic collisions in the proximity of the IPs, and by the electron cloud instability. One also needs to consider the ability of the detector trigger systems to cope with short bunch spacing. Although the bunch gap of 25 ns was designed as a baseline for LHC, the machine was operated with 50 ns bunch spacing in Run 1. This was due to problems in operation mainly from the electron cloud effect. The problems related to 25 ns at LHC was overcome in Run 2. Therefore, we have also adopted 25 ns for the nominal bunch spacing at SPPC. The bunch spacing of 25 ns is defined by the RF system in the p-RCS of the injector chain and preserved from there onward in the following steps of the injection chain [12]. The possibility of shorter bunch spacing will be investigated and is discussed below in **Section 3.2.3**.

Time gaps between bunch trains are needed for beam injection and extraction in both collider rings and the injector chain. Their lengths depend on the practical designs of the injection and extraction (abort) systems, and the rise time of the kickers for beam extraction from SPPC, assumed now to be a few microseconds. The bunch filling is taken to be about 76% of the ring circumference, which is smaller than in the LHC, and is due to the more injection gaps or batch gaps that are needed here. These gaps in the bunch structure have a significant impact on the beam dynamics during collision. On the one hand, the gaps between bunch trains are useful in suppressing collective beam instabilities; on the other hand, some bunches will meet empty bunches at the collision points or the first parasitic collision points, and those irregular bunches are called PACMAN bunches [13].

3.1.4 Beam Sizes at the IPs

The beam sizes at the IPs are determined by the β^* of the insertion lattice and the beam emittance. The initial normalized emittance of 2.4 μm is predefined in the p-RCS of the injector chain and preserved with a slight increase in the course of reaching the top energy of the SPPC due to many different factors such as nonlinear resonance crossings. The nominal beam size is 6.8 μm in rms corresponding to 0.75 m for β^* and 2.4 μm for the transverse emittance. However, at the top energy of 37.5 TeV and in the later part of the acceleration stage, synchrotron radiation will take effect, with damping times about 2.35 and 1.17 h for the transverse and longitudinal emittances, respectively. This will allow emittances to reduce significantly after the collision start time with respect to their initial values at the moment when the beams reach the top energy. Although smaller beam sizes at the IPs are favorable for luminosity, the emittances cannot be allowed to fall freely without limit because of beam-beam tune shift and detector data pileup that is caused by too high number of events per bunch crossing to be handled by the detector. Thus a stochastic emittance heating system is required to limit the synchrotron radiation cooling and control the emittance level during the collision process.

3.1.5 Crossing Angle at the IPs

To avoid parasitic collisions near the IPs producing background for experiments, it is important to separate the two beams, except at the IPs, using a crossing angle between the two beams. The crossing angle is chosen to avoid the beams overlapping at the first parasitic encounters at 3.75 m from the IPs when the bunch spacing is 25 ns At these locations the separation is no less than 12 times the rms beam size. At the SPPC, this crossing angle at the collision energy is about 110 μrad. Compared with head-on collisions, this bunch crossing angle will result in a luminosity reduction of about 15%. The crossing angle could be increased later in a more realistic design, and would have to be increased if smaller bunch spacing were to be adopted, as discussed in **Section 3.2.3**. The crossing angle has also an important impact to the dynamic aperture caused by the beam-beam effects, which is discussed in **Section 3.3**.

With a small bunch spacing the crossing angle must be larger, and the reduction of luminosity would be larger. However, there will be no luminosity loss with crossing angles when crab cavities are used. The crossing angle may be different at injection due to different lattice settings and larger emittance.

At the superconducting quadrupole triplets, the two beams are separated from each other by the crossing angle, and the apertures of the quadrupoles are increased significantly.

3.1.6 Turnaround Time

Turnaround time is about the total time period in a machine cycle when the beams are out of collision, including the programmed countdown checking time before injection, the final check with a pilot shot, the beam filling time with SS beam pulses, the ramping up (or acceleration) time, and the ramping down time. Filling one SPPC ring requires 10 SS beam pulses, which means a minimum filling time of about 14 min including pilot pulses. The ramping up and down times are each about 12.4 min. Altogether, the minimum turnaround time is 48 min, or about 0.8 h. However, the experience at LHC and other proton colliders [14] shows that only about one third of the operations from injection to the top energy are successful, thus the average turnaround time is taken to be 2.4 h. This is considered acceptable, and with a luminosity run time of 4–8 h, during which the beams are in collision, it gives a total cycle time of about 7–11 h.

3.1.7 RF Parameters

The main acceleration system at SPPC uses 400 MHz superconducting cavities. However, an additional RF system of 200 MHz is considered helpful for the longitudinal matching from the SS to the collider during injection, and a dual RF system of 400 and 800 MHz is found beneficial in suppressing instabilities and increasing luminosity at collision [12]. Although the ramping-up time is mainly defined by the superconducting magnets, the RF system must provide sufficient voltage during the process to keep up the acceleration rate with a large longitudinal acceptance. When nearing the final stage of acceleration, synchrotron radiation will play a significant role. About 10 MV in RF voltage is needed to compensate the synchrotron radiation, and the situation is similar during the collisions (and the preparation phase bringing the beams into collision). A total RF voltage of either 24 or 32 MV per beam will be provided by the 400 MHz system. Stochastic noise must be introduced to raise the longitudinal emittance to provide the long bunches required to avoid detector pileup, and avoid instabilities.

3.2 Synchrotron Radiation Related Effects

3.2.1 Synchrotron Radiation at Collision

Synchrotron radiation (SR) power is proportional to the fourth power of the Lorentz factor and the inverse of the radius of curvature in the dipoles, and becomes an important effect at multi-TeV energies using high-field superconducting dipoles. With the beam current of 0.73 A and magnetic field of 12 T, the synchrotron radiation power reaches about 12.8 W/m per aperture in the arc sections, about sixty times of that at LHC. The critical photon energy is about 1.8 keV. There is also a synchrotron radiation effect in the high-gradient superconducting quadrupole magnets.

At the SPPC, synchrotron radiation imposes severe technical challenges to the vacuum system and a probable limit on the circulating current. If absorbed at the liquid helium temperature of the magnet bores, the synchrotron radiation's heat load would be excessive, so it must be absorbed at a higher temperature. A beam screen, or other photon capture system, must be situated between the beam and the vacuum chamber. This limits the beam tube aperture, raises the beam impedance, and/or increases the required superconducting magnet bore radius. There is also a problem related to the electron cloud formation by synchrotron radiation. The technical challenges of the vacuum system and beam screen are discussed in detail in **Section 4.2**.

TABLE 3 | Relevant parameters during operation are summarized for six cases: (a) a fixed tune shift; (b) allowing the tune shift to rise to 0.03; (c) as in (b) but leveling the luminosity to its initial value; (d) as for (c) but with bunch spacing of 10 ns; (e) as for (d) but reducing β^* proportional to the emittance down to 25 cm; (f) as for (e) but with bunch spacing of 5 ns. All values are for run times maximized for a turnaround time of 2.4 h, except for the parenthesized average luminosities and integrated annual luminosities that are for a turnaround time of 0.81 h.

Case	Collis. time	Bunch spacing	Events/ crossing	Luminosity at each IP	Int. ann. lumi. at each IP	Norm. emittance	Protons/ bunch	Tune shift	Beta*
	h	ns		10^{35} cm^{-2}s^{-1}	ab^{-1}	mm-mrad	10^{11}		cm
(a)	6.86	25	418	Max. 1.20		Init. 2.4	Init. 1.50	Init. 0.015	Init. 75
				Ave. 0.68 (0.85)	0.66 (0.82)	Final 1.35	Final 0.85	Final 0.015	Final 75
(b)	5.72	25	624	Max. 1.80		Init. 2.4	Init. 1.50	Init. 0.015	Init. 75
				Ave. 1.00 (1.28)	0.97 (1.24)	Final 0.52	Final 0.64	Final 0.03	Final 75
(c)	7.64	25	418	Max. 1.20		Init. 2.4	Init. 1.50	Init. 0.015	Init. 75
				Ave. 0.88 (1.06)	0.85 (1.03)	Final 0.45	Final 0.56	Final 0.03	Final 75
(d)	8.41	10	217	Max. 1.56		Init. 2.4	Init. 0.60	Init. 0.006	Init. 75
				Ave. 0.84 (1.00)	0.81 (0.97)	Final 0.17	Final 0.20	Final 0.03	Final 75
(e)	6.12	10	358	Max. 2.58		Init. 2.4	Init. 0.60	Init. 0.006	Init. 75
				Ave. 1.34 (1.62)	1.30 (1.62)	Final 0.09	Final 0.11	Final 0.03	Final 25
(f)	8.24	5	133	Max. 1.91		Init. 2.4	Init. 0.30	Init. 0.003	Init. 75
				Ave. 1.05 (1.22)	1.02 (1.21)	Final 0.07	Final 0.063	Final 0.021	Final 25

Case (a): with the emittance heating to keep the beam-beam tune shift per IP to its initial value of 0.015. The bunch population, emittance and luminosity, all fall exponentially with time. The peak and the average luminosities are 1.20 × 10^{35} cm^{-2}s^{-1}, 0.68 × 10^{35} cm^{-2}s^{-1}, respectively.

Case (b): without the emittance heating in the beginning the beam-beam tune shift rises to 0.03 at the maximum. The average luminosity is now 1.00 × 10^{35} cm^{-2}s^{-1}, and this is a considerable gain, but 624 events per bunch crossing is considered excessive.

Case (c): a combination of Case (a) and Case (b), the peak luminosity is maintained no higher than its initial value for pileup control, but it allows the emittance damping and the maximum beam-beam tune shift to 0.03. The average luminosity is now 0.88 × 10^{35} cm^{-2}s^{-1}, which is significantly better than Case (a) and lower than Case (b).

Case (d): same as Case (c), but the bunch spacing is reduced from 25 to 10 ns and the initial bunch intensity is decreased by the same factor of 2.5 from 1.5 × 10^{11} to 6 × 10^{10}. It has almost the same the average luminosity as Case (c), but the peak luminosity is higher and the maximum number of events per bunch crossing decreases to 217 that almost eliminates the pileup concern.

Case (e): same as Case (d), but by applying dynamic β reduction following the transverse emittance damping. Luminosity leveling is applied to reduce the maximum luminosity, but significantly higher average luminosity of 1.34 × 10^{35} cm^{-2}s^{-1} is obtained.*

Case (f): same as Case (e), but now with a bunch spacing of only 5 ns. Its peak luminosity is lower, corresponding to the maximum event per bunch crossing only 133.

The synchrotron radiation also has an important impact on the beam dynamics. Without intervention, both the longitudinal and transverse emittances will shrink with lifetimes of about 2.35 and 1.17 h, respectively. The short damping times may help eliminate collective beam instabilities. One may exploit this feature to enhance the machine performance by allowing the transverse emittances to fall and to increase the luminosity. Nevertheless, to avoid excessive beam-beam tune shift (see **Section 3.3**), a stochastic transverse field noise systems will have to be installed to control the emittance reduction.

3.2.2 Luminosity Optimization and Leveling

With the nominal collision parameters, the number of events per bunch crossing (418) is higher than current detectors could accept, but it is assumed to be acceptable with future detector technology. It means that no significant pileup will happen. However, synchrotron cooling of the transverse emittance can generate luminosities greater than its initial value, and further raise the numbers of events per bunch crossing. Optimum physics use would then require a constraint on the events per bunch crossing, requiring a mechanism to limit the maximum luminosities but increase the average luminosities. Such mechanism is called luminosity leveling and could be implemented by controlling either the β^* or the emittances using the stochastic heating system. Another option is to vary the crossing angle.

Six different operation scenarios to obtain the average luminosity targets have been considered [15]. **Table 3** and **Figure 2** show the relevant parameters as a function of time. In all cases, the collision times are chosen to give the maximum average luminosities assuming the baseline turnaround time of 2.4 h. The increased average luminosities with an ideal turnaround time of 0.81 h are shown in parentheses in columns 5 and 6 in **Table 3**. Crab cavities are used to avoid luminosity reduction due to the hourglass effect, and this becomes critical if the transverse emittance is damped but a larger longitudinal emittance is maintained.

The shorter bunch spacings in Cases (d), (e), and (f) can be obtained by applying a bunch splitting scheme in MSS [16].

3.2.3 Dynamic β^* Reduction

To avoid beam loss, the beam rms size σ must be kept below a given fraction of the apertures of the inner quadrupole triplets at the IPs. If L^* is the distance from the triplet to the IP, then the beam size is given approximately by $\sigma \sim L^*(\varepsilon/\beta^*)^{1/2}$, which sets a minimum acceptable β^*. As there is no design on the detector and machine-detector interface, L^* is currently assumed to 36 m or 50% more than that of the LHC. However, as the emittance ε falls from synchrotron damping, then the β^* can be reduced in

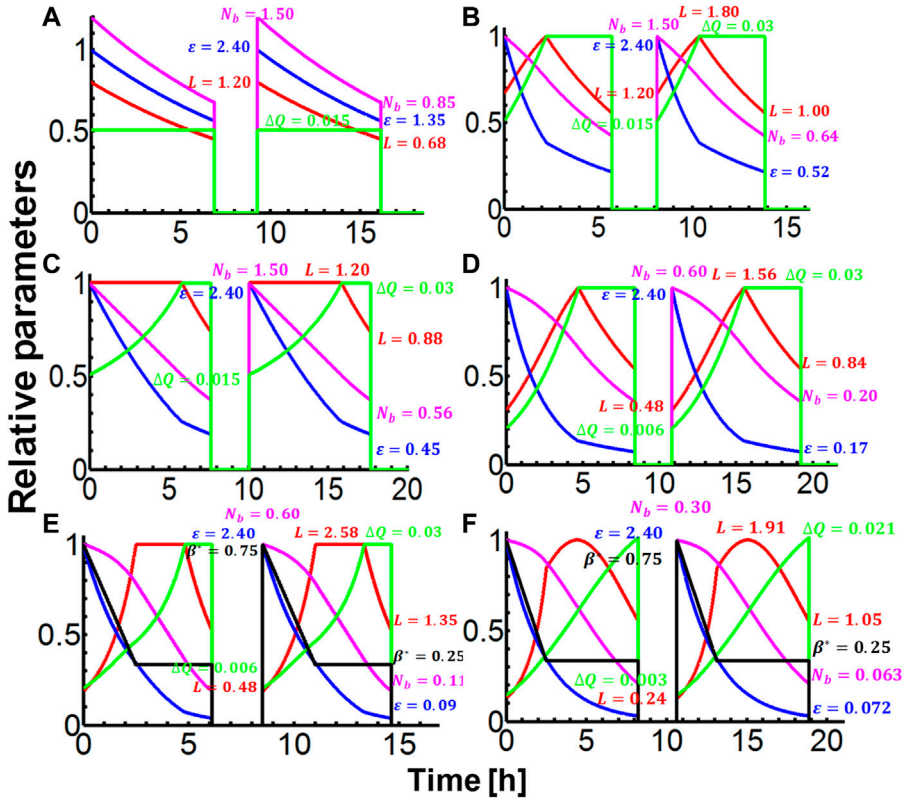


FIGURE 2 | Evolution of parameters vs. time with a turnaround time of 2.4 h and bunch spacing of 25 ns. Red: luminosity, magenta: number of protons per bunch, blue: transverse emittance, green: beam-beam tune shift, black: beta* at the IP. **(A)** with fixed tune shift; **(B)** allowing the tune shift to rise to 0.03; **(C)** as in **(B)** but with the luminosity "leveled" at its initial value; **(D)** as in **(C)** but bunch spacing of 10 ns; **(E)** as for **(D)** but reducing beta* in proportion to emittance down to 25 cm; **(F)** as for **(E)** but with bunch spacing of 5 ns. In plots **(A)**, **(B)**, **(C)** and **(D)**, beta* is kept constant at the nominal 0.75 m.

proportion, without increasing σ at the IPs. A lower limit for β^* may be set by lattice considerations, and the spot size at the IPs should not approach too close to the bunch length to avoid serious hour-glass effects. In the examples in **Table 3** and **Figure 2**, β^* reduction was limited to 0.25 m or one third of its initial value of 0.75 m.

3.3 Beam-Beam Effects

The beam-beam effects, which could lead to emittance growth, lifetime drop, and instabilities, have a very important effect on the luminosity of a collider. They come from both head-on interactions and long-range or parasitic interactions. There are several different beam-beam effects affecting the performance of a proton-proton collider: the incoherent beam-beam effects which influence beam lifetime and dynamic aperture; the PACMAN effects which will cause bunch to bunch variation; and coherent effects which will lead to beam oscillations and instabilities.

The nominal parameters given in **Table 2** are used for the preliminary study of the beam-beam effects. It is reasonable to choose a conservative nominal beam-beam tune shift parameter as 0.0075 or 0.015 for two IPs. However, LHC has reported stable operation with a total beam-beam tune shift ΔQ_{tot}~0.03 with three IPs [17], and the simulations in Ref. [18] predict that the beam-beam limit at LHC might be even larger. Thus, this limit was used for the examples in **Figure 2**.

In fact, the non-linear effects from parasitic beam-beam interactions that will be addressed below in detail are even more important in dynamic aperture, which determines the bunch spacing and the crossing angle. It is also important to consider their effects at the injection energy where the geometrical beam emittance is larger.

3.3.1 Incoherent Effects

Each particle in a beam will feel a strong nonlinear force when the beam encounters the counter rotating beam, with deleterious effects on the dynamic behavior of the particle. This nonlinear interaction will lead to an amplitude dependent tune spread for the particles in both transverse planes, which should be studied to keep the tunes away from crossing dangerous resonance lines. Earlier experiences at both the Tevatron [4] and LHC [3], required the total tune spread from all IP crossings to be kept to no more than 0.015. However, operations with larger tune shifts such as 0.02–0.03 have been reported at both Tevatron [19, 20] and LHC [21]. At SPPC, the beam-beam tune footprints from

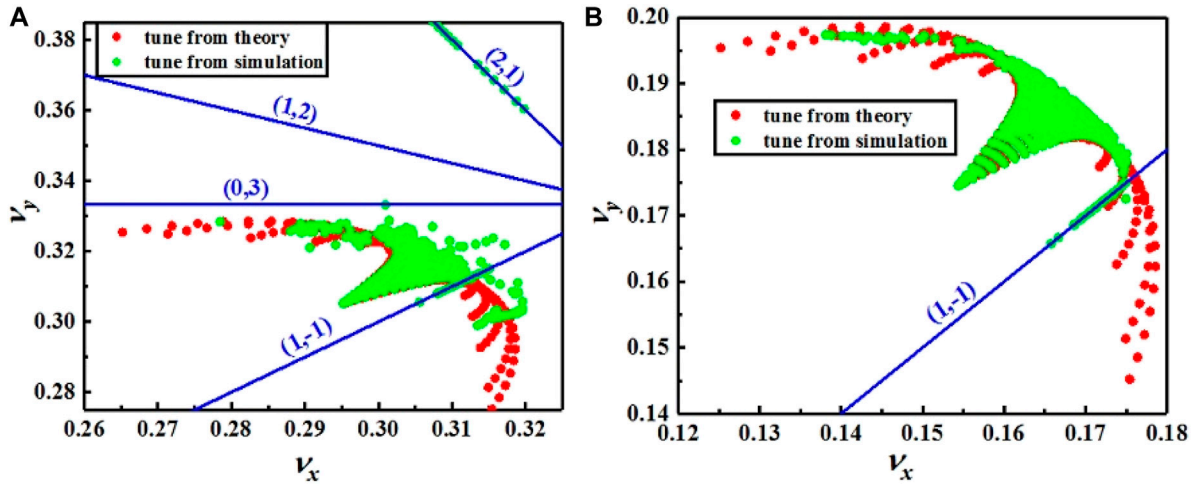


FIGURE 3 | Footprint by the beam-beam interactions for different initial nominal tunes: **(A)** (0.31, 0.32) and **(B)** (0.17, 0.19). Red and green points are from the theoretical calculation and the simulations, respectively. The lines show all nearby resonances up to the fourth order. The numbers (m, n) in brackets define resonances $m\nu_x + n\nu_y = p$, where p is an integer.

theory and simulations for two different initial tunes are shown in **Figure 3**, which include both head-on and parasitic collisions [22].

3.3.2 PACMAN Effects

The circumference and bunch number at the SPPC are 3.7 and 3.5 times of those at the LHC, respectively. With the similar bunch spacing of 25 ns it is expected that the PACMAN effects may have a similar influence as seen at the LHC. Only about half of the total bunches at the SPPC would be regular bunches that will meet the counter bunches at the IPs. The identification of regular bunches is important since the measurements such as tune, orbit or chromaticity should be selectively performed on those.

3.3.3 Coherent Effects

Coherent beam-beam effects would be expected at the SPPC because the two colliding beams are equally strong. Coherent modes of oscillations of the two counter-rotating beams are coupled by the beam-beam interaction; the coherent dipole mode is the most dangerous mode where a bunch oscillates as a rigid object around its nominal orbit. According to the LHC study in Ref. [23], it might be an option to use asymmetric collisions (different bunch intensities) at the SPPC to suppress the excitation of the coherent mode due to the beam-beam effect.

3.3.4 Dynamic Aperture

In order to achieve a higher luminosity, new ideas and technologies are under study, such as the crab waist collision scheme, beam feedback, etc. They could be effective at increasing collider luminosity. New theory and simulation work could guide the study for a luminosity upgrade in the future.

The study on the dynamic aperture at the SPPC shows that the beam-beam interactions are the most influential factor, including

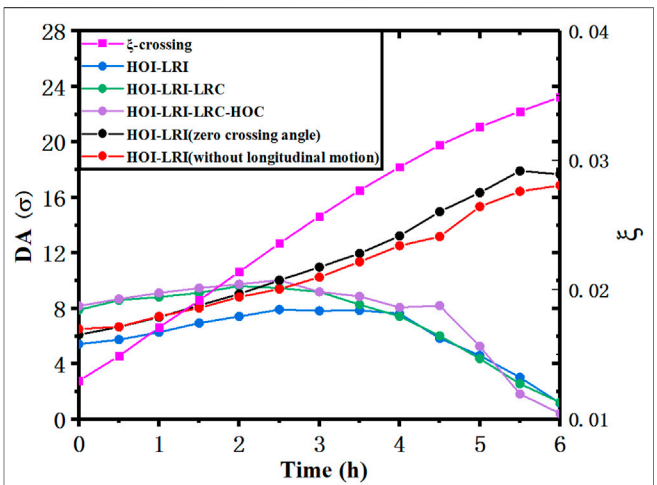


FIGURE 4 | Evolutions of the averaged dynamic aperture (DA, in circles) and beam-beam tune shift parameter (squares) with time including emittance damping and particles burning-off. HOI, LRI, HOC and LRC denote head-on interaction, long-range interaction, head-on compensation and long-range compensation.

both parasitic interactions and head-on interactions [15, 22]. Non-linear forces from parasitic interactions are usually more important in the reduction of dynamic aperture, but the large tune shift from head-on interactions will cause the resonances driven mainly by non-linear magnetic fields. By taking compensation measures such as electron lens for head-on interactions and electric wires for parasitic interactions, dynamic aperture can be restored to an acceptable level, 8 times of the rms beam size (or 8σ). In order to increase the average luminosity, a method to include both the emittance damping and proton burning-off during collision process has

TABLE 4 | Estimates on electron cloud instability for some proton-proton colliders [19–21].

	LHC	FCC-hh	SPPC
Bunch particles (10^{11})	1.15	1.0	0.3/0.6/1.5
Bunch spacing (ns)	25	25	5/10/25
Beam energy (TeV)	7	50	37.5
Pipe radius (mm)	20	13	13
Parameter n	0.165	0.189	1.334/0.331/0.053
Neutralization line density (10^{10}/m)	1.53	1.33	1.995
Neutralization volume density (10^{13}/m^3)	1.22	2.51	3.765
Wake field W/L (10^3/m^2)	1.33	3.15	3.154
Betatron tune	43.3	—	60.3
Synchrotron tune	0.006	0.002	0.005
Growth time (ms)	4.31	—	2.193
Circumference (km)	26.7	100	100
TMCI threshold electron density (10^{13}/m^3)	0.66	0.147	0.27

been studied and found very useful. It is also found that with a reduced beam size at the IPs, the coupling the transverse and longitudinal motions in the presence of a large crossing angle becomes important that limits dynamic aperture, see **Figure 4**. Thus, crab cavities are considered necessary to suppress the coupling and at the same time recover the luminosity loss due to crossing collision.

3.4 Electron Cloud Effects

3.4.1 Electron Cloud Formation

The electron cloud (EC) can cause beam instability. The build-up of accumulated photon electrons and secondary electrons was proved to be one of the most serious restrictions on collider luminosity in PEP II, KEKB, LHC [24, 25], and BEPC. The EC links together the motion of subsequent bunches and induces coupled bunch instability. It also leads to emittance blow-up and luminosity degradation [26, 27]. For the next-generation super proton colliders such as the SPPC with a bunch population higher than 10^{11} and a bunch spacing less than or equal to 25 ns, the EC effect will be critical for reaching the luminosity level of 1.6×10^{35} cm^{-2}s^{-1}.

There are three sources for the electron cloud: photon electrons, residual gas ionization and secondary electron emission. At a vacuum of about 1.3×10^{-7} Pa, the residual gas density is about 2×10^{13} m^{-3}. With an ionization cross section of 2.0 Mb, the electrons produced by gas ionization can be ignored. The necessary condition for electron amplification is that the average secondary electron emission yield (SEY) exceeds one. Electron multipacting occurs if the electrons emitted from the wall reach the opposite side wall just prior to the arrival of the next bunch. The criterion $n = \frac{r^2}{n_b r_e L_{sep}}$ can be used to estimate which kind of electrons is the dominant component in the electron cloud. In the formula, r is the radius of the vacuum pipe, n_b the number of protons in the bunch, L_{sep} is the bunch spacing and $r_e = 2.8 \times 10^{-15}$ m, the classical electron radius. If $n < 1$, some of the primary electrons are lost before the next bunch arrives and secondary electrons dominate the electron cloud; if $n > 1$, the primary electrons interact with more than one bunch and photon electrons compose most of the electron cloud. The estimated parameters n for different proton-proton colliders are listed in **Table 4**. The EC build-up saturates when the attractive beam field at the chamber wall is compensated on the average by the electron space charge field.

Most parameters in **Table 4** are hardly changed if the bunch spacing is reduced, assuming that the average current is maintained: n_b/L_{sep} = constant. However, as the bunch spacing is reduced, the parameter n changes rapidly. For a bunch spacing of 5 ns, n becomes larger than 1.0, and a large n corresponds to an almost electrostatic field that can support an electron cloud, but does not amplify it by multipacting. To produce a bunch spacing of 5 ns, a scheme applying the five-fold bunch splitting in MSS has been studied [16].

3.4.2 Electron Cloud Related Instabilities

The EC links oscillation between subsequent bunches and may lead to coupled bunch instability. The action propagated by the EC between subsequent bunches and the growth rate for the coupled bunch instability can be calculated [28, 29]. The coupled bunch instability can be damped by a feedback system. The EC also drives transverse emittance blow-up, which is very important at lower energy when the synchrotron radiation damping is very weak. For sufficiently long bunches, the single bunch instability manifests itself as strong-tail or transverse mode coupling instability (TMCI). Rough estimates on the TMCI electron density thresholds are summarized in **Table 4**. Some measures such as solenoid magnetic fields, clearing electrodes, or pipe coating should be taken to diminish the electron cloud.

The accumulated electron cloud as a focusing force on the proton beam will cause incoherent tune shift as the counterpart to space charge. A larger tune shift may lead to severe non-linear resonances. At the SPPC, with an average betatron function of about 145 m, the tune shift for an EC density of 1.0×10^{13} m^{-3} is estimated to be about 0.005 which cannot be ignored. Therefore, in the studies of lattice design and dynamic aperture, it will be necessary to consider the tune shift caused by the EC.

3.5 Beam Loss and Collimation

3.5.1 Beam Loss

Beam losses will be extremely important for safe operation in a machine like SPPC where the stored beam energy will be 9.1 GJ per beam. Beam losses can be divided into two classes, irregular and regular [30, 31]. Irregular beam losses are avoidable losses

and are often the result of a misaligned beam or due to a fault in an accelerator element. A typical example is a trip of the RF, which causes loss of synchronization during acceleration and collisions. Vacuum problems also fall into this category. Such losses can be distributed around the machine. A well designed collimator system might collect most of the lost particles, but even a small fraction of the lost particles may cause problems at other locations. Regular losses are non-avoidable and localized in the collimator system or on other aperture limits. They will occur continuously during operation and correspond to the lifetime and transport efficiency of the beam in the accelerator. The lowest possible losses are set by various effects, e.g., Touschek effect, beam-beam interactions, collisions, transverse and longitudinal diffusion, residual gas scattering, halo scraping and instabilities [31]. At the SPPC, the main concerned beam loss mechanisms are listed below, and some of them have been studied in detail.

1) Touschek effect: This, also referred to as intra-beam scattering in the longitudinal plane, is caused by the scattering of charged particles within an individual bunch, and their subsequent loss. It is typically estimated by an average of the scattering rate around the ring [32].
2) Beam-beam interactions: Beam-beam interactions at the IPs produce collisions for physics experiments, but also elastic and inelastic scattering that will lead to emittance blow-up and beam loss [32, 33].
3) Transverse and longitudinal diffusion: Resonance crossings or unstable motion caused by unavoidable field errors and higher order multipoles of ring magnets can cause beam particles to leave the confined phase area and strike the machine aperture. Particles inside the dynamic aperture may also diffuse out from the core of the beam and into the unstable region, e.g., through intra-beam scattering, beam-gas scattering and beam-gas bremsstrahlung [32, 34].
4) Residual gas scattering: This includes inelastic beam-gas nuclear inelastic interactions (both quasi-elastic and diffractive), elastic beam-gas nuclear elastic interactions (both coherent and incoherent), and Coulomb scattering. These effects degrade the beam quality and can also cause immediate beam loss [31, 34].
5) Collimator tails: Collimation is done in both betatron and momentum cleaning insertions, and also as protective measures in other regions such as collision regions. Protons that pass close to, or are only partially stopped by the collimators, can be deflected, and must be intercepted by tertiary and even quaternary collimators. But there is always some inefficiency in these systems leaving tails, also known as "tertiary/quaternary beam halo" that can be lost in other locations in the ring [31, 35].
6) Collective instabilities: A beam becomes unstable when the moments of its distribution exhibit exponential growth (e.g., barycenters and standard deviations in different coordinates) which result in beam loss or emittance growth. There are a wide variety of mechanisms which may produce collective beam instabilities, with the most important ones being the electron cloud effect as described above and coupling impedance.

3.5.2 Collimation

For high-power or high stored-energy proton accelerators, halo particles might potentially impinge on the vacuum chambers and get lost. The radiation from the lost particles will trigger quenching of the superconducting magnets, generate unacceptable background in detectors, damage radiation-sensitive devices, and cause residual radioactivity that prevents hands-on maintenance. These problems can be addressed by collimation systems which confine the particle losses to specified locations where better shielding and heat-load transfer are provided. For high-energy proton-proton colliders with very high stored-energy in the beams, like SPPC, the situation is even more complicated, mainly because extremely high collimation efficiency is required. In addition, it is very difficult to collimate very high energy protons efficiently and the material for the collimators becomes a problem due to impedance and radiation resistance issues.

To illustrate the likely systems needed for the SPPC, we discuss first those used successfully in the LHC, even though it has lower beam energy and stored energy. The LHC uses 98 two-sided and 2 one-sided movable collimators, for a total of 396 degrees of freedom, which provide a four-stage collimation system to tackle 100 MJ of stored energy per beam [36, 37]. LHC is now upgrading the systems for the incoming operation at their design energy of 14 TeV (center-of-mass), and will do additional improvements for the high-luminosity upgrade (HL-LHC) [39]. Two warm interaction regions (IRs) or long straight sections are used to provide betatron collimation and momentum collimation, respectively. Both collimation systems use the sophisticated multi-stage collimation method [39].

With the multi-stage collimation method, the primary collimators of small thickness are the closest to the beam in the transverse phase space and will scatter the primary halo particles. They must be located at large β values to maximize the impact parameters and reduce the out-scattering probability. The secondary and sometime even tertiary collimators will intercept and stop part of the scattered particles; however, they also produce out-scattered particles, which are called secondary and tertiary beam halos. The absorbers will stop the showers from upstream collimators and the additional tertiary or quaternary collimators are used to protect the superconducting quadrupole triplets at the colliding interaction regions directly [37]. The introduction of the collimation system not only demands precious space in the rings, but also increases the coupling impedance, important for collective beam instabilities. However, even this performance is considered not sufficient to prevent the superconducting magnets quenching when the LHC will be upgraded to the HL-LHC with a stored energy up to about 700 MJ [38, 40]. By studying single diffractive effects (SDE) it was found that the problem arises from that the beam loss at the downstream dispersion suppression (DS) section of the betatron collimation insertion (IR7) becomes too important. Thus the LHC is considering to add additional collimators in the DS sections.

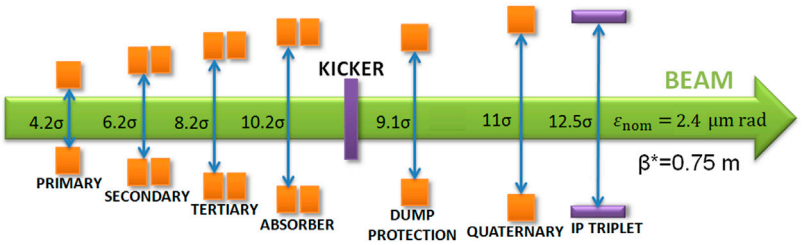


FIGURE 5 | Schematic for the five-stage collimation system at SPPC. The kicker and dump protection, which will operate in the cases of machine mal-functions, are a part of the entire machine protection system.

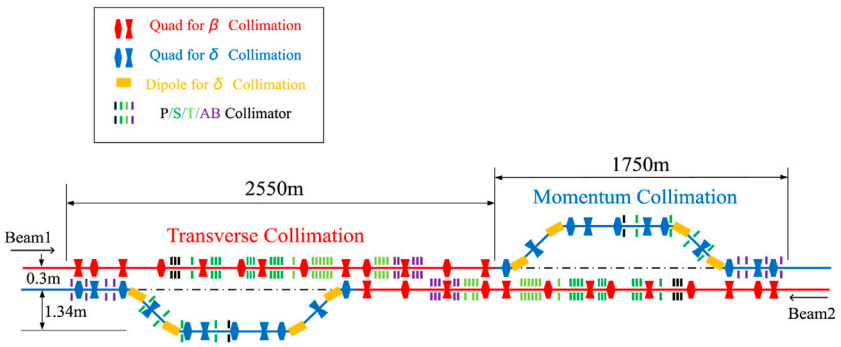


FIGURE 6 | Layout of the collimation insertion. P/S/T/AB denote primary collimator, secondary collimator, tertiary collimator and absorber.

For SPPC, the stored energy in beam is as high as 9.1 GJ per beam, about 23 times that of the LHC at design energy. Therefore, for the same beam loss power, and to prevent frequent SC magnet quenching, the cleaning inefficiency at SPPC should be about 1/23 of that at the LHC. This means a cleaning inefficiency of only 3.0×10^{-6}. Thus five-stage collimation systems for both betatron and momentum collimations are foreseen. **Figure 5** shows the schematic for such a five-stage collimation system, where four stages are in a dedicated collimation section and the fifth is at the IPs to protect the detectors. To avoid the critical SDE which becomes more important at higher energy, we developed a novel concept by combining the betatron collimation and momentum collimation in a same long straight section [41]. In this way, the particles from the SDE at the betatron primary collimators can be cleaned by the momentum collimation system, and we can avoid warm collimators in the downstream DS sections. A chicane-like structure provides dispersion for momentum collimation. One of the two very long straight sections of about 4.3 km is used to host the collimation system, see **Figure 6**. To provide the required phase advance for the four-stage collimation, both dipoles and quadrupoles are superconducting magnets instead of traditionally used warm magnets. These cold quadrupoles are very different from those in the arcs, and they will be designed with enlarged apertures and lower pole strength (not higher than 8 T) with strong radiation shielding, somewhat comparable to the triplet quadrupoles used in the experiment insertions at the LHC. Simulations show that with all the measures taken the radiation is manageable [41]. To further reduce the particle losses in magnets from the SDE, some protective or passive collimators are also used. **Figure 7** shows the simulated proton map in the collimation section. All the cold magnets are kept protected from quench.

Besides the conventional method used at the LHC with primary scattering collimators and absorbers, other novel methods will be considered, including the one studied in CERN and FNAL with bent crystals [42, 43], and the one employing nonlinear magnets to enhance the collimation efficiency [44, 45].

4 KEY TECHNICAL SYSTEMS

4.1 High-Field Superconducting Magnets

4.1.1 Requirements of the High-Field Magnets for SPPC

To bend and focus the high energy proton beams, SPPC needs thousands of high-field dipoles and quadrupoles installed around a tunnel of 100 km in circumference. The nominal aperture in these magnets is 50 mm. The field strength of the main dipoles is 12 T. A field uniformity of 10^{-4} should be attained up to 2/3 of the aperture radius. The magnets are designed to have two beam apertures of opposite magnetic polarity within the same yoke (2-in-1 or twin aperture) to save space and cost. The currently assumed distance between the two apertures in the main dipoles

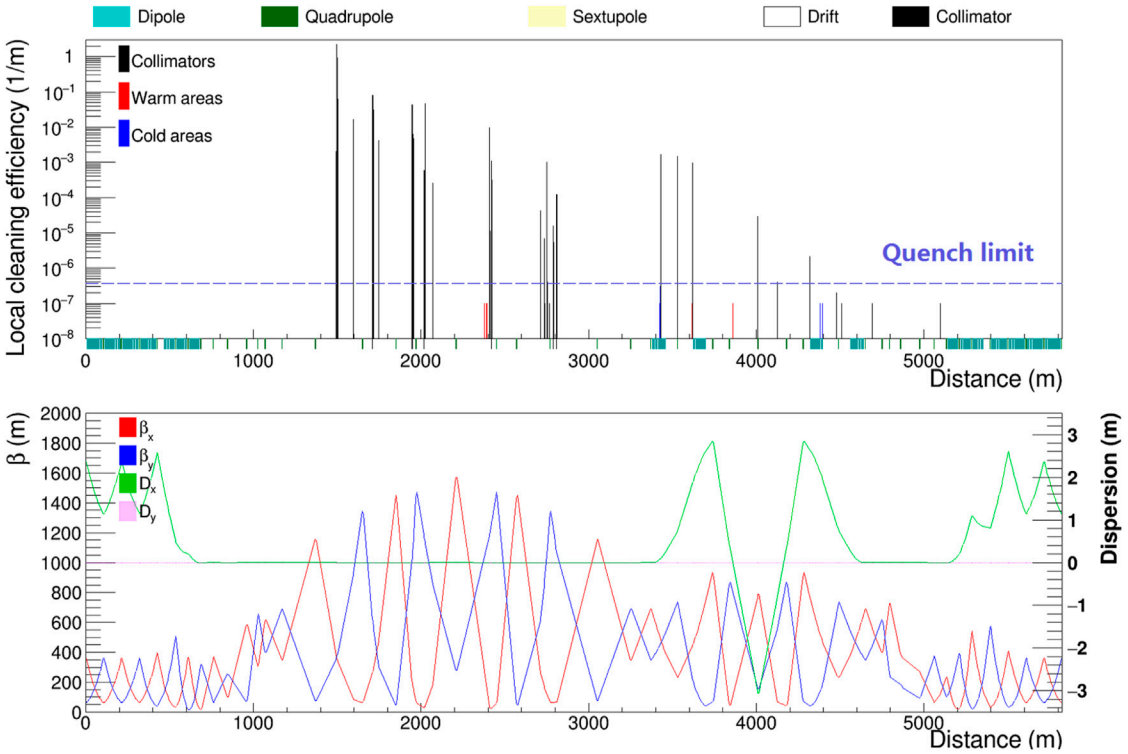


FIGURE 7 | Proton loss map in the collimation insertion with an initial vertical halo distribution (upper) and the lattice (lower).

is 300 mm, but this could be changed based on the detailed design optimization to control the cross-talk effect between the two apertures, and with consideration of overall magnet size. The outer diameter of the main dipole and quadrupole magnets should not be larger than 650 mm, so that they can be placed inside cryostats having an outer diameter of about 1,100 mm. The total magnetic length of the main dipole magnets is about 65.4 km out of the total circumference of 100 km. If the length of each dipole magnet is about 15 m, then about 4,360 dipole magnets are required. High field gradient quadrupoles for SPPC are divided into the following three groups:

1) Type A, at the IPs with single aperture, diameter: 60 mm, pole-tip field: 12 T;
2) Type B, in the matching section with twin aperture, diameter: 60 mm, pole-tip field: 12 T;
3) Type C, in the arcs with twin aperture, diameter: 50 mm, pole-tip field: 12 T.

The inner triplet quadrupoles very close to the IPs are supposed to require very special design to tackle very high radiation in the region.

4.1.2 Current Status of High-Field Accelerator Magnet Technology

One of the most challenging technologies for SPPC is the development of the high field superconducting magnets. Development of superconducting dipole magnets started more than 30 years ago in US laboratories [46, 47]. All the superconducting magnets used in present accelerators are made with NbTi. These magnets work at significantly lower field than the required 12 T (14 T is really required to have an operational margin), e.g., 3.5 T at 4.2 K at RHIC and 8.3 T at 1.8 K at LHC. As shown in **Figure 8**, the critical "engineering" current density J_E of most superconductor wires falls rapidly with the magnetic field. A reasonable design of accelerator magnets requires that the average J_E of the cable should be above 500 A/mm^2 at the desired field. This criterion suggests that it should be possible to develop a dipole with Nb_3Sn of 15–16 T, but the cost of massive production magnets is not promising even in 15–20 years from now. Thus one has to look for alternate superconductors. Fortunately, the advent of High-Temperature Superconductors (HTS) such as YBCO, BI-2212 and BI-2223, in contrary to Low-Temperature Superconductors (LTS) such as NbTi, MgB_2 and Nb_3Sn, whose current carrying capacity decreases only slowly with field (see **Figure 8**), should allow magnets with much higher magnetic fields and much reduced cost in longer term. Among different type of HTS, iron-based HTS has special advantage for isotopic field property and greater potential in future cost reduction. The SPPC strategy is to build iron-based HTS magnets of a modest field of 12 T in Phase-I, and to keep the ultimate upgrading phase to 20–24 T to boost the center-of-mass energy to 125–150 TeV.

Cryogenic is a very costly system related to superconducting magnets. For LTS technology, operation at 1.8 K, instead of 4.2 K, is another option worth study. The quantities of NbTi and Nb_3Sn,

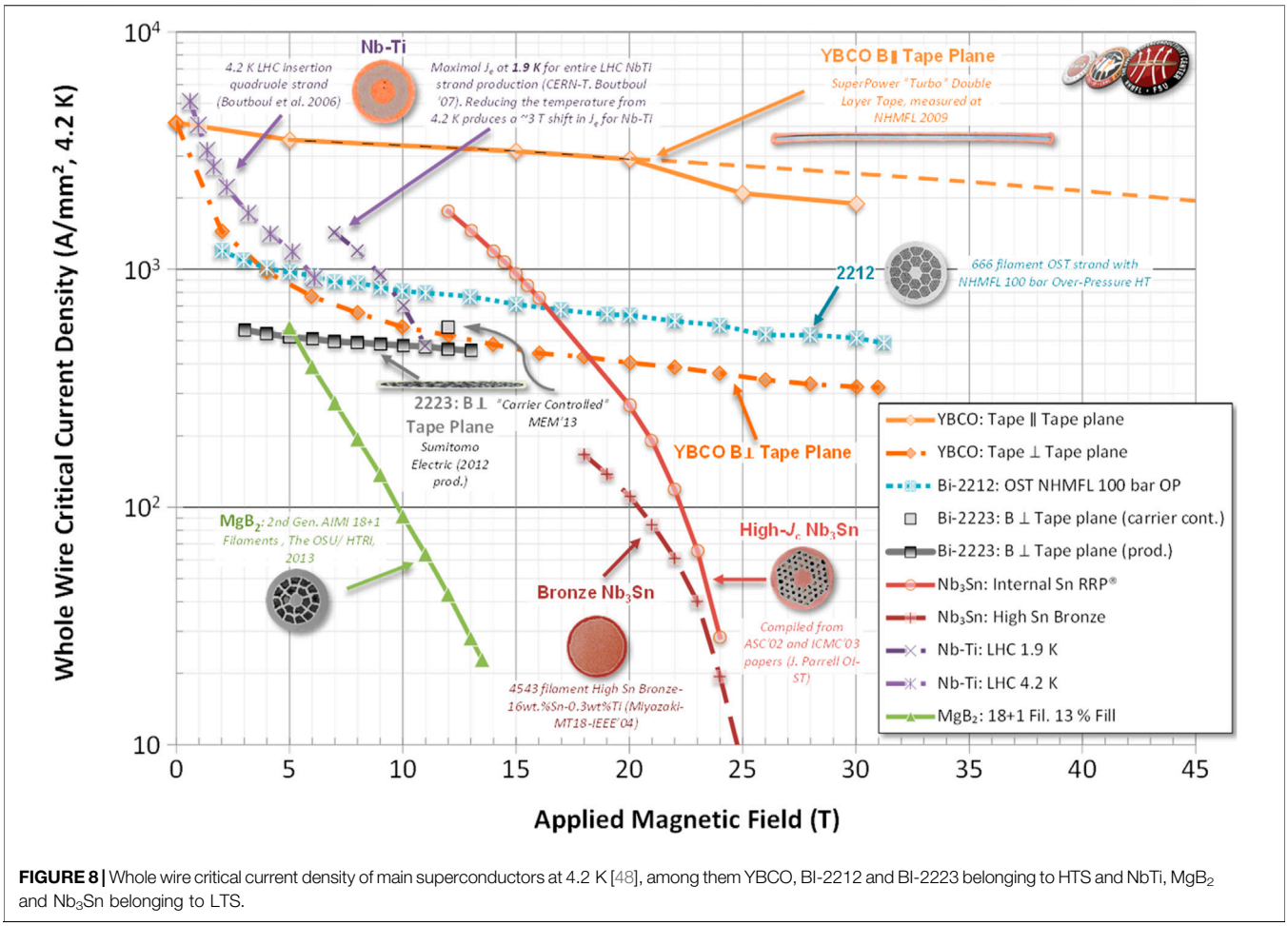


FIGURE 8 | Whole wire critical current density of main superconductors at 4.2 K [48], among them YBCO, BI-2212 and BI-2223 belonging to HTS and NbTi, MgB_2 and Nb_3Sn belonging to LTS.

and their cost, would be reduced, but the cryogenic cost would be greater. An optimization design with 1.8 K solution is required before to have the comparison with the 4.2 K solution on the global cost. Another factor affecting the choice of the temperature is the vacuum design in which the pumping speed is dependent on the temperature.

4.1.3 Development Plan of the High-Field Magnets for SPPC

As mentioned above, the current state-of-the-art magnet technology does not meet the requirements of SPPC, both in technology maturity and cost effectiveness. On the other hand, SPPC is a long-term project aiming for the period of 15–20 years or even longer, thus there is plenty of time for technology development. A roadmap has been established to develop high-field magnet technology in China, especially the one with iron-based HTS superconducting magnets. There are strong common interests from different sectors to develop iron-based HTS superconductors. A consortium consisting of many research institutions and industrial enterprises in China has been formed to develop the technology from basic research to different applications. A few development stages are needed from earlier R&D prototypes of lower field and small aperture, to full-size magnets and to mass production magnets. In the meantime, other HTS and LTS magnet technology will also be studied in parallel, in order to master the magnet structure design, field quality control, quench protection, etc.

4.2 Vacuum and Beam Screen

4.2.1 General Vacuum Considerations

SPPC has three vacuum systems: insulation vacuum for the cryogenic system, beam vacuum for the low-temperature sections, and beam vacuum for the chambers in the room-temperature sections.

4.2.1.1 Insulation Vacuum

The aim here is only to avoid convective heat transfer and there is no need for high vacuum. The room-temperature pressure in the cryostats before cool-down does not have to be better than 10 Pa. Then, so long as there is no significant leak, the pressure will stabilize around 10^{-4} Pa, when the cryostat is at cold conditions. As a huge volume of insulation vacuum is needed at SPPC, careful design is needed to reduce the cost.

4.2.1.2 Beam Vacuum in Cold Sections

In interaction regions or around experiments where superconducting quadrupoles are used, the vacuum has to be

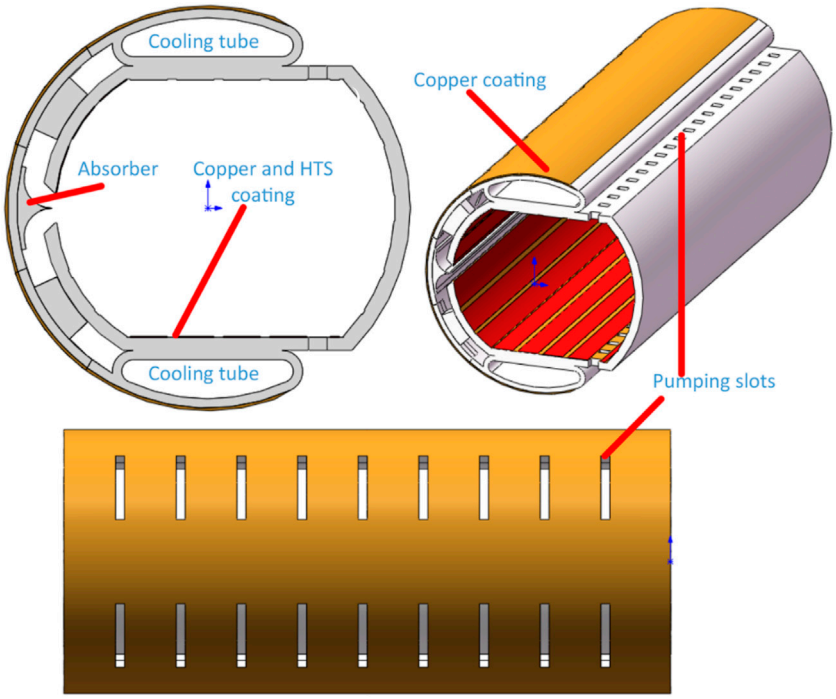


FIGURE 9 | Schematic for the beam screen at SPPC.

very good (less than 10^{13} H_2 per m^3) to avoid creating background in the detectors. But the beams are straight here and there is relatively little synchrotron radiation.

In the arcs, the requirement is based on the beam lifetime, which depends on the nuclear scattering of protons on the residual gas [3]. To ensure a beam lifetime of about 100 h, the equivalent hydrogen gas density should be below 10^{15} H_2 per m^3. The problem here is the huge synchrotron radiation power. If allowed to fall directly on the magnet bore at the magnet temperature of 4.2 K or 1.8 K, the wall power needed to remove it would be grossly too high. It has to be intercepted on a beam screen, which works at a higher temperature, e.g. 40–60 K and is located between the beam and cold bore (see below). This screen, at such a temperature, will desorb hydrogen gas, particularly if it is directly exposed to synchrotron radiation. The space outside the screen will be cryopumped by the low temperature of the bore. Slots must be introduced in the shield to pump the beam space. However, if the core is at 4.2 K, the pumping speed of H_2 is low, thus one may need to use other auxiliary methods, such as cryosorbers used at LHC [49].

4.2.1.3 Beam Vacuum in Warm Sections

The warm regions are used to house the beam collimation, injection, and extraction systems. They use warm magnets or isolated superconducting magnets to tackle with the inevitable beam losses in these locations. They have difficult vacuum pumping requirements due to desorption from the beam losses. The pumping technique with NEG (Non Evaporable Getter) coating is probably required. At least these sections are of limited overall length or much shorter than the cold sections, thus the caused trouble can be managed.

4.2.1.4 Vacuum Instability

Vacuum instability issues need further investigation [50].

4.2.2 Beam Screen

The main function of a beam screen is to shield the cold bore of the superconducting magnets from synchrotron radiation [51]. At SPPC, synchrotron radiation is especially strong because of the very high beam energy and high magnetic field in the arc dipoles. The estimated SR power is about 12.8 W/m per aperture in the arc dipoles. This is much higher than the 0.22 W/m at LHC [52], and greatly increases the difficulty of the beam screen design. The beam screen design should be a compromise to extract the heat load, minimize the occupation of the bore aperture, provide vacuum pumping, reduce coupling impedance and electron cloud, etc. An ideal design of the beam screen might separate the two principal functions—heat load transfer and vacuum pumping, which has been studied at FCC-hh [53] and also at SPPC [54]. The screen itself which encircles the beam, with the slot on the outer side would be run at a relatively lower temperature to control the impedance, while the absorption structures which synchrotron radiation penetrates through the slot would be at a higher temperature to minimize the wall power needed to extract the synchrotron radiation power.

The operating temperature of the screen must be high enough to avoid excessive wall power needed to remove the heat. But not too high to avoid excessive resistivity of the high-temperature superconducting material or copper coating on its inside surfaces, leading to excessive impedance, and to avoid radiating too much power on to the bore at 4.2 K or 1.8 K.

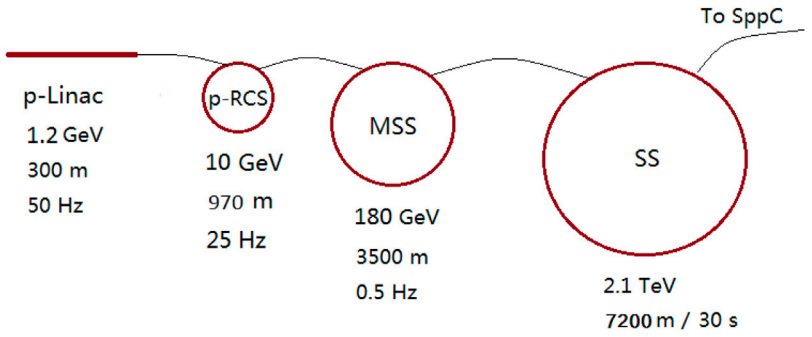


FIGURE 10 | Injector chain for the SPPC.

TABLE 5 | Main parameters for the injector chain at SPPC.

	Energy	Average current	Length/Circum.	Repetition rate	Max. beam power or energy	Dipole field	Duty factor for next stage
	GeV	mA	km	Hz	MW/MJ	T	%
p-Linac	1.2	1.4	~0.3	50	1.6/	—	50
p-RCS	10	0.34	0.97	25	3.4/	1.0	6
MSS	180	0.02	3.5	0.5	3.7/	1.7	13.3
SS	2,100	—	7.2	1/30	/34	8.3	1.3

The design must satisfy requirements of vacuum stability, mechanical support, influence on beam dynamics and refrigeration power. **Figure 9** shows the schematic for the beam screen at SPPC. The main challenges for designing the beam screen are: the working temperature should be balanced between higher temperature for wall power economy due to very high synchrotron radiation load and lower temperature for limiting the impedance; a proper beam screen structure occupies less possible aperture, has sufficient mechanical strength, shields synchrotron radiation from feeding photo-electrons in the beam path, and has a good vacuum pumping; magnet quenches have important impact on the mechanical strength of the beam screen [55], thus the materials for the base (e.g., stainless steel) and coated layers (e.g., YBCO and/or copper) are key factors in balancing the need of low impedance.

4.3 Other Technical Challenges

Besides the two key technologies described above, high-field magnets and vacuum/beam screens, there are other important technologies requiring development in the coming decade in order to build SPPC. Among them are the machine protection system that requires extremely high efficiency collimation, and a very reliable beam abort system. These are important for dumping the huge energy stored in the circulating beams, when a magnet quenches, or another abnormal operating condition occurs. If the extraction system has to be installed in a relatively short straight section, one has to develop more powerful kickers.

A complicated feedback system is required to maintain beam stability. The beam control system also controls emittance blow-up in the collider rings which is important for controlling beam-beam induced instabilities and for leveling the integrated luminosity.

Beam loss control and collimation in the high-power accelerators of the injector chain pose additional challenges. A proton RCS of 10 GeV and a few MW are still new to the community, and needs special care. The gigantic cryogenic system for magnets, beam screens and RF cavities also needs serious consideration.

5 INJECTOR CHAIN

The injector chain by itself is an extremely large accelerator complex. To reach the beam energy of 2.1 TeV required for injection into the SPPC, we require a four-stage acceleration system, with energy gains per stage between 8 and 18. It not only accelerates the beam to the energy for injection into the SPPC, but also prepares the beam with the required properties such as the bunch current, bunch structure, and emittance, as well as the beam fill period.

The four stages are shown in **Figure 10**, with some more parameters given in **Table 5**. The lower the stage, the higher the repetition rate is. The p-Linac is a superconducting linac with a repetition rate of 50 Hz. The p-RCS is a rapid cycling synchrotron with a repetition rate of 25 Hz. The MSS has a relatively lower repetition rate of 0.5 Hz. The SS, which is based on superconducting magnets with maximum dipole field of about 8 T, is even slower. The higher repetition rates for the earlier stages help reduce the SS cycling

period and thus the overall SPPC beam fill time. For easier maintenance and cost efficiency, as well as the physics programs, the first three stages will be built in a relatively shallow underground level, e.g., −15 m, whereas the SS with a much larger circumference will be built in the same level as the SPPC or from −50 m to −100 m.

As shown in **Table 5**, for the SPPC, the different stages are needed for only fractions of the time. They could operate with longer duty cycles, or continuously, to provide high-power beams for other research applications, when they do not serve the SPPC collision. As the present bunch population at the SPPC is limited mainly by the SR power, the accelerators of the injector chain have the potential to load more accumulated particles in a pulse or deliver higher beam power for their own diverse applications when not serving the SPPC. Certainly this capability is also very useful for the future SPPC upgrading.

For such a complex injector system, it may take about 10 years to build and commission stage-by-stage. Thus hopefully the construction of the injector accelerators can be started several years earlier than the SPPC, and this means that it overlaps with the CEPC physics operation.

5.1 p-Linac

Superconducting linacs have undergone tremendous development [56] in the last 2 decades and will presumably make even more progress in the next decade. Hence we have adopted a design of 1.2 GeV in energy and 50 Hz in repetition rate for the p-Linac. The continuous beam power is 1.63 MW. At least half of this could be available for other applications.

5.2 p-RCS

The continuous beam power from p-RCS is 3.4 MW. Only one other proton driver study (for a future Neutrino Factory) has performance close to this [57]. The high repetition rate of 25 Hz will shorten the beam filling time in the MSS. Only a fraction of this power is needed to fill the MSS. Thus most of the beam pulses from the p-RCS could be used for other physics programs. The p-RCS will use mature accelerator technology but be on a larger scale than existing rapid-cycling proton synchrotrons. High-Q ferrite loaded RF cavities are planned to provide very high RF voltage of about 3 MV, and the RF frequency swing of 36–40 MHz is suitable to provide the bunch spacing of 25 ns needed by SPPC.

5.3 MSS

The MSS has beam power similar to the p-RCS but with much higher beam energy and much lower repetition rate. The SPS at CERN and the Main Injector at Fermilab are two good examples for its design. But due to much higher beam power, the beam loss rate must be more strictly controlled. The same RF system as in the p-RCS is planned, but a more sophisticated multi-harmonic RF system is reserved for the future bunch splitting to provide 5-ns bunch spacing. Certainly, the beam from the MSS will find additional physics programs other than only being the injector for the SS.

5.4 SS

The SS will use superconducting magnets similar to those used at the LHC, but with a higher ramping rate. Here, we do not need to consider synchrotron radiation because of the much lower energy. There are no apparent critical technical risks in building the SS. It is still unclear if the beam from the SS can find its own physics programs besides being the SPPC injector.

A dedicated heavy-ion linac (i-Linac) together with a new heavy-ion synchrotron (i-RCS), in parallel to the proton linac/RCS, is needed to provide heavy-ion beams at the injection energy of the MSS, with a beam rigidity of about 36 Tm which is the same as the proton beam of 10 GeV.

6 DISCUSSIONS AND CONCLUSION

A design concept for a future proton-proton collider of 75 TeV in center-of-mass energy has been studied, aiming to achieve high precision in measuring the properties of the Higgs boson and probe the high energy frontier in search for new physics beyond the standard Model. Both the physics potentials and the accelerator scheme are outlined here. The machine performance and key issues on accelerator physics and technology are addressed. Although the CEPC-SPPC project is intended to be hosted by China, the study presented here is totally site-independent.

There are many uncertainties in the physics goals, since the project is supposed to be built in 20 years from now. It is already under discussion how to make a compromised layout to accommodate both CEPC and SPPC in the same tunnel.

On the one hand, the progress in general accelerator technology during the next 2 decades may make those extremely challenging accelerator designs feasible. On the other hand, international efforts to overcome the technical obstacles for building such a machine should be pursued. Fortunately, with the ongoing efforts for the SPPC and FCC-hh projects as a driving force, an international community has already been established to tackle key technical issues such as high-field superconducting magnets, cryogenic beam vacuum or beam screen, beam collimation, etc.

AUTHOR CONTRIBUTIONS

JT reviews the study that has been carried out by the SPPC study group since 2014.

FUNDING

The work is supported by National Natural Science Foundation of China (Projects 11575214, 12035017).

ACKNOWLEDGMENTS

The author would like to thank all the international experts who have provided us with useful comments during the reviews and on

other occasions. Special gratitude goes to the following persons for their contributions to the study: J. Scott Berg (LBNL), Weiping Chai (IMP), Fusan Chen (IHEP), Nian Chen (USTC), Yukai Chen (IHEP), Weiren Chou (IHEP/FNAL), Jianping Dai (IHEP), Haiyi Dong (IHEP), Angeles Faus-Golfe (LAL), Pingping Gan (PKU), Jie Gao (IHEP), Yuanyuan Guo (IMP), Ramesh Gupta (BNL), Tao Han (Pitt), Yang Hong (IHEP), Alexander Krasnov (BINP), Yongbin Leng (SINAP), Guangrui Li (THU), Peng Li (IMP), Zhihui Li (SCU), Baiqi Liu (IHEP), Yudong Liu (IHEP), Xinchou Lou (IHEP/UTD), Yuanrong Lu (PKU), Qing Luo (USTC3, Ernie Malamud (FNAL), Lijun Mao (IMP), James Molson (LAL), Robert B. Palmer (BNL), Quanling Peng (IHEP), Yuemei Peng (IHEP), Qing Qin (IHEP), Stefano Redaelli (CERN), Manqi Ruan (IHEP), GianLuca Sabbi (LBNL), Frank Schmid (CERN), Tanaji Sen (FNAL), Feng Su (IHEP), Shufang Su (UArizona), Diktys Stratakis (BNL), Baogeng Sun (USTC), Meifen Wang (IHEP), Jie Wang (USTC), Liantao Wang (UChicago), Lijiao Wang (IHEP), Xiangqi Wang (USTC), Yifang Wang (IHEP), Yong Wang (USTC), Qingzhi Xing (THU), Qingjin Xu (IHEP), Hongliang Xu (USTC), Wei Xu (USTC), Holger Witte (BNL), Yingbing Yan (SINAP), Yongliang Yang (USTC), Jiancheng Yang (IMP), Ye Yang (THU), Youjin Yuan (IMP), Bo Zhang (USTC), Linhao Zhang (IHEP), Yuhong Zhang (JLAB), Shuxin Zheng (THU), Kun Zhu (PKU), Zian Zhu (IHEP), Ye Zou (IHEP/USTC). Most of the materials were presented in the CEPC-SPPC Preliminary Conceptual Design Report and CEPC Conceptual Design Report [2]. Figures 3 and 4 are from an unpublished paper (Beam-beam effects and mitigation in a future proton-proton collider, L.J. Wang et al.).

REFERENCES

1. Hinchliffe I, Kotwal A, Mangano ML, Quigg C, Wang LT. *Luminosity Goals for a 100-TeV Pp Collider*. arXiv:1504.06108 (2015).
2. The CEPC-SPPC Study Group. CEPC-SPPC Preliminary Conceptual Design Report. In: *2015: The CEPC Study Group, CEPC Conceptual Design Report: Volume 1 - Accelerator*. arXiv:1809.00285 [physics.acc-ph], IHEP-CEPC-DR-2015-01 (2018). p. 01.
3. Brüning O, Collier P, Lebrun P, Myers S, Ostojic R, Poole J, et al. LHC Design Report. In: *The LHC Main Ring*. CERN (2004). p. 003.
4. INSPIRE. *Tevatron Design Report*. Fermilab-design (1983). p. 01.
5. SSC Central Design Group. *Conceptual Design of the Superconducting Super Collider*. SSC-SR-2020 (1986).
6. The VLHC Design Study Group. *Design Study for a Staged Very Large Hadron Collider*. Fermilab TM-2149 (2001).
7. E Todesco F Zimmermann, editors. Proc. Of EuCARD AccNet-EuroLumi Workshop: The Proceedings EuCARD AccNet-EuroLumi Workshop: The High-Energy Large Hadron Collider, 'HE-LHC-2010'. Report CERN-2011-003 (2011). p. 14–6.
8. Ball A. *Future Circular Collider Study Hadron Collider Parameters*. FCC-ACC-SPC-0001 (2014).
9. Abada A. FCC-hh: The Hadron Collider. *Eur Phys J Spec Top* (2019) 228: 755–1107.
10. Zimmermann F. EuCARD-CON-2011-002. In: Proc. Of IPAC-2012. New Orleans: IEEE (2012).
11. Benedikt M. Study Status & Parameter Update, In: FCC Week; 2016 April 11–15; Rome, Italy (2016). p. 11–5.
12. Zhang LH, Tang JY, Hong Y, Chen YK, Wang LJ. Optimization of Design Parameters for SPPC Longitudinal Dynamics. *J Inst* (2021) 16:P03035. doi:10.1088/1748-0221/16/03/p03035
13. Herr W. *Effects of PACMAN Bunches in the LHC*. CERN-LHC-Project-Report-39 (1996).
14. Bruening O. Accelerator Physics Requirements at Commissioning, In: 11th Workshop of the LHC, CERN-SL-2001-003; 2001 Jan 15–19; Chamonix, France (2001).
15. Wang LJ, Tang JY. Luminosity Optimization and Leveling in the Super Proton-Proton Collider. *Radiat Detect Technol Methods* (2021) 5:245–54. doi:10.1007/s41605-020-00233-6
16. Zhang L, Chen M, Tang J. Multifold bunch Splitting in a High-Intensity Medium-Energy Proton Synchrotron. *Results Phys* (2022) 32:105088. doi:10.1016/j.rinp.2021.105088
17. Shiltsev V. *Beam-beam Effects in a 100 TeV P-P Future Circular Collider*. Washington DC: presentation at FCC Week 2015 (2015).
18. Ohmi K, Dominguez O, Zimmermann F. *Beam Beam Studies for the High-Energy LHC*". CERN Yellow Report CERN-2011-003 (2011). p. 93–8.
19. Sen T. Observations and Open Questions in Beam-Beam Interactions. *ICFA Beam Dyn Newslett* (2010) 52:14–32.
20. Shiltsev V. High Luminosity Operation, Beam-Beam Effects and Their Compensation in TEVATRON. In: Proceedings of EPAC08. IEEE (2008). p. 951–5.
21. Papotti G, Buffat X, Herr W, Giachino R, Pieloni T. *Observations of Beam-Beam Effects at the LHC*. arXiv:1409.5208 (2016).
22. Wang LJ. *Study on Luminosity of Super Proton-Proton Collider*. [PhD thesis]. University of Chinese Academy of Sciences (Institute of High Energy Physics, CAS) (2021).
23. Alexahin Y, Grote H, Herr W, Zorzano MP. Coherent Beam-Beam Effects in the LHC. *LHC Project Rep* (2001) 469:1.
24. Miguel A. Furman, "The Electron-Cloud Effect in the Arcs of the LHC". *LHC Project Rep* (1998) 180:1.
25. Zimmermann F, Rumolo G. *Two-stream Problems in Accelerators*. CERN, SL-2001-057 (AP) (2001).
26. Arduini G, Baglin V. Present Understanding of Electron Cloud Effects in the Large Hadron Collider. In: *PAC 2003*. Portland, Oregon, USA: IEEE (2003). p. 12–6.
27. Benedetto E, Schulte D, Zimmermann F, Rumolo G. Simulation Study of Electron Cloud Induced Instabilities and Emittance Growth for the CERN Large Hadron Collider Proton Beam. *Phys Rev ST Accel Beams* (2005) 8: 124402. doi:10.1103/physrevstab.8.124402
28. Ohmi K. Beam-Photoelectron Interactions in Positron Storage Rings. *Phys Rev Lett* (1995) 75:1526–9. doi:10.1103/physrevlett.75.1526
29. Zimmermann F. *A Simulation Study of Electron-Cloud Instability and Beam-Induced Multipacting in the LHC*. LHC-PR-95, SLAC-PUB-7425 (1997).
30. Holzer EB, Dehning B, Effnger E, Emery J, Grishin V, Hajdu C, et al. Beam Loss Monitoring for LHC Machine Protection. *Phys Proced* (2012) 37:2055–62. doi:10.1016/j.phpro.2012.04.110
31. Drozhdin AI, Mokhov NV, Striganov SI. Beam Losses and Background Loads on Collider Detectors Due to Beam-Gas Interactions in the LHC. In: Proceedings of PAC09. Vancouver, BC, Canada: IEEE (2009). p. 2549–51.
32. Wittenburg K. Beam Loss Monitoring and Control. In: Proceedings of EPAC 2002. Paris, France: IEEE (). p. 109–13.
33. Bruce R. *Beam Loss Mechanisms in Relativistic Heavy-Ion Colliders*. Doctoral Thesis, CERN-THESIS-2010-030 (2010). p. 39–45.
34. Mead DJ. The Measurement of the Loss Factors of Beams and Plates with Constrained and Unconstrained Damping Layers: A Critical Assessment. *J Sound Vibration* (2007) 300:744–62. doi:10.1016/j.jsv.2006.08.023
35. Roderik Bruce D. *Off-momentum Collimation and Cleaning in the Energy Ramp in the LHC*. CERN-THESIS-2013-159 (2013). p. 17–23. Doctoral Thesis.
36. Redaelli S. Operational Performance of the LHC Collimation. In: Proc. Of HB2010. Morschach, Switzerland: IEEE (2010). p. 395–9.
37. Salvachua B. Cleaning Performance of the LHC Collimation System up to 4 TeV. In: Proc. Of IPAC2013. Shanghai, China: IEEE (2013). p. 1002–4.
38. Serluca M, Appleby RB, Molson J, Barlow RJ, Rafique H, Toader A. Hi-lumi LHC Collimation Studies with MERLIN Code. In: Proc. Of IPAC2014. Dresden, Germany: IEEE (2014). p. 784–7.
39. Catalan-Lasheras N. *Transverse and Longitudinal Beam Collimation in a High-Energy Proton Collider (LHC)*. [PhD thesis]. Zaragoza (1998).

40. Marsili RBA, Redaelli S. Cleaning Performance with 11T Dipoles and Local Dispersion Suppressor Collimation at the LHC. In: Proc. Of IPAC14. Dresden, Germany: IEEE (2014). p. 170–3.
41. Yang J-Q, Zou Y, Tang J-Y. Collimation Method Studies for Next-Generation Hadron Colliders. *Phys Rev Acc Beams* (2019) 22:023002. doi:10.1103/physrevaccelbeams.22.023002
42. Scandale W, Arduini G, Assmann R, Bracco C, Gilardoni S, Ippolito V, et al. First Results on the SPS Beam Collimation with Bent Crystals. *Phys Lett B* (2010) 692:78–82. doi:10.1016/j.physletb.2010.07.023
43. Zvoda V. Advanced Bent crystal Collimation Studies at the Tevatron (T-980). In: Proc. Of PAC2011. New York, USA: IEEE (2011). p. 73–5.
44. Guo Z, Tang J-Y, Yang Z, Wang X-Q, Sun B. A Novel Structure of Multipole Field Magnets and Their Applications in Uniformizing Beam Spot at Target. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2012) 691:97–108. doi:10.1016/j.nima.2012.06.048
45. Zou Y, Tang J, Yang J. Resonant Slow Extraction in Synchrotrons Using Anti-symmetric Sextupole fields. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2016) 830:150–62. doi:10.1016/j.nima.2016.05.081
46. Dietderich DR, Godeke A. Nb3Sn Research and Development in the USA - Wires and Cables. *Cryogenics* (2008) 48:331–40. doi:10.1016/j.cryogenics.2008.05.004
47. Gupta RC. A Common Coil Design for High Field 2-in-1 Accelerator Magnets. In: Proc. Of PAC 1997. Vancouver, Canada: IEEE (1997). p. 3344–6.
48. Lee P. *A Comparison of Superconductor Critical Currents*. Whole wire critical current density values for long-length superconductors (nationalmaglab.org) (2021)
49. Anashin VV. The Vacuum Studies for LHC Beam Screen with Carbon Fiber Cryosorber. In: Proc. Of APAC2004. Gyeongju, Korea: IEEE (2004). p. 329–31.
50. Gröbner O. Vacuum Issues for an LHC Upgrade. In: *1st CARE-HHH-APD Workshop on Beam Dynamics in Future Hadron Colliders and Rapidly Cycling High-Intensity Synchrotrons*. CERN (2004).
51. Collomb-Patton C. Cold Leak Tests of LHC Beam Screens. *Vacuum* (2010) 84:293–7.
52. Baglin V. Synchrotron Radiation Studies of the LHC Dipole Beam Screen with Coldex. In: Proc. Of EPAC 2002. Paris, France: IEEE (2002). p. 2535–7.
53. Krkotić P, Niedermayera U, Boine-Frankenheim O. High-temperature Superconductor Coating for Coupling Impedance Reduction in the FCC-Hh Beam Screen. *Nucl Instru Meth A* (2018) 895:56–61.
54. Gan P, Zhu K, Fu Q, Li H, Lu Y, Easton M, et al. Design Study of an YBCO-Coated Beam Screen for the Super Proton-Proton Collider Bending Magnets. *Rev Scientific Instr* (2018) 89:045114. doi:10.1063/1.5026932
55. Rathjen C. Mechanical Behaviour of Vacuum Chambers and Beam Screens under Quench Conditions in Dipole and Quadrupole fields. In: Proc. Of EPAC2002. Paris, France: IEEE (2002). p. 2580–2.
56. Ostroumov P, Gerigk F. Superconducting Hadron Linacs. In: AW Chao W Chou, editors. *Review of Accelerator Science and Technology*. Singapore: World Scientific Publishing (2013). doi:10.1142/s1793626813300089
57. Abrams RJ. *Interim Design Report, No. CERN-ATS-2011-216*. arXiv: 1112.2853 (2011). p. 216.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Tang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Physics

ORIGINAL RESEARCH
published: 03 March 2022
doi: 10.3389/fphy.2022.838463

Characterization of Low Gain Avalanche Detector Prototypes' Response to Gamma Radiation

Martin Hoeferkamp[1], *Alissa Howard*[2], *Gregor Kramberger*[2], *Sally Seidel*[1]*, *Josef Sorenson*[1] *and Adam Yanez*[1]

[1]Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM, United States, [2]Department of Experimental Particle Physics, Jozef Stefan Institute and Department of Physics, University of Ljubljana, Ljubljana, Slovenia

Motivated by the need for fast timing detectors to withstand up to 2 MGy of ionizing dose at the High Luminosity Large Hadron Collider, prototype low gain avalanche detectors (LGADs) have been fabricated in a single pad configuration, 2 × 2 arrays, and related p-i-n diodes, and exposed to Co-60 sources for study. Devices were fabricated with a range of dopant layer concentrations, and for the arrays, a variety of inter-pad distances and distances from the active area to the edge. Measurements of capacitance *versus* voltage and leakage current *versus* voltage have been made to compare pre- and post-irradiation characteristics in gain layer depletion voltage, full bulk depletion voltage, and breakdown voltage. Conclusions are drawn regarding the effects of the gammas on both surface and interface states and on their contribution to acceptor removal through non-ionizing energy loss from Compton electrons or photoelectrons. Comparison of the performances of members of the set of devices can be used to optimize gain layer parameters.

Keywords: Low Gain Avalanche Detector, LGAD, silicon detector, particle physics, timing detector

OPEN ACCESS

Edited by:
Petra Merkel,
Fermilab (DOE), United States

Reviewed by:
Gabriele Giacomini,
Brookhaven National Laboratory (DOE), United States
Marco Mandurrino,
National Institute of Nuclear Physics of Turin, Italy

***Correspondence:**
Sally Seidel
seidel@unm.edu

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *17 December 2021*
Accepted: *27 January 2022*
Published: *03 March 2022*

Citation:
Hoeferkamp M, Howard A, Kramberger G, Seidel S, Sorenson J and Yanez A (2022) Characterization of Low Gain Avalanche Detector Prototypes' Response to Gamma Radiation. Front. Phys. 10:838463. doi: 10.3389/fphy.2022.838463

INTRODUCTION

The low gain avalanche detector (LGAD) [1–3], based on the planar technology, produces a signal in response to the generation of free carriers by a charged particle or high-energy photon; when operated, it is depleted by a reverse bias. An evolution of the avalanche photodiode (APD), the LGAD exhibits internal signal gain in the range of up to 100 that is proportional to the applied bias voltage. The profile of the LGAD structure is $n^+/p/p^-/p^+$, where the p-implant below the highly doped n^+ cathode electrode is referred to as the multiplication implant, forming a gain layer, and has dopant concentration in the range a few times 10^{16} cm^{-3} and depth 0.8–2.5 μm. The detection of sub-nanosecond signals produced by minimum ionizing particles is possible with these devices [4].

LGADs are proposed for use in several experiments including upgrades to those at the Large Hadron Collider (LHC) [5, 6]. A typical specification for their operation at the LHC, as components of the High-Granularity Timing Detector (HGTD) in ATLAS or the Endcap Timing Layer (ETL) of CMS, includes tolerance to 2 MGy of the ionizing dose, which will accompany integrated hadron fluence up to about 2.5×10^{15} n_{eq}/cm^2 (this includes a safety factor of 1.5). Thin bulk is preferred, as the minimization of induced current variations due to Landau fluctuations will promote the best timing resolution [7].

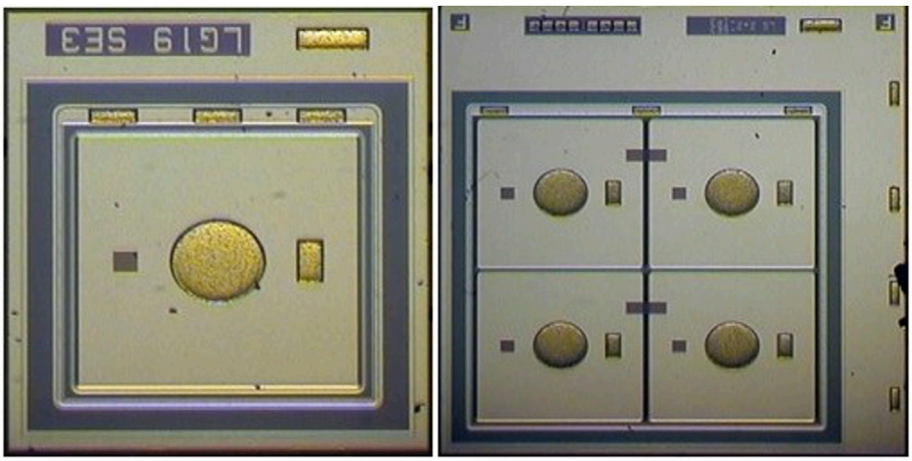


FIGURE 1 | Photographs of (left) a prototype LGAD and (right) a quad sensor prototype.

An active area of research involves the problem of gain decrease as boron-substituted atoms deactivate in response to radiation damage; this is "acceptor removal" [8]. While the primary source of this problem is non-ionizing energy loss (NIEL) due to hadron radiation, a contribution also arises from the associated gammas, which produce point defects in the gain layer through the Compton (and to a lesser extent, photoelectric effect) electrons that they induce.

Gamma radiation motivates a second line of inquiry as well: characterization of oxide charge and interface traps in order to permit optimization of dimensions of the surface features, including interelectrode separation and the distance between the active area and the edge. The goal is to maximize fill factor while ensuring against electrical breakdown under various operating scenarios.

DESCRIPTION OF THE PROTOTYPES

Prototypes of three structures were produced by Hamamatsu Photonics K.K. (HPK) using epitaxial silicon grown on a Czochralski substrate; these are single LGADs, 2 × 2 ("quad") LGAD arrays, and associated p-i-n diodes. All have 50 μm active layer thickness, 200 μm total thickness, and a single guard ring. The pads have dimensions of 1.3 × 1.3 mm^2. The p-type gain layer is approximately 2.5 μm deep and starts very close to the surface. All of the devices include under-bump metallization (UBM). **Figure 1** (left) shows the surface of one such device, which also includes an opening for transient-current technique (TCT) stimulation and a probe needle contact pad. The 2 × 2 arrays have the same features as the smaller devices but variations on inter-pad ("IP") separation (30, 40, 50, and 70 μm) as well as distance from the active area to the edge (300 and 500 μm). **Figure 1** (right) shows a quad prototype. The p-i-n diodes have the same geometry as the LGADs but lack the gain layer. As they can tolerate high bias voltage while sustaining relatively little bulk damage, breakdown in the p-i-ns is indicative of breakdown in the bulk, typically at the guard ring where the field lines are focused. The LGADs and quads were produced with four different options on gain layer dopant concentration. Dopant concentrations of only a few percent difference have previously been shown to lead to very large differences in gain [9].

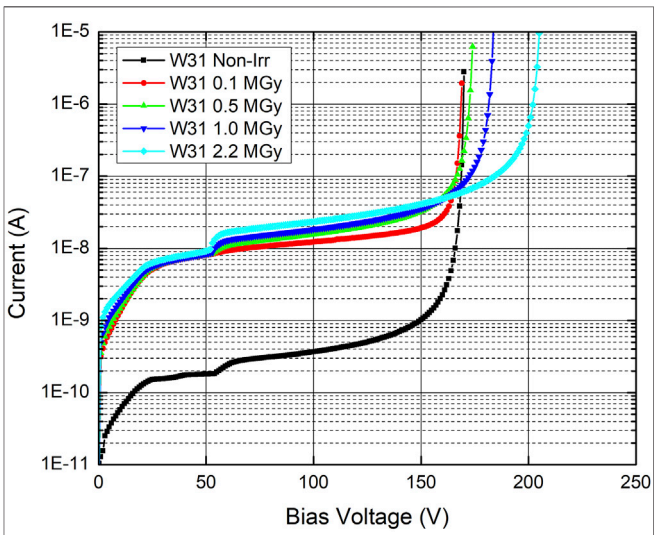


FIGURE 2 | Sample current *versus* voltage characteristics for LGADs from Wafer 31, for various gamma doses.

SINGLE LGAD STUDIES

Measurements

Prototypes were exposed to gammas at the Sandia National Laboratories Gamma Irradiation Facility for total ionizing doses in the range 0.1–2.2 MGy. Measurements of leakage

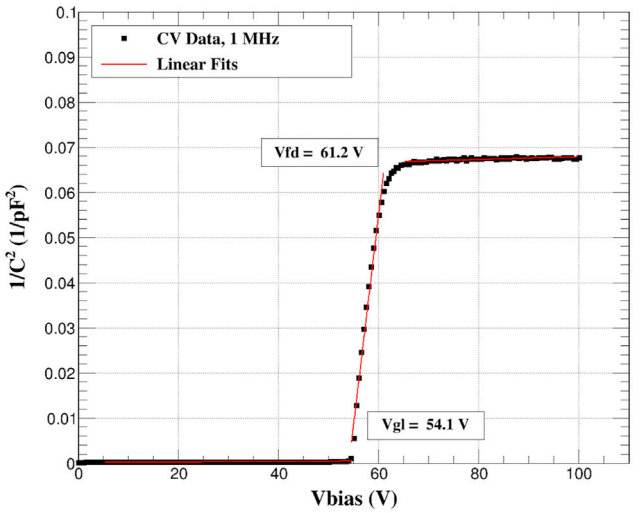


FIGURE 3 | Sample capacitance characteristics *versus* bias voltage, for an unirradiated LGAD from Wafer 31.

current *versus* bias voltage ("IV") and capacitance C *versus* bias voltage ("CV") were carried out before and after the exposure. **Figure 2** shows sample IV curves for a set of devices from Wafer 31; these represent the unirradiated characteristic as well as the response to doses of 0.1, 0.5, 1.0, and 2.2 MGy. The principal features of the curves are representative of all of the wafers studied, although differences were observed for the different gain layer concentrations, and they are discussed below. Noteworthy among the principal features are *1*) the logarithmic rise in leakage current by approximately an order of magnitude during the initial application of bias, due to the surface component; *2*) the foot at approximately 52 V, indicating the depletion of the gain layer; *3*) increase of current (by approximately a factor of 5) which is moderate compared to the increase close to breakdown, which occurs above 160 V; this increase is an indication of the gain; *4*) saturation of the current at the pre-gain layer depletion voltage at 0.1 MGy; and *5*) increase of the breakdown voltage, V_{bd}, with dose, up to about 205 V for the 2.2 MGy sample.

Figure 3 shows an example curve of $1/C^2$ *versus* applied potential V, in this case for an unirradiated LGAD from Wafer 31; data recorded at a temperature of 20°C for applied signal frequency of 1 MHz are shown. The value of the frequency was varied between 1 kHz and 1 MHz, and there was no dependence of the results on the signal frequency provided by the HP4284A LCR meter. The two intercepts of the linear fits to the data in the three regions provide the gain layer depletion voltage V_{gl} and the full bulk depletion voltage V_{fd}. Changes in gain layer depletion have previously been shown to correlate with charge collection performance [10].

Interpretation

Figures 4 and **5** show, respectively, the gain layer depletion voltage V_{gl} and the difference between the full depletion voltage and V_{gl}, which should be proportional to the effective dopant concentration, as a function of total ionizing dose, for single LGAD devices representing all of the gain layer dopant concentrations in wafers with UBM. The data in **Figure 4** are fit to the function $V_{gl} = V_{gl,0}\, e^{-c\phi}$, where ϕ is the total ionizing dose. **Table 1** summarizes the extracted acceptor removal constant (c) values for each wafer, following exposure to 2.2 MGy. Also

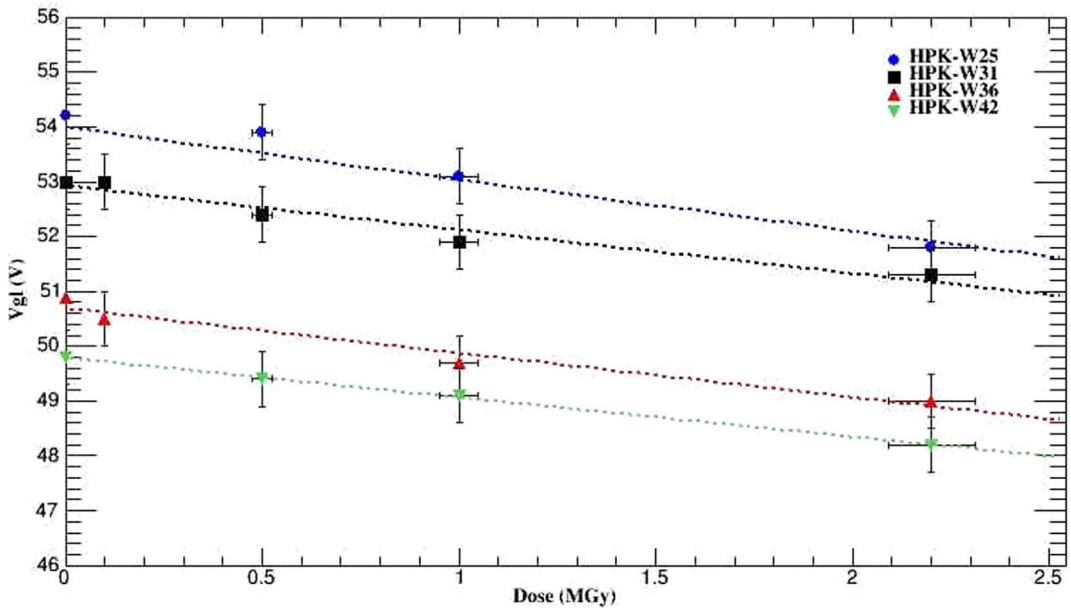


FIGURE 4 | Gain layer depletion voltage as a function of dose, for LGADs from all four wafers. The fits to the function $V_{gl} = V_{gl,0}\, e^{-c\phi}$ are shown, and the resulting values of *c* are reported in **Table 1** as c_γ.

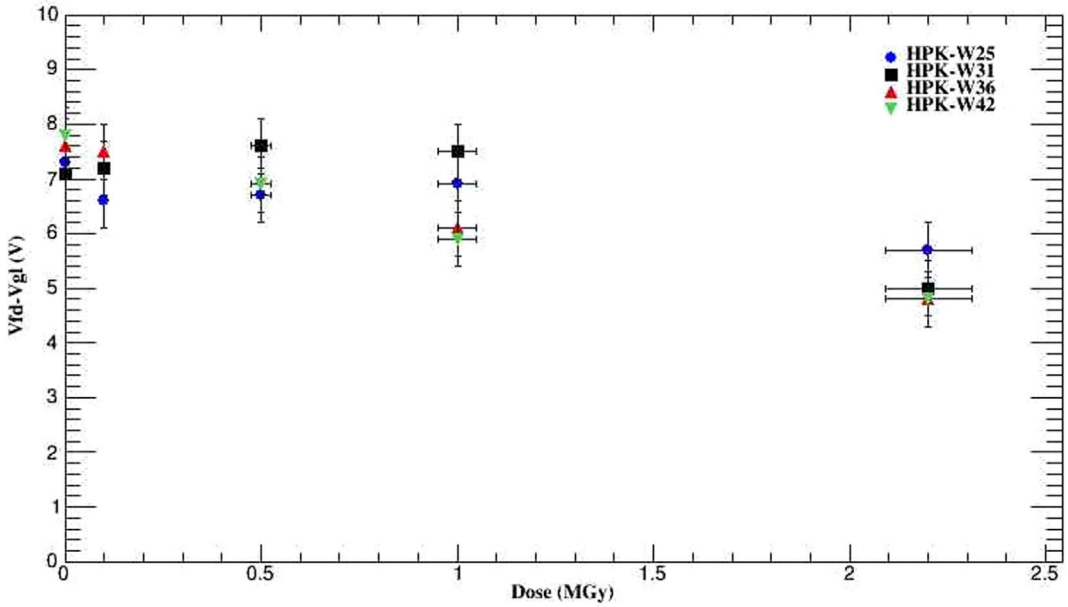


FIGURE 5 | Full bulk depletion voltage minus gain layer depletion voltage, as a function of dose, for LGADs from all four wafers.

TABLE 1 | Properties of the prototype wafers including the acceptor removal constants c_γ of the gain layers, as obtained from a fit of the data in **Figure 4** to the formula $V_{gl} = V_{gl,0}\, e^{-c\phi}$ after the gamma exposure reported here. The initial dopant concentrations of the gain layers (as reflected in the values of $V_{gl,0}$) decrease with increasing wafer number.

Wafer #	$V_{gl,0}$ (V)	Interelectrode separation in the quad sensors (μm)	c_γ [× 10^{-8}/Gy] after exposure to 2.2 MGy
25	54	30, 40, 50, 70	1.79 ± 11.25%
31	53	30, 40, 50, 70	1.53 ± 13.47%
36	51	30	1.62 ± 15.17%
42	50	30	1.47 ± 2.73%

provided there are the pre-irradiation depletion voltages of the gain layers.

The V_{gl} is seen to be only slightly affected by even the highest dose. The full depletion voltage decreases slightly over the same range; this has been observed on epitaxial substrates also for proton-irradiated samples [11]. The observed increase in the breakdown voltage is validated by measurements on the p-i-n diodes (see below). The substantial rise of the surface current below depletion of the gain layer, that is, below the point at which multiplication is possible, can be explained by an increase of the surface recombination velocity. It is not a bulk effect as it does not scale with dose. The fact that it shows signs of saturation above 0.1 MGy is compatible with a surface generation effect. The decrease of V_{gl} in LGADs implies less multiplication. Once the LGAD is fully depleted, further increase of bias voltage adds to the field which can eventually reach breakdown level. The smaller the gain layer depletion voltage, the larger the breakdown voltage is for the device. For gain layer width approximately 2 microns, and active thickness 50 microns, every decrease of V_{gl} by 1 V decreases the breakdown voltage by 25 V.

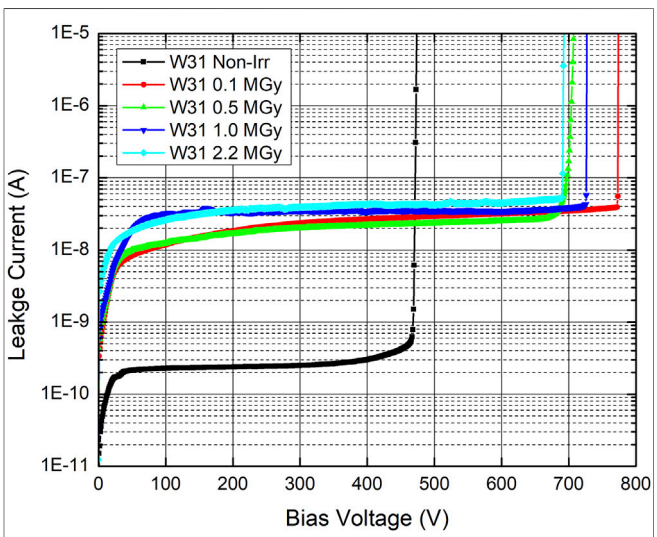


FIGURE 6 | Sample current *versus* voltage characteristics for p-i-n diodes from Wafer 31, for several values of total ionizing dose.

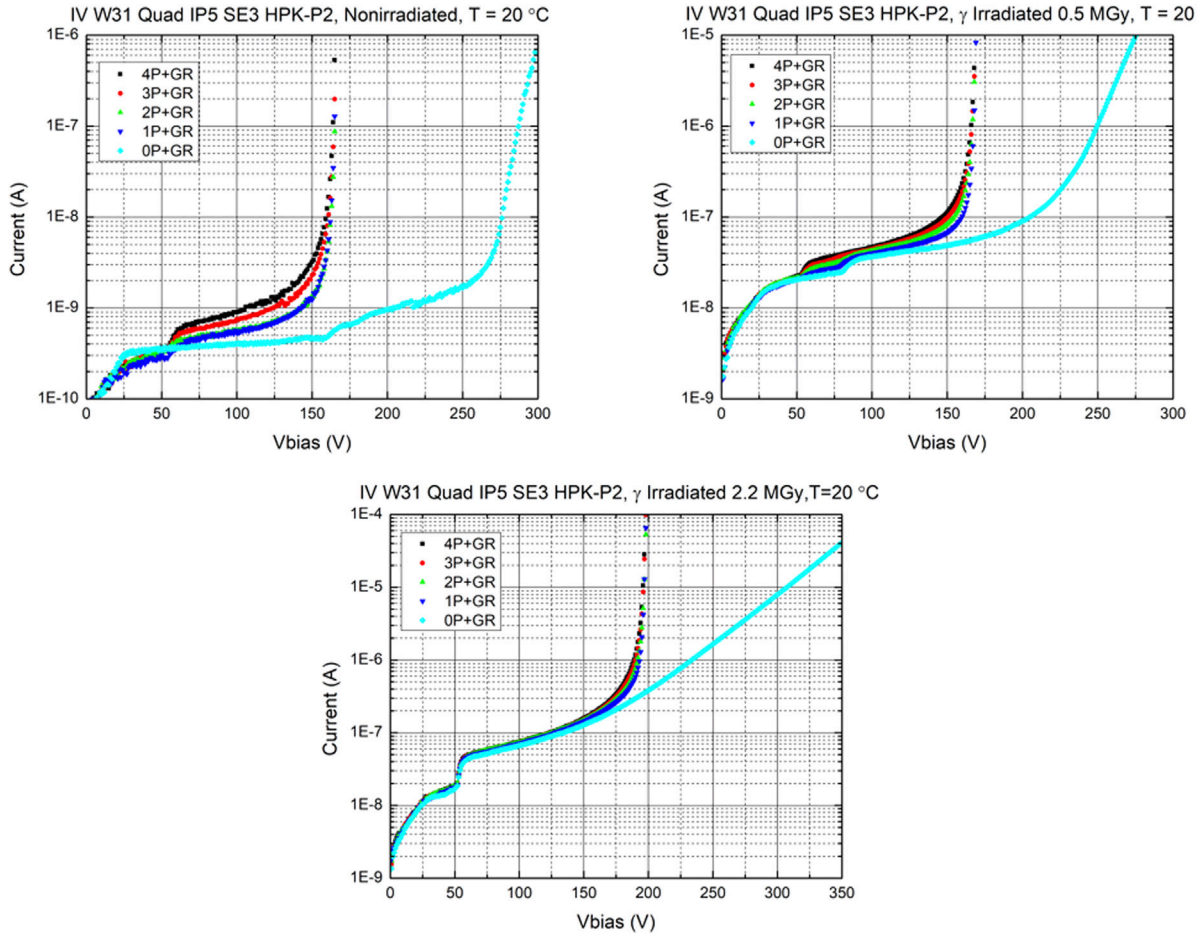


FIGURE 7 | Leakage current *versus* bias voltage for quad sensors taken from Wafer 31, measured (upper left) prior to irradiation, (upper right) after exposure to 0.5 MGy, and (lower) after exposure to 2.2 MGy. In all cases, the temperature during the measurement was 20°C. In each graph, measurements reflect the five modes in which the probes contact the guard ring plus *n* pads, where $n \in \{0,1,2,3,4\}$.

P-I-N DIODE MEASUREMENTS AND INTERPRETATION

Figure 6 shows example IV characteristics for p-i-n diodes from Wafer 25, for total ionizing dose from 0 to 2.2 MGy. In this case, post-irradiation breakdown voltages approaching 800 V are achieved, as irradiation-induced oxide charge moderates the electric field. If the LGADs' primary susceptibility to breakdown were also in the guard ring region, these high values would apply to them as well; however, the LGADs break down at the electrode pads first. These outcomes follow closely the results of measurements made following neutron exposures, reported previously [12]. Between 700 and 800 V, the field in the bulk is sufficiently large that the device breaks down at its weakest point—regardless of whether it is a full LGAD or a p-i-n test structure. For the LGAD, this breakdown is typically through the bulk, whereas for the p-i-n, it is typically at the periphery.

QUAD SENSOR MEASUREMENTS AND INTERPRETATION

If an electrode is floating, its potential is distributed to neighbors by punch-through [13]. This process places a limit on the interelectrode separation, for which the designer must anticipate the consequences in case a lost bump bond leads to breakdown at an electrode, which could then cascade to breakdowns in neighbors.

An IV study involving the quad sensors was carried out to investigate the question of what minimal interelectrode separation will reliably inhibit full punch-through. Bias is applied to the back side of the chip, and leakage current is measured with ground connected to the guard ring plus 0, 1, 2, 3, or all 4 pads. **Figure 7** shows an example set of measurements of this type, for devices taken from Wafer 31, as a function of applied dose.

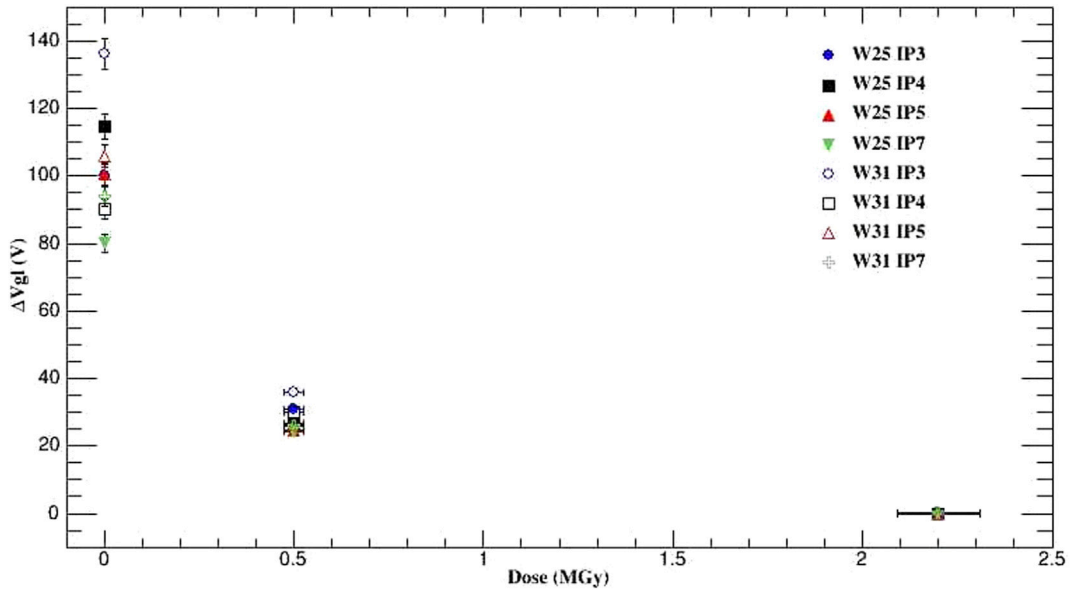


FIGURE 8 | The difference between V_{gl} for the electrodes connected directly to ground, and V_{gl} for electrodes floating, as a function of dose and interelectrode separation. This indicates the voltage needed for punch-through.

Figure 8 shows the punch-through voltage, as a function of dose, for all four interelectrode separations. Punch-through between the guard ring and the pads occurs around 100–140 V prior to irradiation and decreases to nearly 0 V at 2.2 MGy, indicating the loss of resistivity in the region between the pads and the guard ring. At 2.2 MGy, all of the devices' IV curves are similar. In the case of the wafer with 30 μm interelectrode separation, the breakdown voltage for measurement on the guard ring alone was observed to be substantially higher than that in the case in which the guard ring plus any non-zero number of pads are contacted. Among wafers with 30 μm interelectrode separation but differing gains, the voltage at which punch-through occurs increases from approximately 85 V in Wafer 25 to approximately 100 V in Wafer 42.

Figure 9 shows the leakage current *versus* applied bias voltage, for quad devices taken from all wafers (thus with four different initial values of the gain layer depletion voltage). These measurements were made after application of 0.5 MGy, for the measurement configuration indicated above, that is, bias applied to the back side of the chip, and leakage current measured with ground connected to the guard ring plus 0, 1, 2, 3, or all 4 pads. Breakdown occurs at the same bias potential for measurements connecting the guard ring to any number of pads greater than zero. This indicates that the loss of a pad (e.g., disconnection of a bump) will present a danger of breakdown between that pad and its neighbors, for any of the interelectrode separations (30–70 μm) reported here. It is interesting to note that by 2.2 MGy, while the IV curves are identical up to breakdown, the IV measured on the guard ring alone rises with a much slower characteristic.

The resistance between electrodes in the quad sensors was also measured. On each quad sensor, one pad was biased to values in the range 0.5–2.0 V relative to the remaining three grounded pads, and the current drawn on that biased pad from the others was recorded. The back side of the sensor was biased at −100 V with a separate source meter, and the guard ring was allowed to float. A fit to the slope of this IV characteristic yields the inverse of the resistance. **Figure 10** shows the resistance values obtained in this way, for quad sensors representing the four inter-pad spacings, taken from wafers 25 and 31, for doses 0, 0.5, and 2.2 MGy. On all samples, the resistance value is significantly greater than 1 GΩ prior to irradiation; it drops to values around 1 GΩ after application of 0.5 MGy and reaches values in the range 10–100 MΩ after application of 2.2 MGy.

UNCERTAINTIES

The errors on the IV and CV measurements include statistical and systematic uncertainties. Each IV and CV data point shown on the graphs is the average of three to five measurements, and the standard deviation for each is found to be less than 2%. Systematic uncertainties include uncertainties associated with the setup configuration (typically 1.9%), the accuracy of the source and measurement instruments (±0.3% + 100 fA for the Keithley 237; ±0.029% + 300 pA for the Keithley 2410, and ±0.34% for the

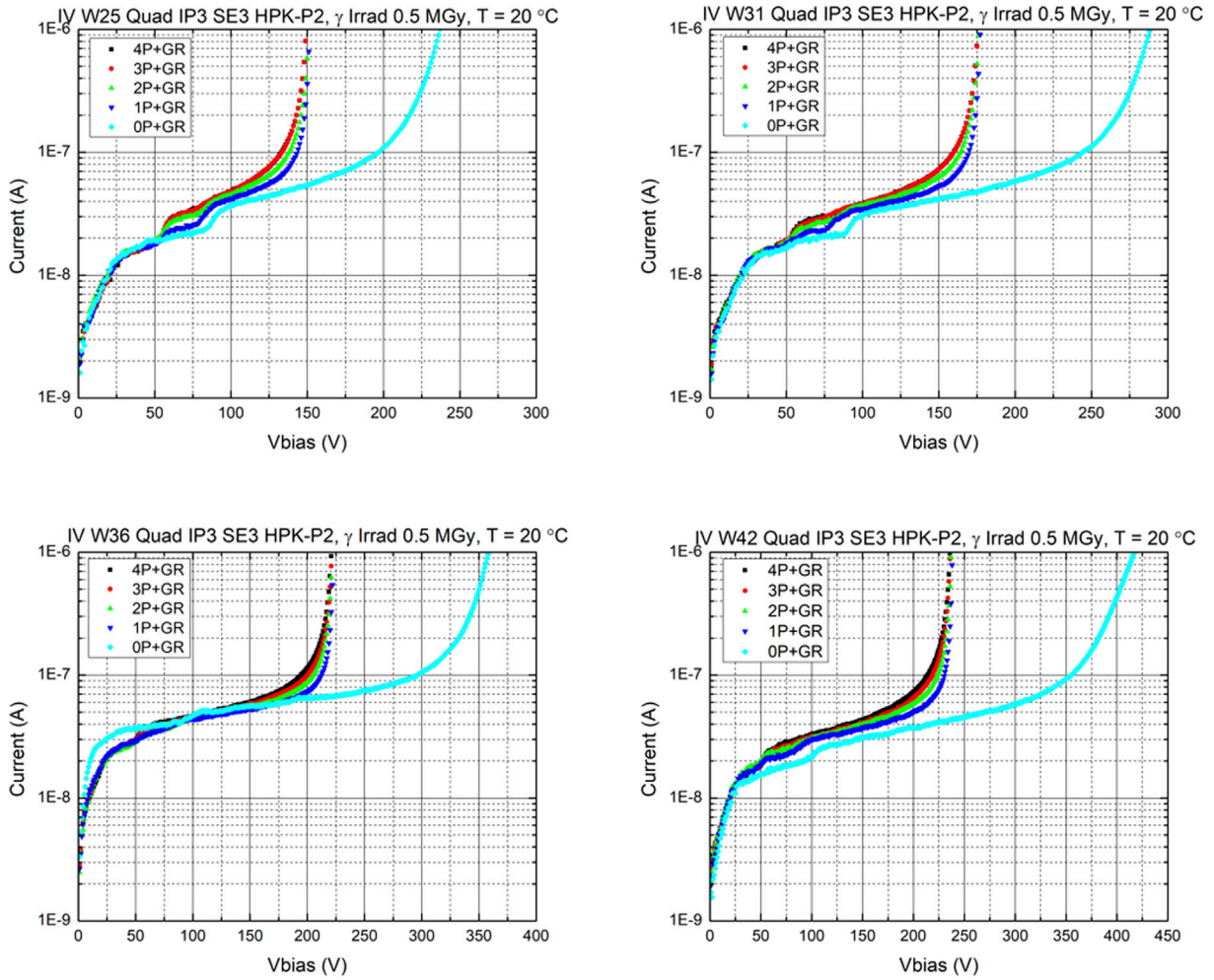


FIGURE 9 | Leakage current *versus* applied bias voltage, for quad devices taken from all wafers (thus with four different initial values of the gain layer depletion voltage). These measurements were made at room temperature after application of 0.5 MGy, for the measurement configuration in which bias is applied to the back side of the chip, and leakage current is measured with ground connected to the guard ring plus 0, 1, 2, 3, or all 4 pads.

HP4284A), the precision of the measurement of the temperature (±0.5°C leads to uncertainty of ±1.82% on leakage current), and the data increment size of 1 V. The uncertainties on V_{gl}, V_{fd}, and pad resistance are derived from the quality of the linear fits and the bias voltage step size of ±1 V. Analyses of the measurements typically require linear fits, on which the uncertainty is typically a small percent.

CONCLUSION

The surface, gain layer, and bulk properties of the LGADs included in this study are found to change after gamma irradiation.

For the single LGADs and the p-i-n diodes, the surface component increases the total leakage current by more than an order of magnitude with only 0.1 MGy dose and saturates at about the same level with a high gamma dose. The single LGADs have a much lower breakdown voltage than the p-i-n diodes, indicating that the LGAD breakdown occurs in the bulk at the electrode pad region. Both the gain layer and full depletion voltages decrease by a small amount even at the highest dose, indicating some damage to the gain layer and bulk. The decrease in ($V_{fd} - V_{gl}$) with dose implies a change in doping concentration in the gain layer and thus acceptor removal. The gamma radiation produced Compton electrons and photoelectrons that led to lattice point defects in the gain layer. The acceptor removal constant was characterized by fitting the data to a decaying exponential function (**Figure 4**), and the resulting values are shown in **Table 1**.

For the 2 × 2 quad LGADs, the punch-through between the guard ring and the pads for all inter-pad separations has been characterized to be over 100 V prior in irradiation; however, after application of gamma irradiation it decreases for all devices studied and reaches nearly 0 V at the maximum 2.2 MGy dose. Pad-to-pad resistance after the maximum

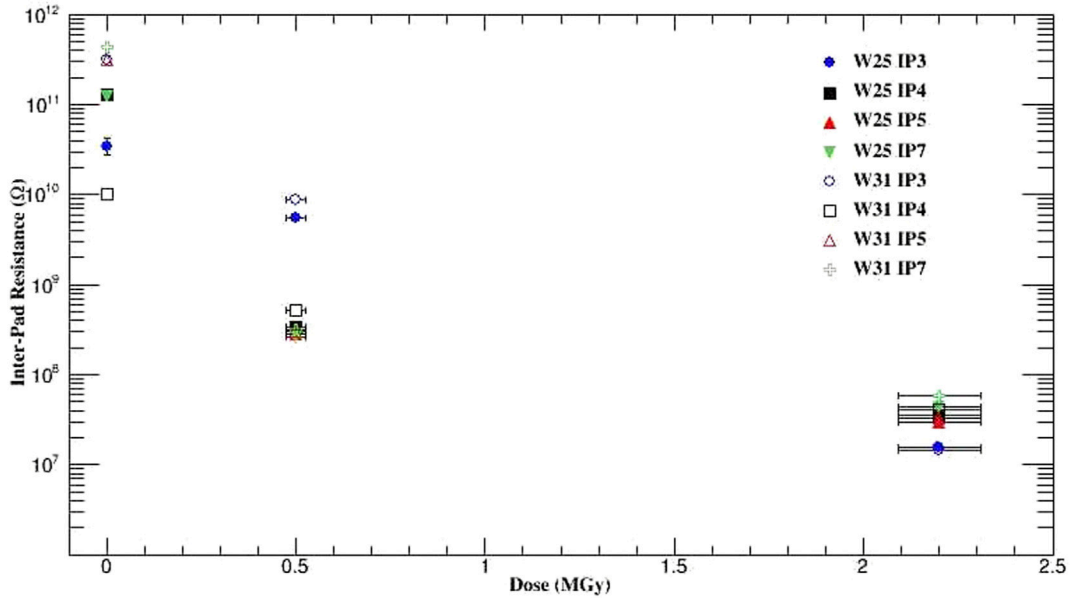


FIGURE 10 | Inter-pad resistance measured on wafers 25 and 31, for all four inter-pad spacings, for doses 0, 0.5, and 2.2 MGy.

dose is found to lie in the range of 10–100 MΩ. Some variations in punch-through voltage and inter-pad resistance are observed for devices from different wafers and different doping concentrations.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

MH, AH, GK, JS, SS, and AY collected and analyzed data. MH and SS carried out irradiations and supervised students' research. GK supervised students' research and initiated the device design. SS composed the manuscript.

FUNDING

This work was made possible by support from the U.S. Department of Energy grant DE-SC0020255 and from ARRS and MIXS, Slovenia. It was also supported by the National Aeronautics and Space Administration (NASA) under Federal award number 80NSSC20M0034.

ACKNOWLEDGMENTS

The opportunity to use the Sandia Gamma Irradiation Facility, made possible by Maryla Wasiolek and Donald Hansen of Sandia National Laboratories, as well as by the University of New Mexico, is gratefully acknowledged. Support from Paulo Oemig of New Mexico State University and the encouragement of Jeremy Perkins and Regina Caputo, both of NASA/GSFC, are deeply appreciated.

REFERENCES

1. Pellegrini G, Fernández-Martínez P, Baselga M, Fleta C, Flores D, Greco V, et al. Technology Developments and First Measurements of Low Gain Avalanche Detectors (LGAD) for High Energy Physics Applications. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2014) 765:12–6. doi:10.1016/j.nima.2014.06.008
2. Cartiglia N. Design Optimization of Ultra-fast Silicon Detectors. *Nucl Instr Meth A* (2015) 796:141.
3. Sadrozinski HF-W. Ultra-fast Silicon Detectors. *Nucl Instr Meth A* (2016) 831:18.
4. Galloway Z, Fadeyev V, Freeman P, Gkougkousis E, Gee C, Gruey B, et al. Properties of HPK UFSD after neutron irradiation up to 6e15 n/cm^2. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2019) 940:19–29. doi:10.1016/j.nima.2019.05.017
5. ATLAS Collaboration. *ATLAS Phase-II Upgrade Scoping Document CERN-LHCC-2015-020* (2015).
6. CMS Collaboration. *CMS Phase-II Upgrade Scoping Document*. CERN-LHCC-2015-9 (2015).
7. Sadrozinski HF-W, Seiden A, Cartiglia N. 4D Tracking with Ultra-fast Silicon Detectors. *Rep Prog Phys* (2018) 81:026101. doi:10.1088/1361-6633/aa94d3
8. Kramberger G, Baselga M, Cindro V, Fernandez-Martinez P, Flores D, Galloway Z, et al. Radiation Effects in Low Gain Avalanche Detectors after Hadron Irradiations. *J Inst* (2015) 10:P07006. doi:10.1088/1748-0221/10/07/p07006
9. Moll M, dissertation Ph.D.. *Radiation Damage in Silicon Particle Detectors: Microscopic Defects and Macroscopic Properties*. Univ. Hamburg (1999).

10. Jin Y, Ren H, Christie S, Galloway Z, Gee C, Labitan C, et al. Experimental Study of Acceptor Removal in UFSD. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2020) 983:164611. doi:10.1016/j.nima.2020.164611
11. Lindstroem G, et al. Epitaxial Silicon Detectors for Particle Tracking - Radiation Tolerance at Extreme Hadron Fluences. *Nucl Instr Meth A* (2006) 568:66–71.
12. ATLAS Collaboration. *Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade*. CERN-LHCC-2020-07 (2020).
13. Lutz G. *Semiconductor Radiation Detectors*. Springer (1999).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Hoeferkamp, Howard, Kramberger, Seidel, Sorenson and Yanez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

REVIEW
published: 25 April 2022
doi: 10.3389/fphy.2022.886473

Electron-Hadron Colliders: EIC, LHeC and FCC-eh

Oliver Brüning[1*†], *Andrei Seryi*[2*†] *and Silvia Verdú-Andrés*[3*†]

[1]European Organization for Nuclear Research, Geneva, Switzerland, [2]Thomas Jefferson National Accelerator Facility, Newport News, VA, United States, [3]Brookhaven National Laboratory, Upton, NY, United States

Electron-hadron colliders are the ultimate tool for high-precision quantum chromodynamics studies and provide the ultimate microscope for probing the internal structure of hadrons. The electron is an ideal probe of the proton structure because it provides the unmatched precision of the electromagnetic interaction, as the virtual photon or vector bosons probe the proton structure in a clean environment, the kinematics of which is uniquely determined by the electron beam and the scattered lepton, or the hadronic final state accounting appropriately for radiation. The Hadron Electron Ring Accelerator HERA (DESY, Hamburg, Germany) was the only electron-hadron collider ever operated (1991–2007) and advanced the knowledge of quantum chromodynamics and the proton structure, with implications for the physics studied in RHIC (BNL, Upton, NY) and the LHC (CERN, Geneva, Switzerland). Recent technological advances in the field of particle accelerators pave the way to realize next-generation electron-hadron colliders that deliver higher luminosity and enable collisions in a much broader range of energies and beam types than HERA. Electron-hadron colliders combine challenges from both electron and hadron machines besides facing their own distinct challenges derived from their intrinsic asymmetry. This review paper will discuss the major features and milestones of HERA and will examine the electron-hadron collider designs of the Electron-Ion Collider (EIC) currently under construction at BNL, the CERN's Large Hadron electron Collider (LHeC), at an advanced stage of design and awaiting approval, and the Future Circular lepton-hadron Collider (FCC-eh).

Keywords: electron, hadron, collider, proton, ion

OPEN ACCESS

Edited by:
Vladimir Shiltsev, Fermi National Accelerator Laboratory (DOE), United States

Reviewed by:
Matthew Wing, University College London, United Kingdom
Allen Caldwell, Max Planck Society, Germany

***Correspondence:**
Oliver Brüning
Oliver.Bruning@cern.ch
Andrei Seryi
seryi@jlab.org
Silvia Verdú-Andrés
sverdu@bnl.gov

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *28 February 2022*
Accepted: *06 April 2022*
Published: *25 April 2022*

Citation:
Brüning O, Seryi A and Verdú-Andrés S (2022) Electron-Hadron Colliders: EIC, LHeC and FCC-eh. Front. Phys. 10:886473. doi: 10.3389/fphy.2022.886473

1 INTRODUCTION

The internal structure of the proton has been a fundamental research topic since the discovery of the proton by Rutherford. First measurements of the proton structure were performed in the 1950s at the Stanford Linear Accelerator (SLAC), yielding a resolution of the proton structure at near femtometer scales and identifying a finite proton radius [1]. Subsequent measurements at SLAC in the 1960s, Fermi National Accelerator Laboratory (FNAL) in the 1970s and the European Organization for Nuclear Research (CERN) in the 1980s improved the resolution of the proton structure by more than two orders of magnitude, revealing the quark structure of the proton and the quark-gluon dynamics inside the proton. The Hadron Electron Ring Accelerator (HERA) at the Deutsches Elektronen-Synchrotron (DESY) laboratory was the first dedicated high center-of-mass energy electron-proton collider and pushed the proton structure resolution down to the attometer scale. The measured proton structure functions from HERA were, in turn, a vital ingredient for the precision measurements at the CERN's Large Hadron Collider (LHC) that led to the discovery of the

Higgs particle in 2012 [2, 3]. Additional insight into the proton structure has been attained at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory (BNL), where high-energy collisions between polarized proton beams enable the study of the contribution of quarks and gluons to the proton spin [4]. Pushing the energy frontier in future hadron collider facilities beyond the TeV scale requires knowledge of the proton structure function at even smaller scales. The Large Hadron Electron Collider (LHeC) and the Future Circular Collider (FCC) could cater to this task. Furthermore, with luminosities above $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at hand, the LHeC could also be a Higgs factory and enable new searches for physics beyond the Standard Model [5]. The HERA collider did not provide polarized proton beams and did not explore the regime of electron-ion collisions. The Electron-Ion Collider (EIC) under construction at BNL will address these aspects in the coming decades and explore QCD aspects not studied by HERA [6, 7], focusing on the central goal of modern nuclear physics – to understand the structure of the proton and neutron directly from the dynamics of their quarks and gluons governed by the theory of their interaction (quantum chromodynamics), and how nuclear interactions between protons and neutrons emerge from these dynamics.

HERA featured asymmetric collisions between beams of different species and energies (most commonly protons at 920 GeV and electrons or positrons at 27 GeV). The operational experience of HERA showed the general feasibility of such an asymmetric collision scheme in circular collider/storage ring configurations and confirmed the necessity – similarly observed by the Super Proton–Antiproton Synchrotron (Sp$\bar{\mathrm{p}}$S) collider at CERN – to match the beam sizes of the colliding beams in order to maintain the proton beam stability and lengthen the proton beam lifetime [8–10]. HERA provided collisions with spin-polarized electron and positron beams, which was essential for electroweak collider physics as well as for the HERMES experiment which studied spin properties in a fixed target mode. The production of polarized electron and positron beams relied on the Solokov-Ternov effect, the mechanism by which high-energy lepton beams become naturally transversely polarized in a storage ring. Spin rotators converted the polarization of the beam from transverse to longitudinal. From a technical point of view, HERA was the first machine that featured, on a large scale, magnets with a cold yoke design [11] and that aimed at minimizing the cycle-dependent persistent current effects of superconducting magnets [12] – a design approach that formed the foundation of the very successful LHC magnet development. HERA was also the first machine to exploit the benefits of superconductivity for both radio-frequency (RF) cavities [13] and magnets [14], and the recipient of the first magnet built using the direct winding technique developed by BNL that has later enabled the manufacturing of low-cost magnets for compact interaction regions (IR) [15].

The EIC will use the RHIC accelerator complex. In operation since 2000, RHIC is the first heavy-ion collider and also the world's only spin-polarized proton collider. A versatile machine, it provides collisions between beams with a wide range of particle species (p ↑, Au, U, . . .) and energies (3.85–100 GeV/u for Au and up to 255 GeV for protons). Transverse and longitudinal stochastic cooling based on microwave technology helps to counteract intra-beam scattering (IBS) of high-energy, bunched ion beams for prolonged luminosity lifetime [16]. The luminosity of low-energy ion collisions is also enhanced thanks to a hadron cooling system that uses RF accelerated electron bunches [17]. Two full Siberian snakes (spin rotators) in each ring enable high luminosity collisions of polarized proton beams at 255 GeV, with a polarization up to 55% averaged over a full store and over the two beams [18]. The EIC will combine the experience from HERA to deliver polarized electron beams with the experience from RHIC to be the first machine that provides the collision of polarized electrons with polarized protons, and at a later stage, polarized ^{2}H and ^{3}He.

The LHeC and FCC-eh collider designs build on the experience of the HERA collider and the recent developments for Energy Recovery Linear accelerators (ERL). The ERL design allows a modular collider design that can be applied to the LHC, the High Energy LHC (HE-LHC) and the FCC hadron storage rings while minimizing the energy requirements for the operation of the collider and providing the maximum performance reach for peak luminosity, promising values in excess of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Both LHeC and FCC-eh implementations are conceived for concurrent operation in parallel to the main program of the hadron colliders and an ensuing dedicated exploitation phase once the main hadron beam physics program has been completed. This paper focuses on the designs of EIC, LHeC and FCC-eh. This is not an exclusive list. Alternatives at a conceptual design stage like the plasma-based particle colliders are looked at in the context of the European Particle Physics Strategy Update (EPPSU) and the associate lab directors group [19], as well as in the ongoing Snowmass 2021 community process [20].

2 DESIGN OPTIONS

An electron-hadron collider has essentially two basic design options: 1) a ring-ring based collider design where both the electron and the hadron beams circulate in opposite directions in storage rings that intersect with each other at one or more locations to enable collisions between the two beams, or 2) a linac-ring based collider design where the hadron beam circulates in a storage ring and collides at one location with electrons from a linear accelerator.

A ring-ring based design is the most efficient collider design as it allows both beams to collide with each other repeatedly. However, its performance is eventually limited by the power loss due to synchrotron radiation of the electron beam, the emittance of the electron beam that is defined by the synchrotron radiation damping and tune shifts caused by the beam-beam interactions at the interaction point. For the LHeC goals, these effects essentially limit the peak performance reach of the collider to luminosities around $10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [21].

Accelerating the electrons through a linear accelerator minimizes the synchrotron radiation losses, but exposes the electron beam only once to a collision with the hadron beam,

thus eliminating the beam-beam tune shift limitation, and allows a better match of the electron beam size to that of the hadron beam as the electron beam emittance is entirely determined by the source. However, a linac-ring design concept allows only one collision of the electron beam with the hadron beam before the electron beams are discarded. The performance reach of this design concept is therefore directly linked to the maximum affordable electron beam power. Again, practical limitations imply for the LHeC concept luminosities around 10^{33} cm^{-2} s^{-1}.

The use of ERLs provides a very appealing alternative design variation of the linac-ring concept. An ERL-based collider design allows, at least in principle, the recuperation of the electron beam power after the collisions with the hadron beam, and thus eliminates the electron beam power limit on the performance. Exposing the electron beam only once to a collision before deceleration and discarding the electron beams after the deceleration further eliminates the limitations of the beam-beam parameter and the electron beam emittances on the performance reach. Assuming an ERL layout with return arcs, where the beam passes through the same RF system during the acceleration and deceleration phases, the total synchrotron radiation power loss is only marginally smaller in an ERL-based design when compared to a ring-ring design of comparable size and bending radius (ca. 70% [21]). However, the elimination of the electron power limit and the beam-beam tune shift limit opens the door to a much higher performance reach and, for the LHeC, opens the door to luminosities above 10^{34} cm^{-2} s^{-1}. This boost in the performance reach comes at the price of relying on a new accelerator concept (ERL) and its challenging operation mode. Any design of a future electron-hadron collider needs therefore to find a good compromise between performance limitations with well-established accelerator concepts and pushing the performance with the help of R&D accelerator concepts.

The above described methodology was fully employed in the development of the United States.based EIC project. While the earlier concepts of the EIC included considerations of a linac-ring concept, where the electron beam was provided by an ERL based on high number of recirculations through a linac and a special Fixed-Field Alternating gradient (FFA) circular beam transport, a comprehensive review [22] found that the "unproven or demanding technical components" of the linac-ring EIC concept with FFAs and large number of recirculations "present both technical and cost risks and will require substantial R&D to be proven reliable and cost effective". Consequently, the more established ring-ring concept was adopted for the United States EIC, which is now an approved project, proceeding to technical design and soon expected to start construction.

3 COMMON CHALLENGES AND ENABLING TECHNOLOGIES

The concept of ERLs was already proposed in the 1960s [23]. A key technological ingredient for the ERL concept is the realization of RF systems with a sufficiently high Q_0 that allow an efficient storage of the electromagnetic fields between deceleration and acceleration cycles. This only became feasible with recent advancements in superconducting RF (SRF) technology that allow accelerating gradients of the order of 20 MV/m with $Q_0 > 10^{10}$ [24–26]. The higher the Q_0 and the lower the losses through Higher Order Modes (HOMs), the more efficient the ERL. Ultimately one substitutes the beam power limitations of a conventional linac with the limitations of the cryogenic system for the ERL that is required for cooling the heat dissipation from the operating mode and HOMs. A second design goal, in addition to the highest possible Q_0, is therefore the development of high-temperature SRF systems that allow a more efficient cooling [e.g., operation at 4 K] and lower cryogenic loads of the SRF system [27, 28]. SRF developments are therefore a critical enabling technology for the ERL-based collider concepts.

A second required enabling technology for ERL-based colliders and cooling systems of future electron-hadron colliders is the development of high average current, high brightness, low emittance electron (and positron) sources with long cathode lifetimes. Record high brightness, high average current non-polarized electron sources use DC guns [29]. Most recently an SRF gun, continuous-wave (CW) operated and equipped with an alkali CsK_2Sb photocathode, provided high-charge electron beams with small transverse emittance and remarkably long lifetimes of up to a month in stable, continued operation [30, 31]. The availability of efficient spin-polarized electron sources is, on the other hand, relevant for electron-hadron colliders like the EIC. GaAs photocathodes are the preferred choice for spin-polarized electron sources [32]. Strained superlattice GaAs cathodes routinely achieve spin polarization above 80% and, in the last years, the use of Distributed Bragg Reflectors (DBR) has boosted their quantum efficiency to a few percentage points [33]. In general, polarized electron sources feature lifetimes of a few days, modest when compared to the non-polarized electron sources.

The introduction of a crossing angle to reduce long-range beam-beam interactions can have a severe impact on the instantaneous luminosity due to the partial geometric overlap of the colliding bunches and the coupling of synchrotron and betatron oscillations [34]. To circumvent this issue, HERA implemented a head-on collision scheme at the expense of adopting a large bunch spacing. The KEK-B electron-positron collider at the High Energy Accelerator Research Organization (KEK) opted instead for a large crossing angle of 22 mrad and demonstrated in the 2000s the full recovery of head-on collisions between electron and positron bunches with crab cavities [35]. Simulations had shown that crab crossing would not only increase luminosity by enabling a full geometric overlap of the colliding bunches but also by its effect on the associated beam-beam tune shift [36]. The High Luminosity LHC (HL-LHC) will be the first collider to use crab cavities for the recovery of head-on collisions between proton bunches [37]. Crabbing of proton bunches was recently demonstrated in the CERN's Super Proton Synchrotron (SPS) [38].

Crab cavities are transversely deflecting RF cavities that provide a transverse kick of the same magnitude but opposite sign to the head and tail of the bunch [39]. Long hadron bunches favor the use of low frequency cavities to avoid a non-linear

crabbing kick [40] that may introduce higher-order synchrobetatron oscillations [41]. Located in the interaction region where typically space constraints are tight, crab cavities are required to be compact and feature a large aperture. High luminosity colliders rely on a high bunch rate that motivates CW operation. Low frequency, large aperture, and CW operation requirements drive the choice of SRF technology for most crab cavities (with the exception of the linear collider CLIC). Compact crab cavity designs have been developed for the HL-LHC crabbing system and the EIC [42–47]. The impedance of the EIC crabbing systems can introduce dangerous transverse instabilities, so a high-gain feedback will reduce the impedance of the crab cavities, to be complemented with a transverse damping system during injection and ramping [48]. The high sensitivity of the hadron beam to RF noise in the crab cavities, in the order of that found for HL-LHC, will also require the implementation of an RF noise feedback [49] as also planned for HL-LHC.

Both HERA and EIC have substantially asymmetric beam energies. To guarantee synchronicity between the colliding beams for a broad range of hadron beam energies outside of the ultra-relativistic regime, at certain energies the EIC hadron beam circulates with a significant radial offset (tens of mm) with respect to the nominal beam pipe size (69 mm diameter) that consequently increases the impedance. In HERA, the minimum energy of the colliding protons (300 GeV) was defined by the maximum orbit change that could be implemented. The different rigidity of the beams also becomes important in the detector solenoid. An uncompensated solenoidal field may compromise the high luminosity as it can introduce closed orbit distortions, transverse beam focusing, transverse coupling, couple longitudinal and transverse planes of crabbed bunches, and impact the polarization of the beams [50, 51]. While the detector solenoid field was uncompensated in HERA, several compensation techniques are planned for the EIC [52, 53].

The typically dense and complex interaction regions of any collider also encounter challenges from the strong synchrotron radiation emitted by leptons and high-energy protons as a result of their passage through separation and focusing magnets. Synchrotron radiation can cause severe heating of equipment and be a direct and indirect source of detector background through backscattered photons and products of the interaction between the beam and photodesorbed molecules. HERA-II had synchrotron radiation absorbers far away from the detector and the low beta quadrupoles to reduce the detector background from backscattered photons [54, 55]. The IR vacuum chambers of the EIC, LHeC and FCC-eh are also a focus of attention [56].

4 EIC

The EIC is a United States DOE project with critical decision 1 (CD-1) approval. The EIC is designed to collide 5–18 GeV polarized electron beams with 41–275 GeV polarized proton beams, polarized light ions with energies up to 166 GeV/u as is the case of ^{3}He, and unpolarized heavy ion beams up to 110 GeV/u [57]. Although the EIC project scope only includes one interaction region, two colliding beam interaction regions and detectors are feasible. The EIC will be built upon the infrastructure of the RHIC complex. Several arcs of the two independent superconducting rings of RHIC will become the hadron storage ring of the EIC. The present injector chain of RHIC will be reused for the EIC. Two electron rings will be added to the tunnel to complete the EIC: a rapid-cycling synchrotron will accelerate the electron beams to storage energy before the beams are injected into the electron storage ring. Strong hadron cooling will be necessary to guarantee the delivery of an electron-proton peak luminosity of 10^{34} cm^{-2} s^{-1} at 105 GeV center-of-mass energy. Construction of the EIC is planned to start in 2024, with first operations beginning early in the next decade. **Figure 1** shows the EIC layout with its main components [58, 59].

Like other high luminosity colliders, the EIC design must address challenges associated with high intensity beams and the attainment of small emittances at the interaction points. The maximum average current for the electron beam circulating in the EIC electron storage ring is limited to 2.5 A for the 5 and 10 GeV electron beams and 0.25 A for the 18 GeV electron beam in order to keep the synchrotron radiation power loss below 10 MW. This power will be restored by seventeen or eighteen 591 MHz single-cell SRF cavities, each equipped with a pair of MW-level CW power couplers [60]. In addition to the challenges derived from high power handling, the variety of beam energies and currents for EIC operation require that these couplers cover a broad range of power levels and couplings. The high luminosity scenario of the EIC will use a 1 A average current proton beam with 10 ns-spaced, 6.9×10^{10} ppb bunches. The EIC will be the electron-hadron collider with smallest bunch spacing. To impede the formation of electron clouds, the vacuum chamber of the EIC hadron storage ring – that of RHIC – will need surfaces with low secondary electron yield (SEY). Arc quadrupoles and sextupoles, which can behave as magnetic bottles for the electron cloud forming after the passage of off-centered beams, present particularly low SEY thresholds close to 1. Electron clouds were effectively suppressed by NEG coating the warm beamline sections of RHIC [61] and the vacuum chamber of the EIC hadron storage ring will as well use NEG coating in the warm beamline sections of the EIC hadron storage ring. The cold sections of RHIC currently have a stainless steel beam pipe, a material with large electric resistivity and high SEY. As part of the conversion of RHIC into the EIC hadron storage ring, the vacuum chamber of the reutilized cold beamline sections will be equipped with a beam screen featuring a copper layer for reduced resistive-wall power losses under a thin amorphous carbon layer for electron cloud suppression [62]. Amorphous carbon is an electron cloud mitigator with SEY close to unity developed over the last decade for use in several proton accelerators of the CERN complex [63].

Acceleration of the hadron beam is currently planned to use two 24.6 MHz quarter-wave resonator (QWR) cavities for capture and acceleration (currently operating at 28 MHz for RHIC, which features a different operation frequency and number of bunches), two 49.3 MHz QWR cavities and two 98.5 MHz QWR cavities for splitting and six existing 197 MHz re-entrant cavities for compression and store, all normal

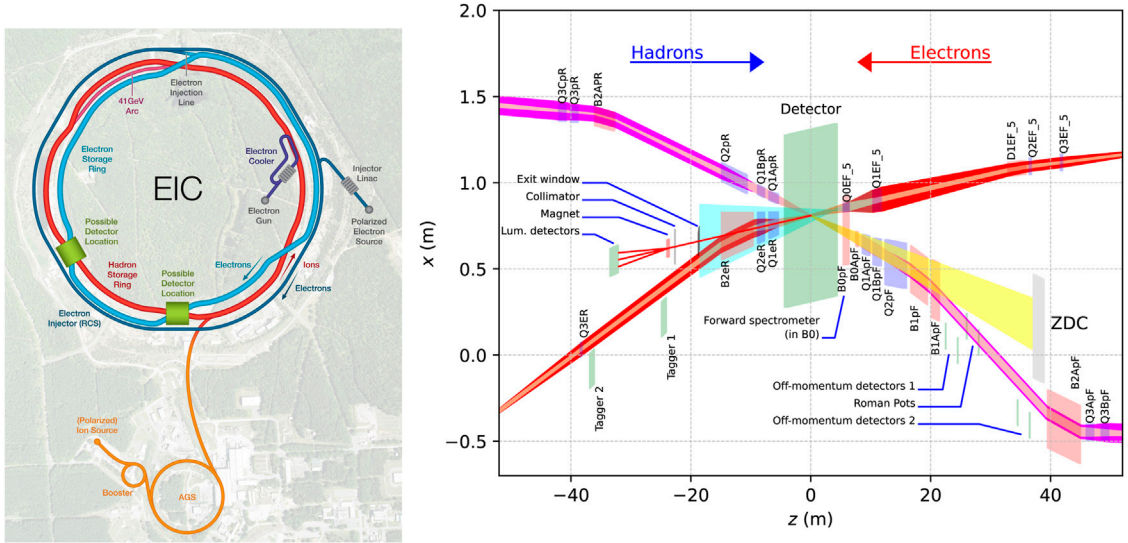


FIGURE 1 | Artistic rendition of the EIC layout and its preliminary interaction region.

conducting, plus one 591 MHz 5-cell SRF cavity [64, 65]. The same 5-cell SRF cavity design is foreseen for acceleration in the RCS (three cavities) and the ERL (ten fundamental frequency cavities and three third harmonic cavities). The 10 MW power lost by the electron beam into synchrotron radiation will be restored by seventeen or eighteen 591 MHz single-cell SRF cavities, to be installed in about nine slots of 8 m length each. While the current design envisions one cavity per cryomodule, putting two cavities per cryomodule is being considered to fit the cavities in the available space, an option also under study for the cooler ERL. The EIC crabbing system follows the local scheme, in which the bunches are crabbed upstream of the interaction point (IP) and un-crabbed downstream of the IP. Crabbing of the electron bunches is realized by one 394 MHz cavity per IP side, each side with about 4 m available. Crabbing of the long hadron bunches uses, for each IP side where up to 15 m are available, four 197 MHz cavities plus two 394 MHz cavities to linearize the crabbing kick along the bunch length. All the crab cavities are SRF radio-frequency dipole (RFD) cavities [66].

The high luminosity of EIC will be achieved by colliding high-charge bunches which are strongly focused at the interaction point. (IBS) in proton bunches will result in longitudinal and horizontal transverse emittance growth time of about 2 h. Strong hadron cooling with a similar cooling time of about 2 h is necessary to attain and preserve the small emittances of the hadron beam and guarantee high luminosity.

Strong hadron cooling (SHC) provides a luminosity increase of a factor 3 to 10 in comparison with performance parameters without cooling. The baseline SHC system for the EIC [67, 68] relies on coherent electron cooling (CeC) based on the micro-bunched electron cooling (MBEC) scheme [69]. In CeC, a dense electron bunch travels together with the proton bunch in a modulator region. Imprints left by protons on the electron beam are amplified by a system of chicanes separated by a quarter plasma wavelength of the electrons, creating density perturbations in the electron beam. The proton bunch lengthens after travel through a chicane, and merges with the electron bunch in a kicker area. Kicks are then applied by the electron beam that are proportional to the average local proton energy offset, resulting in longitudinal cooling. The longitudinal extent of the imprint's "wake" is a few micrometers, resulting in a requirement of sub-micron accuracy and stability of merging of the electron and proton bunches in the kicker region, as well as sub-micron path length stability of their corresponding separate beamlines. Horizontal transverse cooling will be achieved by introducing horizontal dispersion in the SHC area. In addition to SHC, which will operate and keep the emittance stable during collisions, the possibility to pre-cool the injected proton beam vertical emittance at 24 GeV is being evaluated.

To achieve 2 h cooling time, the CeC system needs to operate with a high current beam of 100 mA and up to 150 MeV energy, and electron bunches as long as 14 mm to overlap the hadron bunches. Such an electron beam can only be provided by an ERL. The SHC system uses a single-pass 150 MeV ERL to recover energy from the 100 mA electron beam after it has interacted with the hadron beam. The frequency of the ERL, 591 MHz, is taken after the main RF system of the EIC electron storage ring. No precedent exists for ERL operation with a beam with such high average current and power and with such long, low emittance, high charge bunches. The SHC system of the EIC is possibly the most demanding area reach with accelerator physics challenges and its realization will be an outstanding flagship for the EIC project in particular and for accelerator science and technology in general.

The compact IR design, with challenging space constraints in both transverse and longitudinal planes, will be realized with the use of combined function magnets and assemblies where magnets for the hadron and electron beams share a common iron yoke. To deliver the minimum required aperture defined by the beam size and the synchrotron radiation cone while preserving

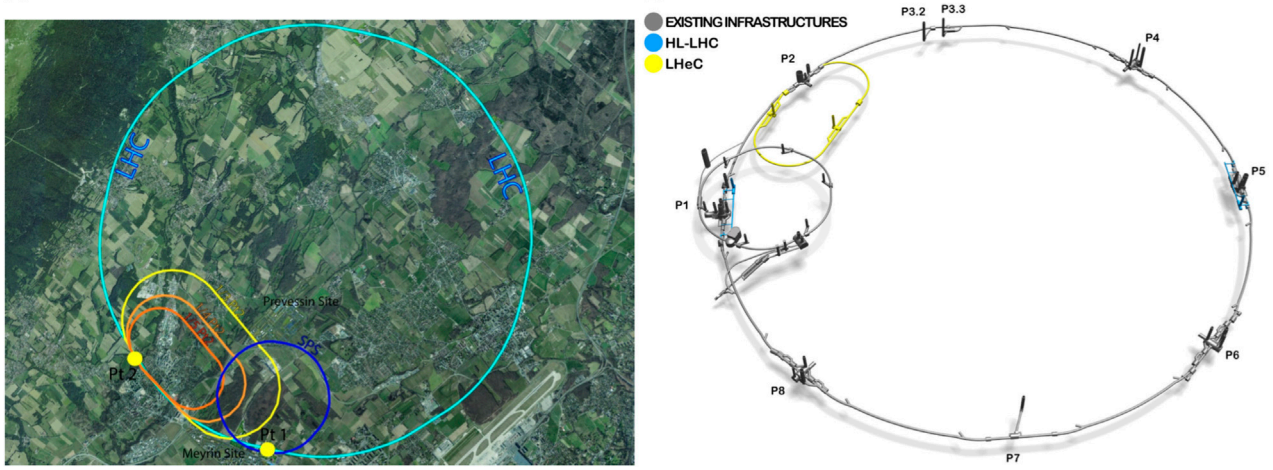


FIGURE 2 | (A) Layout options and footprint of the LHeC in the Geneva basin next to the Geneva airport and CERN. The yellow racetrack corresponds to the LHeC layout that offers optimal performance; in orange, two size variations explored for cost optimization. For reference, the light blue circle depicts the existing tunnel of the LHC; the dark blue circle is the SPS. **(B)** 3D schematic showing the underground tunnel arrangement. The grey sections indicate the existing SPS and LHC tunnel infrastructures and the yellow section the new LHeC installation.

TABLE 1 | The LHeC performance levels during different operation modes.

Parameter	Unit	Run5 period	Run6 period	Dedicated
Brightness $N_p/(\gamma\epsilon_p)$	$10^{17}m^{-1}$	2.2/2.5	2.2/2.5	2.2/2.5
Electron beam current	mA	15	25	50?
Proton β^*	m	0.1	0.07	0.07
Peak Luminosity	10^{34} cm^{-2} s^{-1}	0.5	1.2	2.4
Proton beam lifetime	h	16.7	16.7	100
Fill duration	h	11.7	11.7	21
Turnaround time	h	4	4	3
Overall efficiency	%	54	54	60
Physics time/year	days	160	180	185
Annual integrated luminosity	fb^{-1}	20	50	180

compactness, tapered double-helix or cantered cosine theta (CCT) magnets [70], which rely on the direct-winding technology [15], will be used.

The EIC will be realized over the next decade. RHIC operations will continue until 2025 and the EIC construction will begin soon thereafter. Anticipated EIC completion is planned between 2031 and 2033 [71]. While the SHC hardware and efforts for its initial commissioning are included in the EIC project scope, it is expected that the full performance of SHC and of the EIC will be reached several years after project completion.

5 LHEC

The LHeC proposal aims at maximizing the infrastructure investment of the LHC collider and to expand its physics program by establishing collisions between electrons from a new accelerator infrastructure with one of the hadron beams from the LHC machine [5, 21]. The proposal aims at devising a scheme where electron hadron collisions are generated parasitically to the nominal LHC physics program without changing the baseline operation mode of the LHC. Initial studies looked at both options of a ring-ring based and a linac-ring based implementation and concluded that a linac-ring based collider scheme had the higher performance potential and implied a smaller impact on the nominal LHC infrastructure and operation as compared to a ring-ring based scheme. A concept with an electron accelerator external to the LHC tunnel further provides a modular design concept that can equally and easily be applied to other circular hadron collider options, such as the HE-LHC [72] and the FCC hadron-hadron collider (FCC-hh) [73]. Considering further the energy efficiency of an ERL based linac-ring collider over a conventional linac-ring scheme, the ERL based linac-ring design was adopted as the baseline concept for the LHeC proposal.

The initial LHeC design assumed two superconducting linear accelerators, each being capable of an acceleration of 10 GeV, and three accelerating and three decelerating passages through both linacs for the electron beam. This leads to a racetrack layout of the electron accelerator with a total circumference of ca. 9 km and a

TABLE 2 | Parameter comparison for past and designed electron-hadron colliders. The EIC, LHeC and FCC-eh also include an electron-ion program. Additional parameters can be found in Refs. 5, 8, 58, and 73.

	HERA	EIC	LHeC	FCC-eh
Host site	DESY	BNL	CERN	CERN
Layout	ring-ring	ring-ring	ERL linac-ring	ERL linac-ring
Circumference hadron/lepton (km)	6.3/6.3	3.8/3.8	26.7/[5.3–8.9]	100/[5.3–8.9]
Number of IRs/IPs	4/2	6/1–2	1	1
Max. CM energy (TeV)	0.32	0.14	1.2	3.5
Crossing angle (mrad)	0	22	0	0
Max. peak luminosity (cm^{-2} s^{-1})	5×10^{31}	1×10^{34}	2.3×10^{34}	1.5×10^{34}
Lepton	Electrons, positrons polarized	Electrons polarized	Electrons unpolarized	Electrons unpolarized
Max. average current (A)	0.058	2.5	0.02	0.02
Max. SR power (MW)	7.2	10	45	45
Main RF frequency (MHz)	500	591	802	802
No. main RF cavities/cryomodules	28	17–18/9–18	448/112	448/112
No. crab RF cavities	–	2	–	–
Hadron	Protons unpolarized	Protons polarized	Protons unpolarized	Protons unpolarized
Max. average current (A)	0.163	1.0	1.1	1.1
Main RF frequency (MHz)	208	591	400	400
No. crab RF cavities/cryomodules	–	12/6	8/4	8/4
No. ERL RF cavities	–	13	–	–

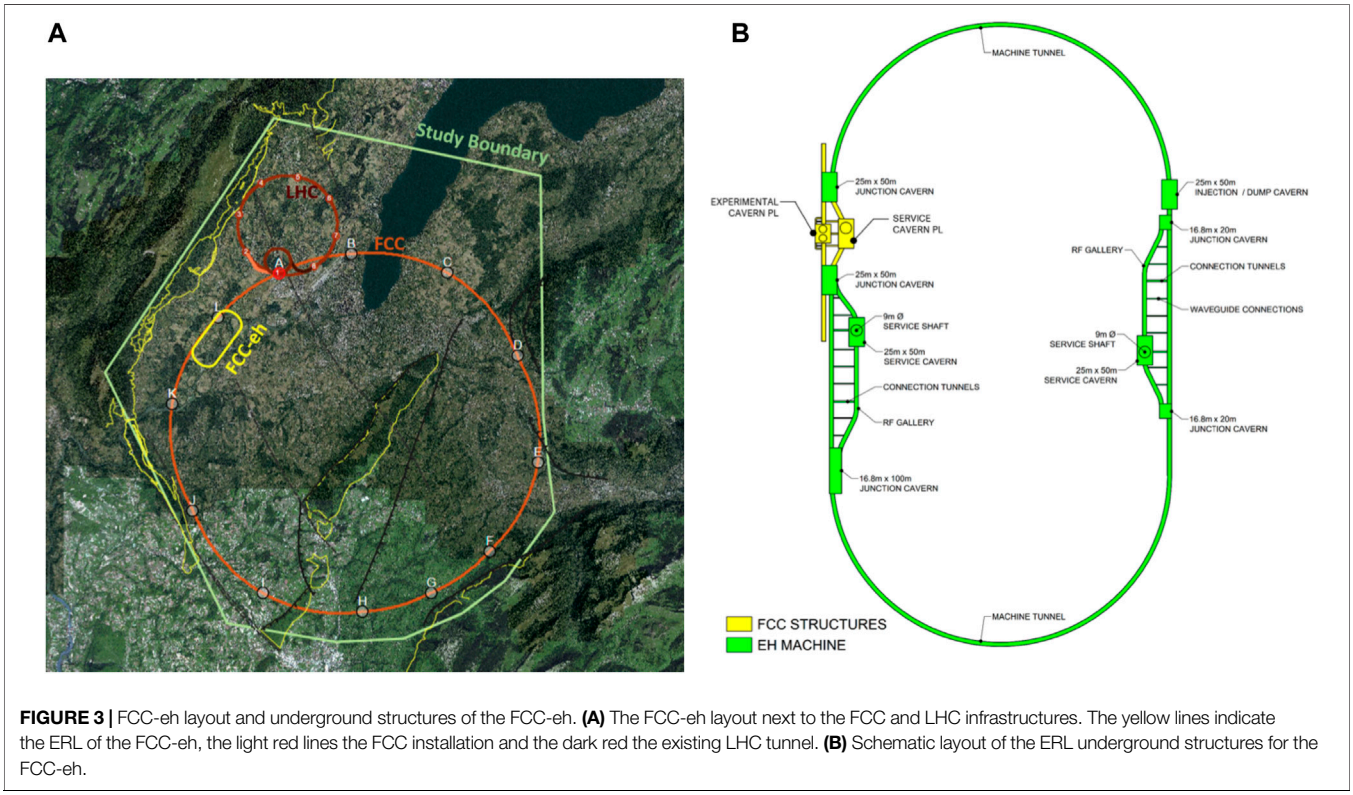


FIGURE 3 | FCC-eh layout and underground structures of the FCC-eh. **(A)** The FCC-eh layout next to the FCC and LHC infrastructures. The yellow lines indicate the ERL of the FCC-eh, the light red lines the FCC installation and the dark red the existing LHC tunnel. **(B)** Schematic layout of the ERL underground structures for the FCC-eh.

maximum electron beam energy of 60 GeV. After the three acceleration passages, the beam is brought into collision with one of the LHC hadron beams before the electron beam enters again the linacs for deceleration. The racetrack shaped electron accelerator can therefore lie tangentially to the existing LHC machine. This layout facilitates the construction of the lepton ERL as the construction can take place largely decoupled from the LHC operation.

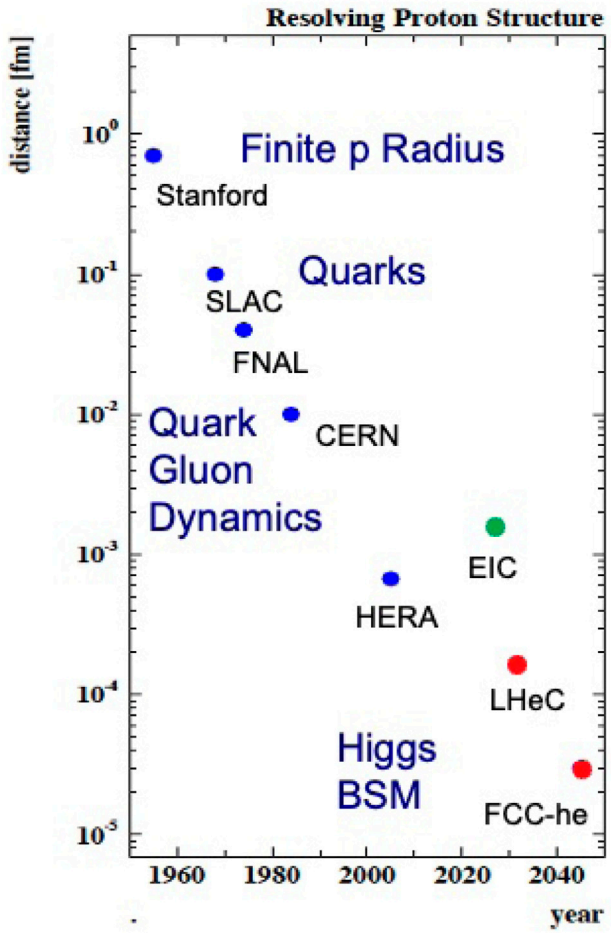


FIGURE 4 | Resolving the proton structure: EIC marked in green with spin resolution and the LHeC and FCC-eh colliders marked in red as potential future colliders [76]. The resolving power is directly related to the maximum $1/Q^2$ achievable at the respective facility. Note: Q^2 is the square of the momentum transferred by the electron to the proton.

Feeding the electron beam into one of the LHC interaction regions and establishing collisions with one of the LHC hadron beams requires the design of novel, asymmetric focusing quadrupole magnets next to the LHeC detector. The focusing quadrupoles have to provide high magnetic fields and sufficient focusing power for the high-energy hadron beams while not affecting the electron beam with its much lower beam energy.

As the LHeC operation is assumed to be performed parasitically on top of the nominal HL-LHC operation, the LHeC does not consider the option of crab cavities acting on the LHC hadron beams. Instead, the head-on collisions in the LHeC are established by integrating a dipole field inside the LHeC detector and gently bending the electron beam onto the trajectory of the hadron beam. Synchrotron radiation originating from the bending of the electron beam onto the LHC hadron beam trajectories poses therefore a challenge for the detector operation and background and needs to be minimized and screened in the LHeC interaction region design. Based on the experience with the HERA operation, the goal is to limit the maximum synchrotron radiation power passing through the LHeC experiment to less then 50 kW. Putting limits on the maximum deflection and bending of the electron beam when entering the interaction region requires the design of novel, asymmetric superconducting focusing quadrupole magnets based on Nb_3Sn technology.

The LHeC design looked at SRF systems based both on the International Linear Collider (ILC) design using 1.3 GHz structures and on the European Spallation Source (ESS) design using 704 MHz structures. Unfortunately these two SRF options do not match to the 40 MHz bunch structure of the LHC hadron beams. The linacs could therefore not use the ILC or ESS SRF cavities as they are, but would require a tuning on the precise RF frequency, which triggered the launch of a new SRF design optimization for the LHeC. Furthermore, beam stability studies showed that an RF frequency of 1.3 GHz would limit the operational current in the RF system and thus the performance reach of the collider. Beam stability and RF power considerations led to the choice of an RF frequency of 802 MHz and first prototype structures produced at the Thomas Jefferson National Accelerator Facility (JLab) exceeded the design criteria in terms of Q_0 and accelerating gradient (18 MV/m with a Q_0 above 3×10^{10}). The chosen SRF frequency is being developed in synergy with the FCC SRF structures.

The initial goal of the LHeC was to provide a beam power in excess of 600 MW at the interaction point with a total wall plug power consumption of 100 MW for the electron beam. Later design considerations aiming on pushing the performance reach beyond a peak luminosity of $10^{34}\,cm^{-2}\,s^{-1}$ and minimizing the total installation cost for the LHeC resulted in shorter linacs, a total circumference of about 5.4 km (1/5th of the LHC circumference with 900 m linac length) but with a slightly higher wall-plug power consumption than the initial 100 MW target. The updated LHeC design features a peak current from the source of 20 mA and total currents within the SRF cavities of more than 120 mA ($2 \times 3 \times 20$ mA) [5]. **Figure 2** shows the potential LHeC size and layout options in relation to the LHC tunnel.

The current LHC planning foresees to extend operation until about 2041 and foresees in total six running periods and five long shutdowns. The nominal LHC operation started in 2010 and extends over three running periods: Run 1 from 2010 until end 2012, Run 2 from 2015 until 2018 and Run three from 2022 until 2025 inclusively. Long Shutdown 1 lasted 2 years from 2013 until 2014 and was used for the consolidation of the inter-magnet splices in order to allow the operation at nominal beam energy of 7 TeV. The second Long Shutdown lasted 3 years and was used for the repair of the diode installation that limited the magnet training after LS1 and the implementation of the LHC Injector Upgrade project. The third LHC run starts in 2022 and is scheduled to extend until 2025 inclusive. A third Long Shutdown extending from 2026 until 2028 will be used for the implementation of the HL-LHC upgrade and the HL-LHC exploitation is assumed to start with Run 4 in 2029. Assuming a 2 year long Long Shutdown 4, the connection of LHeC accelerator complex and the installation of the new LHeC detector could be envisaged during the Long Shutdown 4 in a configuration which may take alternating data on lepton-hadron and on hadron-hadron, as has very recently been shown [74]. For the estimate of the total LHeC performance reach of an integrated luminosity of $1\,ab^{-1}$ it is assumed that the LHeC operates for two runs in parallel with the HL-LHC exploitation followed by one run in a dedicated operation mode where the

operation does no longer have to be assumed to be parasitical to the HL-LHC operation and where the machine performance can be pushed to the maximum value possible. **Table 1** summarizes the parameters and performance reach of the different operation phases for the LHeC [5].

The same machine layout can be used in combination with the HE-LHC, a potential successor project of the HL-LHC that is being studied within the framework of the CERN FCC study. The HE-LHC would use the same tunnel as the LHC and HL-LHC, but would replace the existing superconducting magnets with more powerful high field magnets while using as much as possible the existing LHC infrastructure.

Previous superconducting ERL demonstrators have shown the operability of single-turn, low current ERLs with a maximum beam power of the order of 1 MW. Vital steps for the LHeC development are therefore the demonstration of multi-turn ERL operation with high beam current in the SRF structures and efficient energy recovery and operation efficiency at high beam powers. All three aspects will be addressed with the PERLE ERL demonstrator [75].

6 FCC-EH

The modular design feature of a racetrack shaped ERL that is located tangential to a circular collider, also allows the same design concept (and potentially even the same machine hardware if the projects are staged as separate operation phases) of the LHeC to be used for providing electron-hadron collisions with one of the hadron beams of the FCC-hh [73]. **Figure 3** shows the corresponding machine layout for the FCC-eh implementation. The FCC-eh configuration allows a further push of the center of mass collision energy from the 1.2 TeV at the LHeC to 3.5 TeV when colliding the electron beam from the ERL with one of the hadron beams of the FCC-hh machine. The preferred interaction point for such electron-hadron collisions at the FCC complex is the "L" interaction point close to the main CERN site.

Assuming a staged implementation of the FCC project that starts with a lepton collider in a first phase and a hadron collider in a second phase within the same underground infrastructure could also lead to a scenario where the ERL complex is first used as an injector for the lepton beams for the FCC-ee phase with full energy top-up injection capability for the FCC-ee machine up to the W collision energy, thus sparing the otherwise required booster accelerator within the FCC tunnel. In the second phase of the FCC exploitation with hadron collisions, the ERL complex could then be reconfigured as the electron beam accelerator for electron-hadron collisions during the FCC-hh exploitation phase. Such a staged installation of, first LHeC, followed by an FCC-ee machine with the LHeC ERL [operated in recirculating linac mode without energy recovery] as injector for the lepton beams, followed by an FCC-hh machine with the ERL accelerator providing leptons for lepton-hadron collisions at the FCC represents an interesting scenario for a full exploitation of the infrastructures from the LHC and FCC complex. In addition to providing a unique tool for searching for physics beyond the Standard Model and conducting high precision Higgs studies, the LHeC and FCC-eh provide the finest electron microscope with resolutions of the proton structure down to 10^{-4} fm and 10^{-5} fm, respectively, for the LHeC and the FCC-eh cases, respectively. **Figure 4** shows the potential resolution within reach with the LHeC and FCC-eh colliders in comparison with the achieved resolution in previous experiments. **Table 2** summarizes main design parameters of HERA, EIC, LHeC and FCC-eh.

7 CONCLUSION

The future electron-hadron colliders EIC, LHeC and FCC-eh draw on the experience of the first and only of their kind, HERA, and will build upon RHIC and LHC, two hadron colliders with exceptional versatility and outstanding performances. The urge to produce efficient machines makes the ERL concept an interesting option for future colliders, to the extent that an ERL is part of the hadron cooler system for the EIC and of the acceleration system for the electron beams in LHeC and FCC-eh. In this way the EIC could become the first collider to include an ERL to support regular operations. The PERLE demonstrator will further demonstrate the efficiency and operability of multi-turn high intensity ERLs over the coming years. The quest for high luminosity also requires that future electron-hadron colliders implement crab crossing and guarantee sufficient beam cooling. The solutions implemented to tackle the challenges presented by these high efficiency, high throughput, versatile colliders will be a lodestar to the next generation of particle colliders. A future polarized spin and heavy ion collider, the EIC, and energy frontier electron-hadron colliders, the LHeC and FCC-eh, represent an exciting prospect for novel, luminous deep-inelastic scattering colliders and experiments, to resolve the substructure and dynamics of matter deeper than hitherto and to contribute to the development of particle and nuclear physics with discoveries in the decades ahead.

AUTHOR CONTRIBUTIONS

OB, AS, and SV-A contributed in equal parts to the writing and revision of the manuscript. SV-A edited the last version of the manuscript, which was read and approved by OB and AS.

FUNDING

This study received funding from the Office of Nuclear Physics in the United States Department of Energy (DoE's) Office of Science under contract DE-AC05-06OR23177 with Jefferson Science Associates, LLC and under contract No. DE-SC0012704 with Brookhaven Science Associates, LLC. This study also received funding from CERN for the LHeC and the FCC-eh studies. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

We would like to thank Angelika Drees, Christoph Montag, Kevin Smith, Erdong Wang, Ferdinand Willeke, Holger Witte,

Binping Xiao (BNL), Jiquan Guo, and Bob Rimmer (JLAB) for assistance to gather updated information about the EIC and for fruitful discussions about HERA and the EIC. We are also thankful to Max Klein (CERN and University of Liverpool) who provided details and materials about the LHeC design. We are grateful to Max Klein (CERN and University of Liverpool), Elke-Caroline Aschenauer (BNL), and Todd Satogata (JLab) who kindly reviewed an earlier version of the manuscript and provided useful suggestions to improve the quality and readability of the manuscript.

REFERENCES

1. Hofstadter R. Electron Scattering and Nuclear Structure. *Rev Mod Phys* (1956) 28:214–54. doi:10.1103/revmodphys.28.214
2. Aad G, Abajyan T, Abbott B, Abdallah J, Abdel Khalek S, Abdelalim A, et al. Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC. *Phys Lett B* (2012) 716:1–29. doi:10.1016/j.physletb.2012.08.020
3. Chatrchyan S, Khachatryan V, Sirunyan A, Tumasyan A, Adam W, Aguilo E, et al. Observation of a New Boson at a Mass of 125 GeV with the CMS experiment at the LHC. *Phys Lett B* (2012) 716:30–61.
4. Aschenauer EC, Aidala C, Bazilevsky A, Diehl M, Fatemi R, Gagliardi C, et al. *The RHIC Cold QCD Plan for 2017 to 2023: A Portal to the EIC* (2016). Available at: https://arxiv.org/pdf/1602.03922.pdf (Last Accessed April 4, 2022).
5. Agostini P, Aksakal H, Alekhin S, Allport PP, Andari N, Andre KDJ, et al. The Large Hadron–Electron Collider at the HL-LHC. *J Phys G: Nucl Part Phys* (2021) 48:110501. doi:10.1088/1361-6471/abf3ba
6. Accardi A. Electron Ion Collider: The Next QCD Frontier: Understanding the Glue that Binds Us All. *Eur Phys J A* (2016) 52:268. doi:10.1140/epja/i2016-16268-9
7. National Academies of Sciences. *Engineering, and Medicine. An Assessment of U.S.-Based Electron-Ion Collider Science*. Washington, DC: The National Academies Press (2018). doi:10.17226/25171
8. Voss GA, Wiik BH. The Electron-Proton Collider HERA. *Annu Rev Nucl Part Sci* (1994) 44:413–52. doi:10.1146/annurev.ns.44.120194.002213
9. Brinkmann R, Willeke F. First Experience with Colliding Electron-Proton Beams in HERA. *Proc Int Conf Part Acc* (1993) 5:3742–4. doi:10.1109/PAC.1993.309774
10. Schmüser P, Willeke F. The Electron-Proton Collider HERA. *Elem Particles (Springer) Landolt-Boernstein—Numerical Data Funct Relationships Sci Tech* (2013) 21C:1041–56. New Series.
11. Balewski K, Degele D, Horlitz G, Kaiser H, Lierl H, Mess K-H, et al. Cold Yoke Dipole Magnets for HERA. *IEEE Trans Magn* (1987) 23:1233–5. doi:10.1109/TMAG.1987.1065098
12. Willeke F, Zimmermann F. The Impact of Persistent Current Field Errors on the Stability of the Proton Beam in the HERA Proton Ring. *Conf Rec 1991 IEEE Part Accelerator Conf* (1991) 4:2483–7. doi:10.1109/PAC.1991.165007
13. Dwersteg B, Ebeling W, Moller WD, Proch D, Renken D, Sekutowicz J, et al. Superconducting Cavities for HERA. In: *Proceedings of the 3rd Workshop on RF Superconductivity* (1988). p. 81–94.
14. Wolff S. Superconducting Hera Magnets. *IEEE Trans Magn* (1988) 24:719–22. doi:10.1109/20.11326
15. Parker B, Anerella M, Escallier J, Ghosh A, Jain A, Marone A, et al. Bnl Direct Wind Superconducting Magnets. *IEEE Trans Appl Supercond* (2012) 22:4101604. doi:10.1109/TASC.2011.2175693
16. Blaskiewicz M. Cooling of High-Energy Hadron Beams. *Annu Rev Nucl Part Sci* (2014) 64:299–317. doi:10.1146/annurev-nucl-102313-025427
17. Fedotov AV, Altinbas Z, Belomestnykh S, Ben-Zvi I, Blaskiewicz M, Brennan M, et al. Experimental Demonstration of Hadron Beam Cooling Using Radio-Frequency Accelerated Electron Bunches. *Phys Rev Lett* (2020) 124:084801. doi:10.1103/PhysRevLett.124.084801
18. Fischer W. *Run Overview of the Relativistic Heavy Ion Collider* (2022). https://www.rhichome.bnl.gov/RHIC/Runs/ (Last Accessed December 2, 2022).
19. N Mounet, editor. *European Strategy for Particle Physics - Accelerator RD Roadmap*. Monographs: CERN Yellow Reports (2021). doi:10.23731/CYRM-2022-001
20. Snowmass 2021. *Accelerator Frontier* (2021). Available at: https://snowmass21.org/accelerator/ (Last Accessed April 4, 2022).
21. Abelleira-Fernandez JL, Adolphsen C, Akay A, Aksakal H, Albacete J, Alekhin S, et al. A Large Hadron Electron Collider at CERN. *J Phys G: Nucl Part Phys* (2012) 39:075001. doi:10.1088/0954-3899/39/7/075001
22. Temple E. *NSAC Sub-committee Review of the EIC (Electron Ion Collider) Cost Estimates* (2015). Available at: https://science.osti.gov/-/media/np/nsac/pdf/20150403 (Last Accessed April 4, 2022).
23. Tigner M. A Possible Apparatus for Electron Clashing-Beam Experiments. *Nuovo Cim* (1965) 37:1228–31. doi:10.1007/bf02773204
24. Grassellino A, Romanenko A, Sergatskov D, Melnychuk O, Trenikhina Y, Crawford A, et al. Nitrogen and Argon Doping of Niobium for Superconducting Radio Frequency Cavities: a Pathway to Highly Efficient Accelerating Structures. *Supercond Sci Technol* (2013) 26:102001. doi:10.1088/0953-2048/26/10/102001
25. Grassellino A, Romanenko A, Trenikhina Y, Checchin M, Martinello M, Melnychuk OS, et al. Unprecedented Quality Factors at Accelerating Gradients up to 45 MVm−1in Niobium Superconducting Resonators via Low Temperature Nitrogen Infusion. *Supercond Sci Technol* (2017) 30:094004. doi:10.1088/1361-6668/aa7afe
26. Dhakal P. Nitrogen Doping and Infusion in SRF Cavities: A Review. *Phys Open* (2020) 5:100034. doi:10.1016/j.physo.2020.100034
27. Ganni V. Helium Refrigeration Systems for Super-conducting Accelerators. *AIP Conf Proc* (2015) 1687:040001. doi:10.1063/1.4935327
28. Posen S, Hall DL. Nb3Sn Superconducting Radiofrequency Cavities: Fabrication, Results, Properties, and Prospects. *Supercond Sci Technol* (2017) 30:033004. doi:10.1088/1361-6668/30/3/033004
29. Dunham B, Barley J, Bartnik A, Bazarov I, Cultrera L, Dobbins J, et al. Record High-Average Current from a High-Brightness Photoinjector. *Appl Phys Lett* (2013) 102:034105. doi:10.1063/1.4789395
30. Petrushina I, Litvinenko VN, Jing Y, Ma J, Pinayev I, Shih K, et al. High-brightness Continuous-Wave Electron Beams from Superconducting Radio-Frequency Photoemission Gun. *Phys Rev Lett* (2020) 124:244801. doi:10.1103/physrevlett.124.244801
31. Wang E, Litvinenko VN, Pinayev I, Gaowei M, Skaritka J, Belomestnykh S, et al. Long Lifetime of Bialkali Photocathodes Operating in High Gradient Superconducting Radio Frequency Gun. *Sci Rep* (2021) 11:4477. doi:10.1038/s41598-021-83997-1
32. Maruyama T, Garwin EL, Prepost R, Zapalac GH, Smith JS, Walker JD. Observation of Strain-Enhanced Electron-Spin Polarization in Photoemission from InGaAs. *Phys Rev Lett* (1991) 66:2376–9. doi:10.1103/PhysRevLett.66.2376
33. Liu W, Chen Y, Lu W, Moy A, Poelker M, Stutzman M, et al. Record-level Quantum Efficiency from a High Polarization Strained GaAs/GaAsP Superlattice Photocathode with Distributed Bragg Reflector. *Appl Phys Lett* (2016) 109:252104. doi:10.1063/1.4972180
34. Piwinski A. *Storage Ring Luminosity as a Function of Beam-To-Beam Space and Crossing Angle* (1967).
35. Abe T. Beam Operation with Crab Cavities at KEKB. *Conf Proc C* (2007) 070625:1487. doi:10.1109/PAC.2007.4440798
36. Funakoshi Y. Operational Experience with Crab Cavities at KEKB. In: W Herr G Papotti, editors. *Proceedings of the ICFA Mini-Workshop on Beam–Beam Effects in Hadron Colliders*. CERN–2014–004 (2014). p. 27–36. doi:10.5170/CERN–2014–004.27
37. Calaga R, Tomas R, Zimmermann F. Crab Cavity Option for LHC IR Upgrade. In: W Scandale, T Taylor, F Zimmermann, editors. *3rd CARE-HHH-APD Workshop: Towards a Roadmap for the Upgrade of the LHC and GSI Accelerator Complex: Valencia, Spain 16 - 20 Oct 2006. CERN*. Geneva: CERN (2007). doi:10.5170/CERN-2007-002
38. Calaga R, Alekou A, Antoniou F, Appleby RB, Arnaudon L, Artoos K, et al. First Demonstration of the Use of Crab Cavities on Hadron Beams. *Phys*

Rev Accel Beams (2021) 24:062001. doi:10.1103/PhysRevAccelBeams.24.062001
39. Palmer RB. Energy Scaling, Crab Crossing, and the Pair Problem. *eConf* (1988) C8806271:613–9.
40. Sun Y-P, Assmann R, Barranco J, Tomás R, Weiler T, Zimmermann F, et al. Beam Dynamics Aspects of Crab Cavities in the CERN Large Hadron Collider. *Phys Rev ST Accel Beams* (2009) 12:101002. doi:10.1103/PhysRevSTAB.12.101002
41. Xu D, Hao Y, Luo Y, Qiang J. Synchrobetatron Resonance of Crab Crossing Scheme with Large Crossing Angle and Finite bunch Length. *Phys Rev Accel Beams* (2021) 24:041002. doi:10.1103/PhysRevAccelBeams.24.041002
42. Calaga R, De Maria R, Metral E, Sun Y, Tomas R, Zimmermann F. Crab Cavities. In: *Proceedings of the Chamonix 2010 Workshop on LHC Performance* (2010).
43. Xiao B, Alberty L, Belomestnykh S, Ben-Zvi I, Calaga R, Cullen C, et al. Design, Prototyping, and Testing of a Compact Superconducting Double Quarter Wave Crab Cavity. *Phys Rev ST Accel Beams* (2015) 18:041004. doi:10.1103/PhysRevSTAB.18.041004
44. Verdú-Andrés S, Artoos K, Belomestnykh S, Ben-Zvi I, Boulware C, Burt G, et al. Design and Vertical Tests of Double-Quarter Wave Cavity Prototypes for the High-Luminosity LHC Crab Cavity System. *Phys Rev Accel Beams* (2018) 21:082002. doi:10.1103/PhysRevAccelBeams.21.082002
45. De Silva SU, Delayen JR. Design Evolution and Properties of Superconducting Parallel-Bar Rf-Dipole Deflecting and Crabbing Cavities. *Phys Rev ST Accel Beams* (2013) 16:012004. doi:10.1103/PhysRevSTAB.16.012004
46. De Silva SU, Delayen JR. Cryogenic Test of a Proof-Of-Principle Superconducting Rf-Dipole Deflecting and Crabbing Cavity. *Phys Rev ST Accel Beams* (2013) 16:082001. doi:10.1103/PhysRevSTAB.16.082001
47. Hall B, Burt G, Apsimon R, Lingwood CJ, Tutte A, Grudiev A, et al. Design and Testing of a Four Rod Crab Cavity for High Luminosity LHC. *Phys Rev Accel Beams* (2017) 20:012001. doi:10.1103/PhysRevAccelBeams.20.012001
48. Blaskiewicz M. *Instabilities Driven by the Fundamental Crabbing Mode. Tech. Rep*. Upton, NY, United States: Brookhaven National Laboratory (2021). doi:10.2172/1825736Instabilities Driven by the Fundamental Crabbing Mode
49. Smith K, Mastoridis T, Fuller P, Mahvi P, Matsumara Y. *EIC Transverse Emittance Growth Due to Crab Cavity RF Noise: Estimates and Mitigation. Tech. Rep*. Upton, NY, United States: Brookhaven National Laboratory (2022).
50. Peggs SG. The Projection Approach to Solenoid Compensation. *Part Accel* (1982) 12:219–29.
51. Wei G, Morozov V, Zhang Y, Pilat FC. Evaluation and Compensation of Detector Solenoid Effects in the JLEIC. In: *Proc. 7th Int. Particle Accelerator Conf. (IPAC'16)*. Busan, Korea (2016). p. 2454–6. doi:10.18429/JACoW-IPAC2016-WEPMW015
52. Gamage BR, Michalski TJ, Morozov VS, Rajput-Ghoshal R, Seryi A, Wittmer W, et al. Detector Solenoid Compensation for the Electron-Ion Collider. *Proc 12th Int. Part Acc. Conf.* (2021). doi:10.18429/JACoW-IPAC2021-TUPAB041
53. Berg JS. *First IR DesignEIC Workshop – Accelerator Partnerships* (2021). Available at: https://meetings.triumf.ca/event/254/ (Last Accessed April 4, 2022).
54. Willeke F. The HERA Lepton-Proton Collider. *chap* (2016) 15:225–42. doi:10.1142/9789814436403_0015
55. Schneekloth U. The HERA Ep Interaction Regions Learned Lessons. *Electron-Ion Collider Workshop* (2008). https://web.mit.edu/eicc/Hampton08/ (Last Accessed April 4, 2022).
56. Stutzman M. *Interaction Region, Synchrotron Radiation and Vacuum EIC Workshop – Accelerator Partnerships* (2021). Available at: https://meetings.triumf.ca/event/254/ (Last Accessed April 4, 2022).
57. ECCE Physics Wiki. *ECCE Physics Wiki* (2021). (last edited on 30 October 2021).
58. Willeke F, Beebe-Wang J. Upton, NY, United States: Brookhaven National Laboratory (2021). doi:10.2172/1765663Electron Ion Collider Conceptual Design Report 2021*Tech Rep*
59. Montag C. Design Status Update of the Electron-Ion Collider. JACoW. *Proc 12th Int Part. Acc Conf (Jacow)* (2021) 2585–8. doi:10.18429/JACoW-IPAC2021-WEPAB005
60. Xu W, Fite J, Holmes D, Conway Z, Smith K, Zaltsman A. High Power Coupler Development for EIC. In: *Proc. 2021 International Conference on RF Superconductivity (SRF'21)* (2021). doi:10.2172/1809066
61. Fischer W, Blaskiewicz M, Brennan M, Huang H, Hseuh H-C, Ptitsyn V, et al. Electron Cloud Observations and Cures in RHIC. In: *Proc. 22nd Particle Accelerator Conference (PAC)*. Albuquerque, New Mexico: JACoW (2007). p. 759–63. doi:10.1109/PAC.2007.4441092
62. Verdú-Andrés S, Brennan J, Blaskiewicz X, Mand G, Gupta R, Hershcovitch A, et al. A Beam Screen to Prepare the RHIC Vacuum Chamber for EIC Hadron Beams: Conceptual Design and Requirements. JACoW. In: *Proc. 12th Int. Particle Acc. Conf*. São Paulo, Brazil: JACoW (2021). p. 2066–9. doi:10.18429/JACoW-IPAC2021-TUPAB260
63. Yin Vallgren C, Arduini G, Bauche J, Calatroni S, Chiggiato P, Cornelis K, et al. Amorphous Carbon Coatings for the Mitigation of Electron Cloud in the CERN Super Proton Synchrotron. *Phys Rev ST Accel Beams* (2011) 14:071001. doi:10.1103/PhysRevSTAB.14.071001
64. Rimmer R, Preble J, Smith K, Zaltsman A. HOM Damper Design for BNL EIC 197 MHz Crab Cavity. JACoW. In: *Proc. 12th Int. Particle Acc. Conf*. São Paulo, Brazil: JACoW (2021). p. 1179–81.
65. Preble J. EIC RF Systems Overview. *EIC Workshop – Accelerator Partnerships* (2021). Available at: https://meetings.triumf.ca/event/254/.
66. Xiao B, Delayen J, De Silva SU, Li Z, Rimmer R, Verdú-Andrés S, et al. HOM Damper Design for BNL EIC 197 MHz Crab Cavity. JACoW. In: *Proc. 2021 International Conference on RF Superconductivity*. Grand Rapids, MI, United States: JACoW (2021).
67. Stupakov G. Micro-bunched Electron Cooling (MBEC) for Future Electron-Ion Colliders. In *61st ICFA ABDW on High-Intensity and High-Brightness Hadron Beams*. Daejeon, Korea: JACoW (2018). p. 379–81.
68. Wang E. The Accelerator Design Progress for EIC strong Hadron Cooling. JACoW. In: *12th Int. Particle Acc. Conf*. São Paulo, Brazil: JACoW (2021). p. 1424–7.
69. Ratner D. Microbunched Electron Cooling for High-Energy Hadron Beams. *Phys Rev Lett* (2013) 111:084802. doi:10.1103/PhysRevLett.111.084802
70. Witte H, Parker B, Palmer R. Design of a Tapered Final Focusing Magnet for eRHIC. *IEEE Trans Appl Supercond* (2019) 29:1–5. doi:10.1109/TASC.2019.2902982
71. Hallman T. DOE NP Perspectives on the Electron-Ion Collider (EIC) (2021). EIC Workshop – Accelerator Partnerships. Available at: https://meetings.triumf.ca/event/254/ (Last Accessed April 4, 2022).
72. Abada A, Abbrescia M, AbdusSalam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. HE-LHC: The High-Energy Large Hadron Collider. *Eur Phys J Spec Top* (2019) 228:1109–382. doi:10.1140/epjst/e2019-900088-6
73. Abada A, Abbrescia M, AbdusSalam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. FCC-hh: The Hadron Collider. *Eur Phys J Spec Top* (2019) 228:755–1107. doi:10.1140/epjst/e2019-900087-0
74. André KDJ, Bella LA, Armesto N, Bogacz SA, Britzger D, Brüning OS, et al. An experiment for Electron-Hadron Scattering at the LHC. *Eur Phys J C* (2022) 82:40. doi:10.1140/epjc/s10052-021-09967-z
75. Klein M, Stocchi A. *PERLE: A High Power Energy Recovery Facility for Europe. A Contribution to the Update of the European Strategy on Particle Physics* (2018).
76. Klein M. *Private Communication* (2022).

Conflict of Interest: AS is employed by Jefferson Science Associates, LLC and SV-A is employed by Brookhaven Science Associates, LLC.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Brüning, Seryi and Verdú-Andrés. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

MINI REVIEW
published: 23 May 2022
doi: 10.3389/fphy.2022.888395

Accelerator Technology and Beam Physics of Future Colliders

Frank Zimmermann *

CERN, Geneva, Switzerland

We review key challenges for future next and next-next (-next) generation particle colliders and possible technological paths to address them.

Keywords: hadron collider, lepton collider, accelerator R&D, future circular collider, gamma factory, muon collider, linear collider, synchrotron radiation

1 FUTURE COLLIDER LANDSCAPE

High-energy physics calls for particle colliders with much higher energy and/or luminosity than any past or existing machine. Various types of future particle colliders are being proposed and under development.

Technically closest to construction are the International Linear Collider (ILC) in Japan, the Future Circular electron-positron Collider (FCC-ee) in Europe, and the Circular Electron Positron Collider (CEPC) in China. The ILC is a refinement of the former TESLA collider design, with 1.3 GHz superconducting radiofrequency cavities as underpinning technology, and, as such, it is grounded in more than 30 years of dedicated and successful R&D efforts. Another type of linear collider, CLIC, is based on higher-gradient normalconducting RF cavities, and powered with a novel two-beam acceleration scheme. The two circular collider designs, FCC-ee and CEPC, build on 60 years of experience with operating colliding-beam storage rings, and in particular, they include ingredients of the former LEP collider at CERN, and of the KEKB, PEP-II and SuperKEKB B factories. Combining successful concepts and introducing a few new ones allows for an enormous jump in performance. For example, FCC-ee, when running on the Z pole is expected to deliver more than 100,000 times the luminosity of the former LEP collider. The circular lepton colliders FCC-ee and CEPC would be succeeded by energy frontier hadron colliders, FCC-hh and SPPC, respectively, providing proton collisions at a centre-of-mass energy of about 100 TeV or higher.

Several colliders based on energy-recovery linacs (ERLs) also are under discussion. A Large Hadron electron Collider [1], with an electron beam from a dedicated ERL, could extend the physics programme at the LHC. Recently, high-energy, high-luminosity ERL-based versions of the FCC-ee [2] and of the ILC [3] have been proposed.

The above proposals are complemented with still others, presumably in the farther future, such as photon colliders, muon colliders, or colliders based on plasma acceleration.

Aside performance, technical feasibility, affordability, and sustainability are further questions which the collider designers may need to address.

Five major challenges are driving the design and, ultimately, the feasibility of future high-energy colliders. These are: 1) synchrotron radiation, 2) the bending magnetic field, 3) the accelerating gradient, 4) the production of rare or unstable particles (positrons or muons), and 5) cost and sustainability.

2 SYNCHROTRON RADIATION

A charged particle deflected transversely to its velocity vector emits electromagnetic radiation which, if caused by the influence of an external magnetic field, is called synchrotron radiation. Denoting the

OPEN ACCESS

Edited by:
Alessandro Tricoli, Brookhaven National Laboratory (DOE), United States

Reviewed by:
Alexander Valishev, Fermi National Accelerator Laboratory (DOE), United States
Wolfram Fischer, Brookhaven National Laboratory (DOE), United States

****Correspondence:***
Frank Zimmermann
frank.zimmermann@cern.ch

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *02 March 2022*
Accepted: *05 April 2022*
Published: *23 May 2022*

Citation:
Zimmermann F (2022) Accelerator Technology and Beam Physics of Future Colliders. Front. Phys. 10:888395. doi: 10.3389/fphy.2022.888395

charge of the particle by e, its relativistic Lorentz factor by γ, and considering a particle that follows a circular orbit of bending radius ρ, the energy loss per turn is given by

$$U_0 = \frac{e^2}{3\epsilon_0}\frac{\gamma^4}{\rho}. \quad (1)$$

If there is not a single particle but a beam with current I_{beam}, the power of the emitted synchrotron radiation becomes

$$P_{\mathrm{SR}} = \frac{I_{\mathrm{beam}}}{e}U_0. \quad (2)$$

To provide some examples, the maximum synchrotron radiation power at the former Large Electron Positron collider (LEP) was about 23 MW, while for the proposed future circular electron-positron collider FCC-ee a total constant value of 100 MW has been adopted as a design constraint.

For the same particle energy, the Lorentz factor of protons is much (about 2000 times) lower than for electrons. Consequently, until now, synchrotron radiation power for proton beams has been much less significant, even if not fully negligible. For the Large Hadron Collider (LHC), it amounts to about 10 kW. However, this value increases to a noticeable 5 MW for the proposed future circular hadron collider FCC-hh. Removal of this heat, from inside the cold magnets of the collider arcs, requires more than 100 MW of electrical cryoplant power. These numbers reveal that for both future electron-positron and hadron colliders, synchrotron radiation alone implies more than 100 MW of electrical power needs.

Possible mitigation measures to limit or suppress the synchrotron radiation include:

- increasing the bending radius ρ, which translates into a large(r) circular collider, and is a key part of the FCC concept;
- the construction of a linear collider, which features only minor arcs, but still faces the issues of radiation in the final quadrupole magnets (Oide effect) and in collision (beamstrahlung)—see below;
- the construction of a muon collider;
- miniaturizing the beam vacuum chamber of a large ring; and
- shaping the beam to suppress radiation.

We will now look at these five possibilities in greater detail.

2.1 Size of Circular Colliders

The construction cost of different collider elements increases or decreases with the size of the ring. The optimum size is a function of the maximum beam energy. In 1976, B. Richter performed a cost optimisation of circular electron-positron colliders [4]. For a maximum c. m. energy of about 365 GeV (top quark production), he found that a collider diameter of 100 km is close to the optimum. A similar circumference value of about 90 km is obtained when extrapolating from the size and energy of more recent machines (PETRA, TRISTAN and LEP) [5].

Serendipitously, a circumference of 90–100 km is exactly the size required for a 100 TeV hadron collider. Namely, the beam energy of a hadron collider is given by

$$E = ecB\rho, \quad (3)$$

where B is the dipole field, ρ the bending radius. Doubling the field compared with the LHC, and increasing the radius or circumference by a factor 3–4 yields a factor 6–8 increase in proton energy to about 100 TeV in the centre of mass.

In addition, the size of 90–100 km required for both FCC lepton and hadron colliders also matches the local topology of the Lake Geneva basin, where possible tunnel locations are bounded on two sides by the Jura and (Pre-)Alpes, respectively, and where, in addition, the collider should pass around the Salève mountain.

2.2 Linear Colliders

A linear collider still features moderate arcs in its beam delivery system, and also faces the issues of synchrotron radiation emitted in the final quadrupole magnets (Oide effect) and in collision (beamstrahlung), which ultimately limit the achievable beam size and the maximum beam energy of such colliders.

Indeed, some bending magnets are an integral part of the beam delivery systems, e.g., for the collimation of off-energy particles, and for the chromatic correction of the final focus. Synchrotron radiation emitted in these bending magnets can increase the beam size at the interaction point (IP), either directly due to the resulting increase of the horizontal emittance, or due to incomplete chromatic correction for particle energy changes that occur within the system [6]. These effects call for reduced bending as the beam energy is increased. At the same time, at higher energy the incoming geometric beam emittance adiabatically decreases, allowing for stronger sextupole magnets. In consequence, the geometry and the length of the beam delivery system change with beam energy. Two historical examples from the CLIC beam delivery design in **Figure 1** illustrate the beam-delivery footprint and length changes that may be required when increasing the collision energy from 500 GeV to 3 TeV. The initial tunnel layout should accommodate and provide space for the high-energy geometry. Even with the modified, optimised geometry synchrotron radiation is by no means negligible. For example, synchrotron radiation in the bending magnets caused a factor of about two loss in luminosity in the 2003 CLIC BDS design at 3 TeV (**Figure 1**, left picture) [7]; a similar situation was found for the SLC at a beam energy of only 45.6 GeV [8]. Such questions will also need to be addressed for a proposed 3 TeV energy upgrade of the International Linear Collider [9], or for upgrades of linear colliders to even higher energies, based on plasma acceleration.

A second limit set by synchrotron radiation in linear colliders arises in the final quadrupole magnets, where photon emission leads to an energy change, and thereby to a different focal length and increase in the vertical spot size ("Oide effect") [10], as is illustrated in **Figure 2**.

The third, and perhaps most important limitation due to synchrotron radiation at linear colliders relates to the one

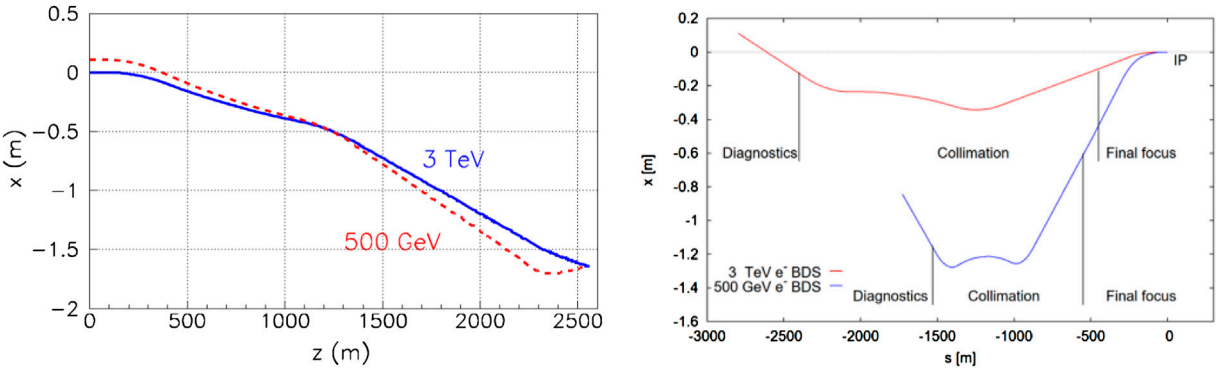


FIGURE 1 | Historical footprints of CLIC 3 TeV and 500 GeV beam delivery systems from 2003 [7] (left) and 2010 [69] (right), illustrating the layout changes required due to synchrotron radiation as a function of beam energy.

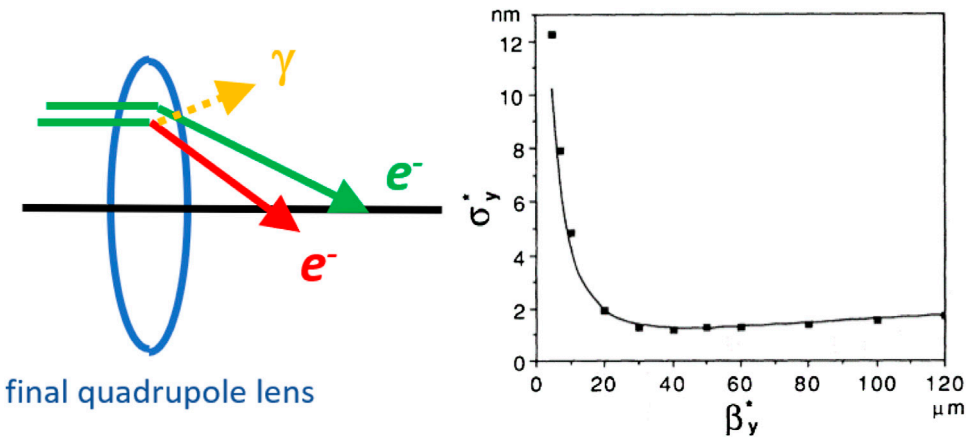


FIGURE 2 | Illustration of the Oide effect, where photon emission in the final quadrupole lens results in a minimum possible spot size for an optimized value of β_y^* [10], and an example of vertical rms beam size versus β_y^* from Ref. [6].

emitted during the collision in the electromagnetic field of the opposite beam, also called "beamstrahlung". The strength of the beamstrahlung is characterized by the parameter Υ, defined as [11,12] $\Upsilon \equiv \gamma B/B_c = (2/3)\hbar\omega_c/E_e$, with $B_c = m_e^2c^2/(e\hbar) \approx 4.4$ GT the Schwinger critical field, $\hbar\omega_c = (3/2)\hbar c\gamma^3/\rho$ the critical photon energy as introduced by Sands [13], E_e the electron (or positron) energy before radiation, B the local magnetic field, $\rho = e/(pB)$ the local bending radius, γ the relativistic Lorentz factor corresponding to E_e, $p \approx E_e/c$ the particle momentum, e the electron charge, and c the speed of light. The average Υ during the collision of three-dimensional Gaussian bunches is

$$\langle\Upsilon\rangle = \frac{5r_e^2}{6\alpha}\frac{N_b}{\sigma_z\left(\sigma_x^\star + \sigma_y^\star\right)}, \tag{4}$$

where α denotes the fine structure constant ($\alpha \approx 1/137$), $r_e \approx 2.8 \times 10^{-15}$ m the classical electron radius, N_b the bunch populaiton σ_z the rms bunch length, and $\sigma_{x(y)}^\star$ the rms horizontal (vertical) spot size at the collision point.

In the classical regime $\Upsilon \ll 1$, and for flat Gaussian beams, the number of photons emitted per beam particle during the collision is [14].

$$n_\gamma \approx 2.12\frac{\alpha N_b r_e}{\sigma_x^\star + \sigma_y^\star}. \tag{5}$$

The parameter n_γ is important, since it describes the degradation of the luminosity spectrum. Namely, the emission of beamstrahlung photons changes the energy of the emitting electron or positron, and thereby the energy of its later collision. The fraction of the total luminosity L_{tot} at the target centre-of-mass energy L_0 is determined by n_γ as [15].

$$\frac{L_0}{L_{\text{tot}}} = \frac{1}{n_\gamma^2}(1 - e^{-n_\gamma})^2, \tag{6}$$

To illustrate this degradation with an example, for CLIC at 380 GeV 60% of the total luminosity lie within 1% of the target energy, while at 3 TeV this fraction decreases to only 34%.

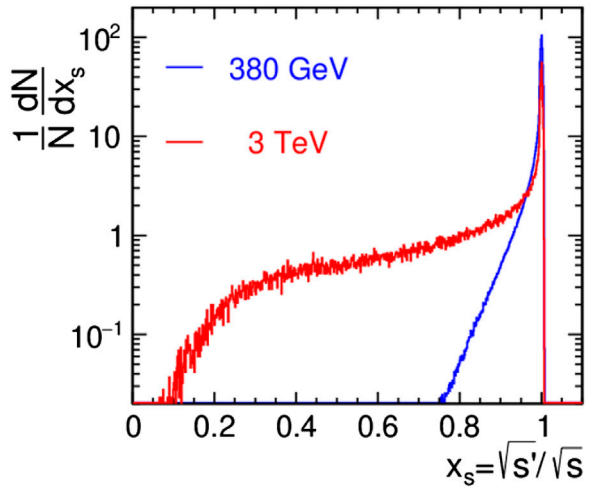


FIGURE 3 | Differential luminosity as a function of normalized centre-of-mass energy, $x_s = \sqrt{s'}/\sqrt{s}$ for CLIC at nominal centre-of-mass energies $\sqrt{s}$ of 380 GeV and 3 TeV [16].

Figure 3 presents the respective luminosity spectra [16]. In this way, at TeV energies, e^+e^- collisions in linear colliders lose their distinct energy precision.

2.3 Muon Colliders

The muon is about 200 times heavier than the electron, which, according to **Eq. 1**, implies close to 2×10^9 times less radiation at the same energy and bending radius. On the other hand, muon beams have two drawbacks: their production is not trivial, and the muons decay, with a rather short lifetime of only 2.2 μs at rest. In **Section 5.2**, we will present an innovative approach to the muon collider.

2.4 Shielding the Radiation

The radiation emission is suppressed at wavelengths larger than $\lambda_{sh} \approx 2\sqrt{d^3/\rho}$ with d signifying the pipe diameter [17]. Therefore, miniature accelerators with extremely small beam pipe on the micron or nanometre scale, combined with a large bending radius ρ could suppress almost all radiation. An extreme case would be the use of bent-crystals, where d becomes comparable to the inter-atom distance in the crystal lattice.

2.5 Shaping the Beam

It is noteworthy that classically a uniform time-independent beam does not emit any synchrotron radiation [18,19]. For example, the CERN ISR operated with high-current stationary beams. In the case of such a coasting beam, residual radiation could arise from shot noise or from beam instabilities. The shot noise might be reduced by suitable manipulations—see e.g., [20] — or by stochastic cooling. The shot noise and, therefore, the associated synchrotron radiation can be markedly reduced in case the cooling is so strong as to produce a crystalline beam [21]. Accelerating a "DC" (or near-DC) beam may be accomplished by induction acceleration [22].

3 HIGH-FIELD MAGNETS

The energy reach of hadron colliders is determined by their size and by the magnetic field—see **Eq. 3**.

All SC hadron storage rings built to date used magnets based on Nb-Ti conductor, for which the maximum reachable magnetic field is 8–9 T, as for the LHC dipole magnets. To go beyond this field level, the High Luminosity LHC (HL-LHC) upgrade foresees the installation of a few tens of higher-field magnets made from Nb_3Sn superconductor, with a design peak field of 11–12 T. The FCC-hh is designed with a few 1,000 of Nb_3Sn magnets with a higher field of 16 or 17 T, which is close to the maximum field that can be reached with this type of conductor. To achieve even higher fields, high-temperature superconductors are under consideration. At CERN magnets based on REBCO are being developed. In China iron-based superconductor, with a field of up to 24 T, is the material of choice for the SPPC.

The coils of the SC magnets for future hadron colliders must withstand extreme pressure and forces, without any quench and without any degradation in performance. The horizontal forces per quadrant in dipole accelerator magnets approach 10 MN/m for a field of 20 T [23].

4 ACCELERATING SYSTEMS

4.1 SC Radiofrequency Systems

As for the bending fields, also for the accelerating systems, superconducting materials have gained widespread use. Superconducting radiofrequency (RF) cavity systems underpin many modern facilities, the latest examples being the European XFEL at DESY Hamburg, the LCLS-II at SLAC, and FRIB in Michigan. Accelerating fields have been increased from a few MV/m to more than 30 MV/m for multicell cavities, and close to twice this value for single cells. Most SC cavities to date have been based on bulk Nb or in Nb-on-Cu cavities. New cavity treatments (nitrogen doping or nitrogen infusion [24]), innovative production methods (chemical vapor deposition [25], high impulse power magnetron sputtering [26]) and new materials, e.g., Nb_3Sn [27], as for the magnets, etc. promise further significant advances in performance, by factors of 2–10 in quality factor Q_0 and of 2–3 in maximum accelerating gradient. As an example, for Nb_3Sn, the theoretical ultimate "superheating" field [28] corresponds to a maximum accelerating gradient of ~ 100 MV/m, about twice the corresponding value for Nb, while the latter is not far from the currently achieved peak values of about 50 MV/m for Nb cavities [27].

4.2 Plasma Acceleration and Crystals

Other advanced accelerating concepts can reach much higher gradients. For example, plasma acceleration routinely achieves fields of 100 GV/m, which is 3,000 times higher than the Nb cavities proposed for the International Linear Collider. The accelerating plasma waves can be driven either by a high-energy charged particle beam or by a laser. Comprehensive

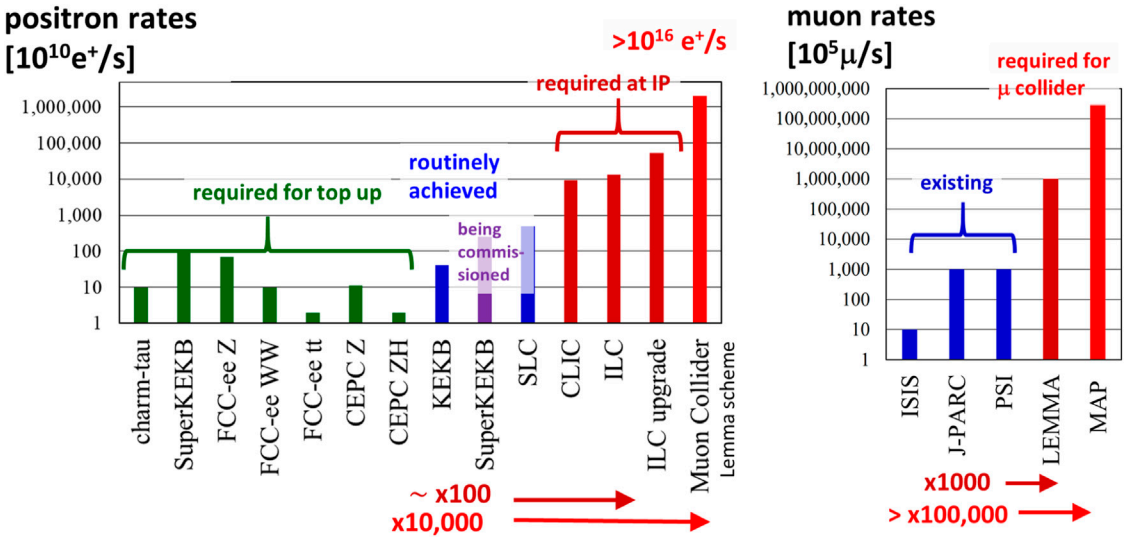


FIGURE 4 | Production rates of positrons at dedicated e^+e^- colliders [68] (left) and of muons at non-collider facilities [70] (right) presently achieved (blue) and required for various future collider projects (red, brown and green). Note the logarithmic scale on the vertical axes.

concepts have been developed for electron-positron colliders based on either beam-driven [29,30] or laser-driven plasma acceleration [31,32]. Beam quality, pulse-to-pulse stability, and energy efficiency of plasma accelerators [33] are critical issues addressed by ongoing R& D programs. High-energy colliders are arguably the most demanding application of plasma acceleration. Possible ultimate limits of plasma acceleration arise from the scattering of beam particles off plasma nuclei and plasma electrons, and from the emission of betatron radiation [34]. Both of these effects might be partially mitigated by accelerating in a hollow plasma channel. For realizing e^+e^- colliders, not only electrons but also positrons must be accelerated in the plasma, while preserving the beams' transverse and longitudinal emittance. For this purpose, more complex plasma excitation schemes may need to be developed, e.g., [35, 36].

Thanks to their higher electron density, even larger gradients can be generated in crystals. The maximum field is given by [37].

$$E_0 \approx \frac{m_e c \omega_p}{e} \approx 100\left[\frac{\text{GeV}}{\text{m}}\right]\sqrt{n_0[10^{18}\ \text{cm}^{-3}]}, \qquad (7)$$

with ω_p the angular plasma frequency and n_0 the electron density. With $n_0 \approx 10^{22}$ cm^{-3} to 5×10^{24} cm^{-3} in a crystal, peak gradients of 10–1000 TV/m would be within reach. Accelerating crystal waves could be excited by X-ray lasers [37].

5 PRODUCTION OF UNSTABLE OR RARE PARTICLES

Several future colliders require unprecedented production rates of positrons (linear colliders) and muons (muon collider), while future circular colliders need positrons at a level already demonstrated, as is illustrated in **Figure 4**.

The present world record positron production rate of about 5×10^{12} e^+ per second was established at the SLC in the 1990s. Even achieving, or reproducing, this SLC rate is not trivial. The SLC target failed after 5 years of operation. For a dedicated failure analysis performed at LANL, the failed SLC positron target was cut into pieces and metallographic studies were carried out to examine the level of deterioration of material properties due to radiation exposure. The hardness of the target material in units of kg/mm^2 was found to be decreased by about a factor of 2, over the first 10 mm. However, whether this degradation had been due to radiation damage, work hardening, or temperature cycling could not be clearly resolved.

To push the production rate of e^+ and μ's much beyond the state of the art, a candidate ultimate source of positrons and muons is the Gamma factory [38], which we discuss in the following subsection.

5.1 Gamma Factory

The Gamma factory [38] is based on resonant scattering of laser photons off partially stripped heavy-ion beam in the existing LHC or in the planned FCC-hh. Profiting from two Lorentz boosts, the Gamma factory acts as a high-stability laser-light-frequency converter, with a maximum photon frequency equal to $\nu_{\gamma,\text{max}} = 4\gamma^2\nu_{\text{laser}}$, where γ is the relativistic Lorentz factor of the partially stripped ion beam. This allows the production of intense bursts of gamma rays with photon energies of up to several 100 MeV.

In particular, the Gamma factory can serve as a powerful source of e^+ (yielding 10^{16}–10^{17} e^+/s—five orders of magnitude higher than the state of the art), μ (10^{11}–10^{12}/s), π, etc. The positron rate available from the Gamma factory would be sufficient for a LEMMA type muon collider [39,39]. The Gamma factory would also allow for doppler laser cooling of

high-energy beams, and, thereby, provide an avenue to a High Luminosity LHC based on laser-cooled isocalar ion beams [41].

5.2 Induction Acceleration and Positron Annihilation in Plasma Target

The LEMMA scheme for a muon collider is based on the annihilation of positrons with electrons at rest [39]. The cross section for continuum muon pair production $e^+e^- \rightarrow \mu^+\mu^-$ has a maximum value of about 1 μb at a centre-of-mass energy of ~0.230 GeV, which corresponds to a positron beam energy of about 45 GeV, exactly as required for the FCC-ee operating as a TeraZ factory and provided by the FCC-ee full-energy booster [42].

Challenges with the LEMMA-type muon production scheme relate to the emittance preservation of muons and muon-generating positrons upon multiple traversals through a target, and the merging of many separate muon bunchlets, due to production by many separate positron bunches or positron bunch passages.

These challenges may potentially be overcome by [43]:

- Operating the FCC-ee booster with a barrier bucket and induction acceleration, so that all positrons of a cycle are merged into one single superbunch [44], instead of ~ 10,000 separate bunches.
- Sending the positron superbunch from the booster into a plasma target, where, during the passage of the positron superbunch, the electron density is enhanced 100–1,000 fold without any significant density of nuclei, hence with beamstrahlung and Coulomb scattering absent.

Since the positron bunch will be mismatched to the nonlinear plasma channel, filamentation and significant transverse emittance growth may result [43].

For a typical initial plasma electron density of $n_e = 10^{23}$ m^{-3}, and assuming a density enhancement by a factor of 1,000, due to the electron pinch in the positive electric field of the positron beam, the positrons annihilate into muon pairs at a rate of 10^{-8} m^{-1}.

As described in the CDR [45], the FCC-ee booster can accelerate 3.5×10^{14} positrons every 50 s. Using the much more powerful Gamma Factory positron source, with a rate of 10^{16}–10^{17} $e^+ s^{-1}$ [38], and injecting into the booster during one or a few seconds, of order 10^{17} e^+ can be accumulated, at the booster injection energy of ~20 GeV. The positrons can be captured into a single barrier RF bucket, with a final length of ~ 5 m, at which the longitudinal density would be about 1,000 times higher than the peak bunch density in the collider ring (without collision), possibly compromising the beam stability.

Accelerating the long positron superbunch containing 10^{17} e^+ by 25 GeV, from 20 to 45 GeV, requires a total energy of 0.4 GJ, or, if accelerated over 2 s, about 200 MW of RF power. This translates into an induction acceleration voltage of ~2 MV per turn, which is three orders of magnitude higher than the induction voltage of the KEK digital accelerator [46], but about 10 times lower than the induction RF voltage produced at the LANL DARHT-II [47], at much higher or lower repetition rate, respectively. On the ramp and at top energy, the full bunch length l_b can conceivably be compressed to the assumed $l_b \approx 5$ m, by squeezing the gap of the barrier bucket (which requires substantially more voltage for the barrier RF system)—also see [21,46]. Tentative parameters of the positron superbunch are compiled in **Table 1**. We assume that the booster ring runs near the coupling resonance so that the emittance is shared between the two transverse planes.

TABLE 1 | Tentative parameters of the positron superbunch sent onto the plasma target.

Parameter	Symbol	Value	Unit
beam energy	E_b	45	GeV
total bunch length	l_b	5	m
bunch population	N_b	1,000	10^{14}
transverse rms emittance	$\varepsilon_{x,y}$	135	pm
initial beta function at plasma entrance	$\beta_{x,y}$	0.1	m
initial rms beam size	$\sigma_{x,y}$	3.7	μm

When the accelerated and compressed positron bunch is sent into the plasma channel, we consider that the plasma electron distribution quickly acquires a nearly stationary shape, while any remaining plasma ions are slowly repelled away from the positron beam. In the stationary phase, the electron distribution approaches a shape that mimics the one of the positron beam, with a density

$$n_{e,\mathrm{stat}} \approx \frac{N_b}{2\pi l_b \sigma_\perp^2}, \tag{8}$$

so as to neutralize the electric field. With an average rms size of $\sigma_\perp \approx 10$ μm, we obtain $n_{e,\mathrm{stat}} \approx 10^{26}$ m^{-3}. Considering a 100 m long plasma channel yields ~ 10^{11} μ pairs, with an initial muon energy of ~22 GeV, and an initial lifetime of 0.5 ms at this energy.

In particular, once the electron distribution is nearly stationary, the longitudinal fields inside the plasma can be neglected. The resulting transverse emittance of the produced muons can be optimized by adjusting positron beam parameters and the optical functions at the entrance to the plasma [43]. In addition, a phase rotation (bunch compression) of the muons may be required, since the initial bunch length ~ 5 m, of the positrons or resulting muons, will still be too long for collider operation.

Overall, the described scheme, sketched in **Figure 5**, would produce about 10^{12} muon pairs per cycle, with a cycle length of order 3 s. Even at an energy of 50 TeV, the muons would decay with a lifetime of only 1.1 s. This kind of cycle/lifetime ratio of about 3:1 might still be considered acceptable. On the other hand, for collision at a muon beam energy of seven TeV in the existing LHC ring, the muon lifetime would be only 0.15 s, and the scheme would be considerably more challenging.

6 COST AND SUSTAINABILITY

6.1 Efficient RF Power Sources

Radiofrequency (RF) systems are used to keep a charged particle beam bunched, and to feed energy to the beam, be it for purposes of acceleration or to compensate for the energy lost due to

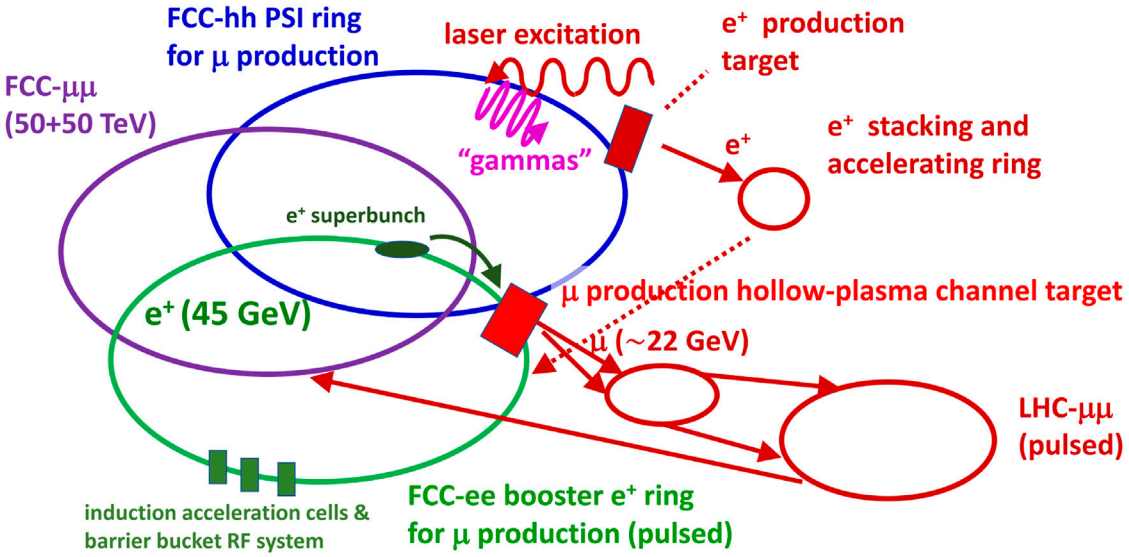


FIGURE 5 | Concept of a 100 TeV μ collider based on FCC-hh and FCC-ee. In one of the FCC-hh rings, partially stripped heavy ions are collided with a pulsed laser to generate intense high-energy gamma rays that are converted into positrons, which are accumulated, then accelerated, and injected into a barrier bucket in the FCC-ee booster ring. The resulting superbunch is brought to a positron energy of 45 GeV, with induction acceleration, where the superbunch is extracted and sent into a plasma target, leading to a plasma electron density enhancement and, thereby, amplified annihilation into muon pairs. The muons are accelerated in a modifed SPS and LHC, to be finally injected and accelerated in the second FCC-hh collider ring. This is a modified version of the scheme presented in Ref. [40].

synchrotron radiation. In superconducting continuous-wave RF cavities, almost no power is lost to the cavity wall and all RF power entering the cavity can be transferred to the beam highly efficiently. Then, in the overall power budget, the RF power source is the most inefficient element. For RF frequencies above about 400 MHz, and for high power applications, historically klystrons have been the RF power source of choice on particle accelerators.

It is most remarkable that about 80 years after the invention of the klystron by the Varian brothers, a revolution in klystron technology is underway. Using advanced bunching techniques, it is expected that the klystron efficiency can be raised from the present 50–60% level to about 90%, which would translate into a significant energy saving [48]. Prototypes of such novel highly-efficient klystrons are being manufactured both by CERN, in collaboration with industry, for FCC, CLIC and ILC, and, in China, for the CEPC project.

In parallel, the efficiency of alternative RF power sources, such as inductive output tubes or solid-state amplifiers [49], is also being improved.

While at present the RF power sources are the dominant contributors to overall grid-to-beam power transmission inefficiency, a few percent additional losses each occur in the electrical network between utility high-voltage interconnect point and RF power source, and in the wave guides and couplers feeding the generated RF power into the accelerating cavities, respectively.

6.2 Efficient Magnets

For high fields, superconducting magnets are most efficient, as no energy is lost, and electric power is mostly required for the cryogenic system. For lower fields, up to of order 1 T, permanent magnets are most energy efficient. An example is the Fermilab Recycler Ring [50], which was built almost entirely from permanent magnets. Even adjustable permanent magnets have been designed and built for applications at light sources, colliders, and plasma accelerators [51]. Other ingenious solutions for energy saving can be found, depending on the respective application. For example, for the FCC-ee double-ring collider, twin dipole and quadrupole magnets at low field (of order 0.05 T, for the dipoles) have been designed [52], which promise a significant power reduction compared with the magnets of comparable fields at earlier colliders.

6.3 Energy Recovery Linacs

Recovering the energy of the spent beam after one or several collisions is another effective measure to improve overall energy efficiency, if a significant fraction of the overall electric power is stored in the beam, as typically is the case for beams accelerated in superconducting linacs [53].

A comparison of ERL-based colliders proposed half a century ago with several recent concepts is presented in **Table 2**. The main differences between proposals from the 1970s and today are the collision of flat beams instead of round beams, much smaller (vertical) beam sizes, combined with higher beam current, yielding, on paper, of order ~10,000 times higher luminosity than the proposals from half a century ago.

6.4 Beam Loss Control and Machine Protection

Also minimisation of beam loss can improve the energy efficiency of accelerators, such as ERLs. For proposed future higher-energy

TABLE 2 | A comparison of ERL-based colliders proposed in the 1960s [53] and 1970s [71,72], and in the recent period 2019–2021 [3,2].

Proposal	Tigner 1965 [53]	Amaldi 1976 [71]	Gerke-Steffen 1979 [72]	Litvinenko et al. 2020 [2]		Telnov 2021 [3]	
c.m. energy [GeV]	1–6	300	200	240	600	250	500
av. beam current [mA]	120	10	0.3	2.5	0.16	100	100
vert. rms IP beam size [nm]	40,000 (round)	2,000 (round)	900 (round)	6	5	6.1	7.4
luminosity [10^{34} cm^{-2}s^{-1}]	0.0003	0.01	0.004	73	8	90	64

facilities, machine protection and beam collimation systems become ever more challenging due to their unprecedented beam power or stored energy. For example, the FCC-hh design features a stored beam energy of 8.3 GJ [54], which is more than a factor 20 higher than for the LHC.

7 NOVEL APPLICATIONS AND DIRECTIONS

Storage rings constructed as high energy physics colliders could also serve for other intriguing applications. In this section, we mention a few examples.

7.1 Ultimate Light Sources

Large circular storage rings like the FCC-ee, and even the FCC-hh, can serve as ultimate storage-ring light sources, with diffraction limited emittances down to photon wavelengths of

$$\lambda_{\min} \approx 4\pi\varepsilon_x. \tag{9}$$

For FCC-ee the geometric emittance ε_x, of the collider or of the full-energy booster, scales as γ^2, and the lowest value of $\varepsilon_x \approx 50$ pm is reached at the injection energy of 20 GeV, resulting in $\lambda_{\min,ee} \approx 650$ pm. With a beam current of 1.5 A or higher, this could represent a formidable light source. Conversely, for FCC-hh the normalized proton beam emittance $\gamma\varepsilon_x$ shrinks during proton beam storage at 50 TeV to ~ 0.2 μ m [54], corresponding to a geometric emittance of 4 pm, and the associated minimum wavelength is $\lambda_{\min,ee} \approx 50$ pm, still more than an order of magnitude lower than for the FCC-ee. The FCC-hh design beam current is 0.5 A.

The FCC-ee ring emittance could be further reduced by factors of 10–100 through the addiition of damping wigglers, pushing the accessible wavelength into the 10 pm regime.

A more detailed study of synchrotron light produced by such low-emittance FCC-ee beams passing through realistic undulator configurations has been performed recently [55]. For hadron storage rings, their use as a light source was discussed in the past, e.g., for the Superconducting Super Collider (SSC) [56].

In addition, also Free Electron Lasers (FELs) based on ERLs designed for high-energy physics colliders can offer outstanding performance in terms of average brightness, and in their wavelength reach down into the few picometre range [57], e.g., in the case of the LHeC-ERL based FEL, with a beam current of ~20 mA.

7.2 Detection of Gravitational Waves

Various approaches have been suggested for using beams in a storage ring for the detection of gravitational waves [58–61] including the construction of special optics with regions of extremely high beta functions that would serve as gravitational wave antennae [61,62]. Exploration of such possibilities continues.

7.3 Storage Rings as Quantum Computers

With advanced cooling and manipulation schemes, storage rings might eventually be used as quantum computers [63,63]. Indeed, combining the storage rings of charged particles with the linear ion traps used for quantum computing and mass spectrometry would enable a large leap in the number of ions serving as qubits in the quantum computing. Such an approach holds the promise of significant advances in general quantum calculations and, especially, in simulations of complex quantum systems.

8 BEYOND THE EARTH

To reach the Planck scale of 10^{28} eV, linear or circular colliders would need to have a size of order 10^{10} m, which is about a 10th of the distance between the Earth and the Sun, if operated close to the Schwinger critical field [65,66].

Following the FCC a possible next or next-next step in this direction could be a circular collider on the Moon (CCM) [67]. With a circumference of about 11 Mm, a centre-of-mass energy of about 14 PeV (1,000 times the energy of the LHC), based on 6×10^5 dipoles with 20 T field, either ReBCO, requiring ~7–13 ktons of rare-earth elements, or iron-based superconductor (IBS), requiring of order a million tons of IBS [67]. Many of the raw materials required to construct machine, injector complex, detectors, and facilities can potentially be sourced directly on the Moon. The 11,000-km tunnel should be constructed a few 10–100 m under lunar surface to avoid lunar day-night temperature variations, cosmic radiation damage, and meteoroid strikes. A “Dyson band” or “Dyson belt” could be used to continuously collect Sun power. Operating the collider would require the equivalent of 0.1% of the Sun power incident on Moon surface [67].

9 SUMMARY AND OUTLOOK

Particle colliders boast an impressive 70 years long history, with dramatic improvements in performance, and they will also be the cornerstone for a long and exciting future in high-energy physics. Future colliders should heed the lessons from the previous generations of colliders, like LEP, SLC, KEKB, PEP-II, LHC, and SuperKEKB.

Present collider-accelerator R&D trends include the development of more powerful positron sources; the widespread application of energy recovery; "nanobeam" handling—with stabilisation, positioning, and tuning; the polarization control at the 0.1% level; monochromatization; the use of machine learning and artificial intelligence, e.g., for automated design and for accelerator operation; and the introduction of novel uses such as for probing gravity or developing high-throughput quantum computing; plus, last not least, bringing advanced acceleration schemes to maturity.

Considering the desired higher intensity and energy for future machines, a major challenge will be to make the future colliders truly "green," that is energy-efficient and sustainable. In this context, suppressing synchrotron radiation or mitigating its impact becomes a key objective for the long term. Concerning the near term, it is important to observe that the Future Circular lepton Collider, FCC-ee, is the most sustainable of all the proposed Higgs and electroweak factory proposals, in that it implies the lowest energy consumption for a given value of total integrated luminosity [68], over the collision energy range from 90 to 365 GeV.

For the Future Circular Collider (FCC) effort, the next concrete steps encompass a specific local and regional implementation scenario worked out in collaboration with host states, machine design optimization, physics studies and technology R&D, performed via a global collaboration and supported by the EC H2020 FCC Innovation Study, to prove the FCC feasibility by 2025/26.

AUTHOR CONTRIBUTIONS

The work reported reflects half a century of accelerator R&D and progress with charged particle colliders. The author confirms being the sole author of this review article, which cites relevant contributions by others, and has approved it for publication.

FUNDING

This work was supported, in parts, by funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No. 101004730 (iFAST).

ACKNOWLEDGMENTS

The author would like to thank R. Aßmann, M. Benedikt, G. Franchetti, K. Oide and many other colleagues for helpful discussions.

REFERENCES

1. Abelleira Fernandez JL, Adolphsen C, Akay AN, Aksakal H, Albacete JL, Alekhin S, et al. A Large Hadron Electron Collider at CERN: Report on the Physics and Design Concepts for Machine and Detector. *J. Phys. G: Nucl. Part. Phys* (2012) 39:075001. doi:10.1088/0954-3899/39/7/075001
2. Litvinenko VN, Roser T, Chamizo-Llatas M. High-Energy High-Luminosity e^+e^- Collider Using Energy-Recovery Linacs. *Phys Lett B* (2020) 804:135394. doi:10.1016/j.physletb.2020.135394
3. Telnov VI. A High-Luminosity Superconducting Twin e^+e^- Linear Collider with Energy Recovery. *JINST* (2021) 16:P12025. doi:10.1088/1748-0221/16/12/P12025
4. Richter B. Very High Energy Electron-Positron Colliding Beams for the Study of Weak Interactions. *Nucl Instr Methods* (1976) 136:47–60. doi:10.1016/0029-554X(76)90396-7
5. Myers S. FCC: Building on the Shoulders of Giants. *Eur Phys J Plus* (2021) 136:1076. doi:10.1140/epjp/s13360-021-02056-w
6. Oide K. A Final Focus System for Flat-Beam Linear Colliders. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1989) 276:427–32. doi:10.1016/0168-9002(89)90567-6
7. Aleksa M, Assmann RW, Burkhardt H, Jeanneret JB, Redaelli S, Risselada T, et al. CLIC Beam Delivery System. *CERN Rep CLIC-Note-551* (2003).
8. Zimmermann F. *Magnet Alignment Tolerances in the 1994 SLC Final Focus System*. SLAC SLC Collider Note SLAC-CN-398 (1994).
9. Padamsee H. ILC Upgrades to 3 TeV. *arXiv 2108.11904* (2021).
10. Oide K. Synchrotron-Radiation Limit on the Focusing of Electron Beams. *Phys Rev Lett* (1988) 61:1713–5. doi:10.1103/PhysRevLett.61.1713
11. Yokoya K. Quantum Correction to Beamstrahlung Due to the Finite Number of Photons. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1986) 251:1–16. doi:10.1016/0168-9002(86)91144-7
12. Yokoya K, Chen P. Beam-Beam Phenomena in Linear Colliders. *Lect Notes Phys* (1992) 400:415–45. doi:10.1007/3-540-55250-2_37
13. Sands M. The Physics of Electron Storage Rings: An Introduction. *Conf Proc C* (1969) 6906161:257–411.
14. Chen P. Review of Linear Collider Beam-Beam Interaction. *AIP Conf Proc* (1989) 184:633–79. doi:10.1063/1.38047
15. Chen P. Differential Luminosity Under Multiphoton Beamstrahlung. *Phys Rev D* (1992) 46:1186–91. doi:10.1103/PhysRevD.46.1186
16. Abramowicz H, Alipour Tehrani N, Arominski D, Benhammou Y, Benoit M, Blaising JJ, et al. Top-Quark Physics at the CLIC Electron-Positron Linear Collider. *J High Energ Phys* (2018) 2019:3. doi:10.1007/JHEP11(2019)003
17. Warnock RL. Shielded Coherent Synchrotron Radiation and its Effect on Very Short Bunches. In: 4th Advanced ICFA Beam Dynamics Workshop: Collective Effects in Short Bunches. SLAC Publication SLAC-PUB-5375 (1990).
18. Schwinger J. On Radiation by Electrons in a Betatron. *Quantum Legacy* (1945) 1945:307–31. doi:10.2172/1195620
19. Arzimovitch L, Pomeranchuk I. The Radiation of Fast Electrons in the Magnetic Field. *J Phys (Moscow)* (1945) 9:267.
20. Ratner D, Stupakov G. Observation of Shot Noise Suppression at Optical Wavelengths in a Relativistic Electron Beam. *Phys Rev Lett* (2012) 109:034801. doi:10.1103/PhysRevLett.109.034801
21. Primack H, Blümel R. Synchrotron Radiation of Crystallized Beams. *Phys Rev E* (1999) 60:957–67. doi:10.1103/PhysRevE.60.957
22. Takayama K, Briggs RJ, (editors). Induction Accelerators. *Particle Acceleration and Detection*. Heidelberg, Germany: Springer (2011). doi:10.1007/978-3-642-13917-8
23. Bottura L. High Field Magnets. In: Future Collider Forum: 1st Workshop, DESY; 6–8 October, 2021 (2021).
24. Dhakal P. Nitrogen Doping and Infusion in SRF Cavities: A Review. *Phys Open* (2020) 5:100034. doi:10.1016/j.physo.2020.100034
25. Ge M, Arrieta V, Gruber T, Kaufman JJ, Liepe M, Maniscalco JT, et al. CVD Coated Copper Substrate SRF Cavity Research at Cornell University. In: Proc. 19th Int. Conf. RF Superconductivity (SRF'19). Geneva: JACoW Publishing (2019). p. 381–6. doi:10.18429/JACoW-SRF2019-TUFUB8
26. Sayeed M, Elsayed-Ali H, Eremeev GV, Kelley MJ, Reece CE, Pudasaini U. Magnetron Sputtering of Nb3Sn for SRF Cavities. In: Proc. 9th Int. Particle Accelerator Conf. (IPAC'18). Geneva: JACoW Publishing (2018). p. 3946–9. doi:10.18429/JACoW-IPAC2018-THPAL129

27. Posen S, Liepe M, Eremeev G, Pudasaini U, Reece CE. Nb3Sn Superconducting Radiofrequency Cavities: A Maturing Technology for Particle Accelerators and Detectors. In: 2022 Snowmass Summer Study (2022).
28. Padamsee H, Shepard KW, Sundelin R. Physics and Accelerator Applications of Rf Superconductivity. *Annu Rev Nucl Part Sci* (1993) 43:635–86. doi:10.1146/annurev.ns.43.120193.003223
29. Adli E, Delahaye J-P, Gessner SJ, Hogan MJ, Raubenheimer T, An W, et al. A Beam Driven Plasma-Wakefield Linear Collider: From Higgs Factory to Multi-TeV. *arXiv 1308.1145* (2013). doi:10.2172/1074154
30. Chen JBB, Schulte D, Adli E. e^+e^- Beam-Beam Parameter Study for a TeV-Scale PWFA Linear Collider. *arXiv 2009.13672* (2021).
31. Leemans W, Esarey E. Laser-Driven Plasma-Wave Electron Accelerators. *Phys Today* (2009) 62:44–9. doi:10.1063/1.3099645
32. Schroeder CB, Benedetti C, Esarey E, Leemans WP. Laser-Plasma-Based Linear Collider Using Hollow Plasma Channels. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2016) 829:113–6. doi:10.1016/j.nima.2016.03.001
33. Lebedev V, Burov A, Nagaitsev S. Efficiency versus Instability in Plasma Accelerators. *Phys Rev Accel Beams* (2017) 20:121301. doi:10.1103/PhysRevAccelBeams.20.121301
34. Zimmermann F. Possible Limits of Plasma Linear Colliders. *J Phys Conf Ser* (2017) 874:012030. doi:10.1088/1742-6596/874/1/012030
35. Xu ZY, Xiao CF, Lu HY, Hu RH, Yu JQ, Gong Z, et al. New Injection and Acceleration Scheme of Positrons in the Laser-Plasma Bubble Regime. *Phys Rev Accel Beams* (2020) 23:091301. doi:10.1103/PhysRevAccelBeams.23.091301
36. Hue CS, Cao GJ, Andriyash IA, Knetsch A, Hogan MJ, Adli E, et al. Efficiency and Beam Quality for Positron Acceleration in Loaded Plasma Wakefields. *Phys Rev Res* (2021) 3:043063. doi:10.1103/PhysRevResearch.3.043063
37. Shiltsev VD. High-Energy Particle Colliders: Past 20 years, Next 20 years, and Beyond. *Phys.-Usp.* (2012) 55:965–76. doi:10.3367/UFNe.0182.201210d.1033
38. Krasny MW. The Gamma Factory Proposal for CERN. *arXiv 1511.07794* (2015).
39. Antonelli M, Boscolo M, Di Nardo R, Raimondi P. Novel Proposal for a Low Emittance Muon Beam Using Positron Beam on Target. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2016) 807:101–7. doi:10.1016/j.nima.2015.10.097
40. Zimmermann F. LHC/FCC-Based Muon Colliders. *J Phys Conf Ser* (2018) 1067:022017. doi:10.1088/1742-6596/1067/2/022017
41. Krasny M, Petrenko A, Placzek W. The Gamma Factory Path to High-Luminosity LHC with Isoscalar Beams. *PoS* (2021) ICHEP2020:690. doi:10.22323/1.390.0690
42. Benedikt M, Zimmermann F. Future Circular Collider: Integrated Programme and Feasibility Study. *Submitted Front Phys* (2022).
43. Zimmermann F, Latina A, Blondel A, Antonelli M, Boscolo M. *Muon Collider Based on Gamma Factory, FCC-Ee and Plasma Target* (2022). Submitted to IPAC'22.
44. Blondel A. *Private Communication* (2021).
45. The FCC Collaboration. FCC-ee: The Lepton Collider. *Eur Phys J Spec Top* (2019) 228:261. doi:10.1140/epjst/e2019-900045-4
46. Takayama K, Yoshimoto T, Barata M, Wah LK, Xingguang L, Iwashita T, et al. Induction Acceleration of Heavy Ions in the KEK Digital Accelerator: Demonstration of a Fast-Cycling Induction Synchrotron. *Phys Rev ST Accel Beams* (2014) 17:010101. doi:10.1103/PhysRevSTAB.17.010101
47. Nielsen K, Barraza J, Kang M, Bieniosek F, Chow K, Fawley W, et al. Upgrades to the DARHT Second Axis Induction Cells. *Conf Proc C* (2007) 070625:2385. doi:10.1109/PAC.2007.4441258
48. Baikov AY, Marrelli C, Syratchev I. Toward High-Power Klystrons with RF Power Conversion Efficiency on the Order of 90. *IEEE Trans Electron Dev* (2015) 62:3406–12. doi:10.1109/TED.2015.2464096
49. Barov N, Chang X, Newsham DJ, Wu D. Development of the Energy-Efficient Solid State RF Power Source for the Jefferson Laboratory CEBAF Linac. *Conf Proc C* (2012) 1205201:3455–7.
50. Volk JT. Experiences with Permanent Magnets at the Fermilab Recycler Ring. *J Inst* (2011) 6:T08003. doi:10.1088/1748-0221/6/08/t08003
51. Shepherd B. Permanent Magnets for Accelerators. *JACoW* (2020) IPAC2020:MOVIRO05. doi:10.18429/JACoW-IPAC2020-MOVIRO05
52. Milanese A. Efficient Twin Aperture Magnets for the Future Circular e^+/e^- Collider. *Phys Rev Accel Beams* (2016) 19:112401. doi:10.1103/PhysRevAccelBeams.19.112401
53. Tigner M. A Possible Apparatus for Electron Clashing-Beam Experiments. *Nuovo Cim* (1965) 37:1228–31. doi:10.1007/BF02773204
54. The FCC Collaboration. FCC-hh: The Hadron Collider. *Eur Phys J Spec Top* (2019) 228:755. doi:10.1140/epjst/e2019-900087-0
55. Casalbuoni S, Benedikt M, Doser M, Zimmermann F. First Thoughts on the Synergetic Use of the FCC-ee Collider and its Injector Complex for Photon Science and Other Applications. *CERN/FCC Intern Note* (2020).
56. Dutt SK. Synchrotron Radiation from Protons. In: Workshop on 4th Generation Light Sources (1992). SSC Pub SSCL-Preprint-182.
57. Nergiz Z, Mirian NS, Aksoy A, Zhou D, Zimmermann F, Aksakal H. Bright Ångstrom and Picometer Free Electron Laser Based on the Large Hadron Electron Collider Energy Recovery Linac. *Phys Rev Accel Beams* (2021) 24:100701. doi:10.1103/PhysRevAccelBeams.24.100701
58. Zer-Zion D. Possible Detection of High Frequency Gravitational Waves in Storage Rings: Speculation on Future Applications. *Astropart Phys* (2000) 14(3):239–43.
59. van Holten JW. *Cyclotron Motion in a Gravitational Wave Background* (1999). doi:10.48550/arXiv.gr-qc/9906117
60. Rao S, Brüggen M, Liske J. Detection of Gravitational Waves in Circular Particle Accelerators. *Phys Rev D* (2020) 102:122006. doi:10.1103/PhysRevD.102.122006
61. Berlin A, Brüggen M, Buchmueller O, Chen P, D'Agnolo RT, Deng R, et al. Storage Rings and Gravitational Waves: Summary and Outlook. In: ARIES WP6 Workshop: Storage Rings and Gravitational Waves (2021). arXiv 2105.00992.
62. Oide K. *Response of a Storage-Ring Beam to a Gravitational Wave - Preliminary Considerations* (2021). Presented at ARIES SRGW2021 Workshop.
63. Brown KA, Roser T. Towards Storage Rings as Quantum Computers. *Phys Rev Accel Beams* (2020) 23:054701. doi:10.1103/PhysRevAccelBeams.23.054701
64. Shaftan T, Blinov BB. Cold Ion Beam in a Storage Ring as a Platform for Large-Scale Quantum Computers and Simulators: Challenges and Directions for Research and Development. *Phys Rev Accel Beams* (2021) 24:094701. doi:10.1103/PhysRevAccelBeams.24.094701
65. Chen P, Noble RJ. Crystal Channel Collider: Ultra-High Energy and Luminosity in the Next Century. *AIP Conf Proc* (1997) 398:273–85. doi:10.1063/1.53055
66. Zimmermann F. Future Colliders for Particle Physics-"Big and Small". *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2018) 909:33–7. doi:10.1016/j.nima.2018.01.034
67. Beacham J, Zimmermann F. A Very High Energy Hadron Collider on the Moon. *New J Phys* (2022) 24:023029. doi:10.1088/1367-2630/ac4921
68. Benedikt M, Blondel A, Janot P, Mangano M, Zimmermann F. Future Circular Colliders Succeeding the LHC. *Nat Phys* (2020) 16:402–7. doi:10.1038/s41567-020-0856-2
69. Zamudio G, Tomas R. *Optimization of the CLIC 500 GeV Final Focus System and Design if a New 3 TeV Final Focus System with L*=6.0 M*. Geneva: Tech. Rep., CERN (2010).
70. Zimmermann F. Shaping the Future of High-Energy Physics - Accelerators. *CERN Acad Train Lectures* (2021).
71. Amaldi U. A Possible Scheme to Obtain e-e- and e+e- Collisions at Energies of Hundreds of GeV. *Phys Lett B* (1976) 61:313–5. doi:10.1016/0370-2693(76)90157-X
72. Gerke H, Steffen K. Note on a 45-GeV - 100-GeV 'Electron Swing' Colliding Beam Accelerator. *DESY Rep DESY-PET-79-04* (1979).

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zimmermann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

REVIEW
published: 27 May 2022
doi: 10.3389/fphy.2022.864823

Quantum Computing Applications in Future Colliders

Heather M. Gray[1,2]* *and Koji Terashi*[3]

[1]Department of Physics, University of California, Berkeley, Berkeley, CA, United States, [2]Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States, [3]International Center for Elementary Particle Physics (ICEPP), The University of Tokyo, Tokyo, Japan

High-energy physics is facing a daunting computing challenge with the large amount of data expected from the HL-LHC and other future colliders. In addition, the landscape of computation has been expanding dramatically with technologies beyond the standard x86 CPU architecture becoming increasingly available. Both of these factors necessitate an extensive and broad-ranging research and development campaign. As quantum computation has been evolving rapidly over the past few years, it is important to evaluate how quantum computation could be one potential avenue for development for future collider experiments. A wide variety of applications have been considered by different authors. We review here selected applications of quantum computing to high-energy physics, including topics in simulation, reconstruction, and the use of machine learning, and their challenges. In addition, recent advances in quantum computing technology to enhance such applications are briefly highlighted. Finally, we will discuss how such applications might transform the workflows of future collider experiments and highlight other potential applications.

Keywords: quantum computing, future collider experiments, quantum machine learning, quantum annealers, digital quantum computer, pattern recognition

OPEN ACCESS

Edited by:
Petra Merkel,
Fermi National Accelerator Laboratory (DOE), United States

Reviewed by:
Jean-Roch Vlimant,
California Institute of Technology, United States
Gabriel Nathan Perdue,
Fermi National Accelerator Laboratory (DOE), United States

***Correspondence:**
Heather M. Gray
heather.gray@berkeley.edu

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *28 January 2022*
Accepted: *03 May 2022*
Published: *27 May 2022*

Citation:
Gray HM and Terashi K (2022) Quantum Computing Applications in Future Colliders. Front. Phys. 10:864823. doi: 10.3389/fphy.2022.864823

1 INTRODUCTION

The fields of particle physics and computing have long been intertwined. The success of particle physics depends on the use of cutting-edge computing technology and in certain cases, the requirements of particle physics experiments have stimulated the development of new technologies in computing. Perhaps the most notable example is the development of the world wide web by physicists at the European Council for Nuclear Research (CERN), see [1], and other examples include the introduction of distributed grid computing in [2].

The state-of-the-art collider in particle physics is the Large Hadron Collider (LHC) [3], which is located at CERN just outside Geneva in Switzerland. The first proton-proton collisions in the LHC were recorded in 2010 and a key achievement was the discovery of the Higgs boson in 2012 by the two general purpose experiments, ATLAS and CMS. The third data-taking run of the LHC (Run 3) is planned to start in 2022 and to continue for four years. Run 3 is scheduled to be followed by a three year long shutdown during which the accelerator and experiments will undergo significant upgrades. The high-luminosity LHC (HL-LHC), as discussed in [4], is expected to start delivering proton-proton collisions in 2029.

The HL-LHC will deliver collisions at an instantaneous luminosity a factor of five to seven higher than the original LHC design luminosity. These collisions will increase the number of additional interactions, or pile up, by up to a factor of five to reach approximately 140–200 average pile up

interactions per bunch crossing. The experiments will include new detectors with more readout channels, which will increase the size of the recorded events by a factor of four to five. In addition, upgrades to the trigger systems will increase the event rate by up to an order of magnitude. These extensive upgrades herald the start of a decades long program in precision and discovery physics. At the same time, they place strong demands on computing.

Current projections for future computing budgets for the HL-LHC follow the so-called flat-budget scenario, in which only small increases to the budgets are foreseen to account for expectations from inflation. Extrapolations of the current computing model to the HL-LHC show a large deficit compared to the requirements [5–8]. In addition, computing for particle physics experiments has relied on Moore's Law. Moore's Law is the observation that the number of transistors on integrated circuits doubles every two years. However, over the past decade processor speeds have become limited by power density such that speed increases are due to increases in the number of cores rather than the speed of individual cores. In addition, the hardware landscape for computing processors has become increasingly heterogeneous. A range of development efforts are ongoing to explore how HEP software can be adapted to efficiently exploit heterogeneous computing architectures.

Although we have not yet entered the HL-LHC era, given the timescales required to build accelerators and detectors for particle physics experiments, the field is undergoing a series of international and national review processes to determine the future collider facilities to follow the HL-LHC. Although consensus has not yet been reached, many of the future colliders under consideration would make even more extensive demands on computing. One such collider, the Future Circular Collider in [9] would collide protons at a center-of-mass energy of 100 TeV with up to a thousand pile up interactions per bunch crossing.

The ideas of exploiting quantum mechanics to build a computer first began to be explored more than four decades ago. Initial ideas were focused on how a quantum computer could be used to simulate quantum mechanical systems. A decade later further interest was stimulated when quantum algorithms, which could be used to solve classically intractable problems, were introduced. One of the earliest of these, and one of the most famous, is Shor's algorithm for the factorization of prime numbers. At approximately the same time, the first quantum computers were built based on existing techniques from nuclear magnetic resonance. It is sometimes said that we are currently in the Noisy Intermediate-Scale Quantum (NISQ) era as introduced in [10]. Quantum computers in the NISQ era have orders of tens to hundreds physical qubits, and have been shown to surpass classical computers but only for specifically constructed problems. They also experience significant noise associated with the hardware and the electronics for qubit control.

There are two types of quantum computer: quantum annealers (QA) and circuit-based quantum computers. Quantum annealers are specifically designed to solve a single class of problem: minimization problems, and, in particular, to minimization problems that can be expressed as quadratic unconstrained binary optimization (QUBO) problems. Circuit-based quantum computers, on the other hand, can be programmed to execute more general quantum circuits, and are more similar in concept to the classical computers. Quantum computers use a number of different technologies for the qubits including superconducting transmon qubits, ion traps, photons and topological qubits and their current status is dicussed in [11, 12]. A typical state-of-the-art quantum annealer is the one produced by D-Wave with up to 5,000 qubits, while typical state-of-the-art circuit-based quantum computers are those from IBM and Google which have $\mathcal{O}(60-120)$ superconducting transmon qubits. Thanks to the extensive investment from governments and companies into the development of quantum computers, their capabilities have increased rapidly over the past years. See [13] for a recent review of the field of quantum computation.

Given the computational challenges faced by high-energy physics [14] now and in the future, and given the rapid development and exciting potential of quantum computers, it is natural to ask whether quantum computers can play a role in the future computing at HEP. Fault-tolerant quantum computers with sufficient number of qubits and gates are still decades away and very likely beyond the HL-LHC era, yet, future high-energy physics experiments will place even more extensive demands on computing. This article discusses selected studies exploring the application of quantum computers to HEP. In **Section 2** we discuss how quantum computers might be used to improve the quality of simulation for high-energy physics events. In **Section 3** we discuss one of the key computational challenges in HEP, the reconstruction of charged particle trajectories, and explore how quantum computing could be used for such problems. Finally, in **Section 4** we discuss the exciting field of machine learning with quantum computers and focus on applications of such techniques to physics analysis. **Section 5** concludes the article with a short summary and future outlook.

2 APPLICATIONS OF QUANTUM COMPUTING TO SIMULATION

One of the most promising applications for quantum computing is to simulate inherently quantum-mechanical systems, such as systems described by Quantum Field Theory (QFT) in particle physics. Quantum algorithms may perform particle scatterings in QFTs in polynomial resources using a universal quantum computer, as proposed by Jordan, Lee and Preskill in [15] and Preskill in [16]. Since then, a number of pioneering studies have been done in the contexts of simulation of particle systems, e.g., neutrino or neutral Kaon oscillations [17–19], heavy-ion collisions [20], parton distributions inside proton [21] as well as low-energy effective field theory [22] and quantum electrodynamics [23]. However, as represented by the Jordan-Lee-Preskill algorithm, a full QFT simulation generally requires prohibitively large resources and therefore cannot be implemented on near-term quantum computers. An

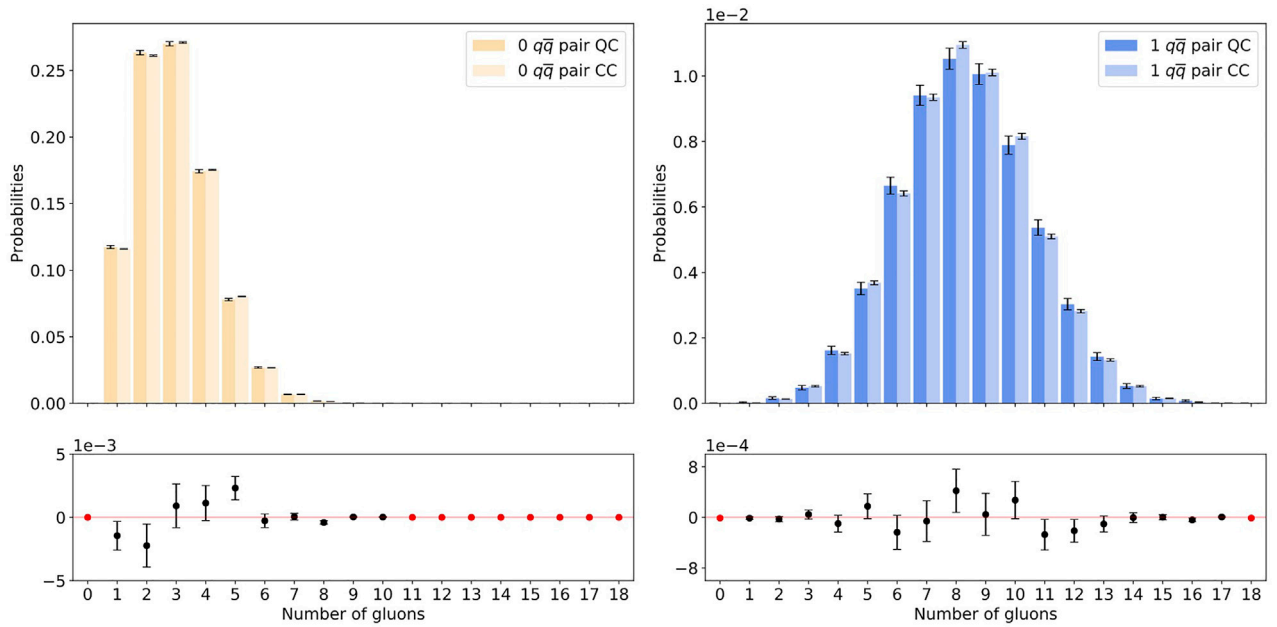


FIGURE 1 | Probability distributions of the number of gluons measured after a 31-step parton shower for the classical (CC) and quantum (QC) algorithms for the scenario where there are zero quark-antiquark pairs (left) and exactly one quark-antiquark pair (right) in the final state. From [25].

alternative approach for quantum computing applications of QFT is to break down particle scattering processes into pieces and exploit quantum computations in the place where conventional classical calculations are intractable.

In Monte Carlo simulations of high-energy hadron collisions, physics processes are treated by factorizing hard scatterings which occur at short distances and parton evolutions which occur at long distances inside hadrons. This factorization property allows the short- and long-range subprocesses to be calculated independently. The hard scattering between partons creates a large momentum transfer and produces a cascade of outgoing partons called the parton shower (PS). The PS simulation is a prototypical Markov Chain Monte Carlo (MCMC) simulation and describes the evolution of the system from the hard interaction to the hadronization scale. This technique for PS simulation has been successfully validated by comparing the MC simulation with experimental data, but quantum interference and correlation effects are neglected. Even though these effects are small compared to the current experimental accuracy, such limitations may become a bottleneck when the measurement will reach an unprecedented precision in the HL-LHC era.

Ref. [24] develops a quantum circuit that describes quantum properties of parton showers, in particular, the quantum interference that arises from different intermediate particles using a simplified QFT. They start with a system of n fermions that can have either one of two flavors, f_1 and f_2. These fermions can radiate a scalar particle ϕ, which itself can split into a $f\bar{f}$ pair of the same flavor or different flavors ($\bar{f}$ is an antifermion). In this setup, [24] demonstrates that quantum interference effects are successfully simulated using a quantum simulator and hardware, when the ϕ coupling to a $f_1\overline{f_2}$ or $\overline{f_1}f_2$ pair is turned on, due to unobserved intermediate states with mixed fermion flavors. This behavior is not properly captured by a classical MCMC simulation. It is also shown that the quantum circuit has capability to sample parton showers from a full probability distribution in polynomial time, which is otherwise difficult with the classical MCMC approach.

Another approach based on a quantum random walk is used in [25] to simulate parton showers. During a random walk, the movement of a particle termed *walker* is controlled by a coin flip operation that determines the direction the walker will move and a shift operation which moves the walker to the next position. The quantum analogue of the random walk performs these operations in a superposition of the basis states for the coin and shift operations, therefore allowing all possible shower histories to be generated when applied to the parton shower. In [25], the emission probabilities are controlled by the coin flip operations, and updating the shower content with a given emission corresponds to the shift operation of the walker in the position space. With this novel approach, [25] manages to simulate a collinear 31-step parton shower, implemented as a two-dimensional quantum walk with gluons and a quark-antiquark pair, as shown in **Figure 1**. This figure shows the probability distributions of the number of gluons after a 31-step parton shower for the classical and quantum algorithms, for the scenario where there are zero (left) and exactly one (right) quark-antiquark pair in the final state. A very good agreement is seen between the two algorithms. The quantum random walk approach demonstrates a significant improvement in the shower depth with fewer number of qubits from the algorithm in [24].

The calculation of the hard scattering also requires significant computational resources in the conventional techniques based on squaring the scattering amplitudes. [26] performs the calculation of hard interactions *via* helicity amplitudes by exploiting the equivalence between qubits and helicity spinors. This relies on the mapping between angles used to parameterize the helicity spinors and the qubit degrees of freedom in the Bloch sphere representation. The operators acting on the spinors are encoded as quantum circuits of unitary operators. With these helicity-qubit encodings, [26] demonstrates the construction of two quantum algorithms. The first is the helicity amplitude calculation in the $q \rightarrow qg$ process and the second are the helicity calculations for the *s*- and *t*-channel amplitudes of a $q\bar{q} \rightarrow q\bar{q}$ process. The algorithms highlight a unique advantage of quantum computations by exploiting a superposition of helicity qubits and allow a simultaneous calculation of the positive and negative helicities of each particle and the *s*- and *t*-channel amplitudes in the $2 \rightarrow 2$ process.

Despite recent improvements in the implementation of PS algorithms, running them on NISQ devices and simulating realistic parton showers involving many shower steps are currently challenging. This is largely due to the fact that the shower simulation is performed by repeating many times a circuit corresponding to a single PS step and this often results in a long circuit, which is hard to implement to NISQ device due to limited coherence time, qubit connectivity and hardware noise. One strategy to improve the performance is to mitigate errors through modifications to quantum state operation and measurement protocols. A number of readout and gate error mitigation techniques have been proposed in the literature, e.g., zero-noise extrapolation technique with identity insertions for gate errors, originally proposed in [27] and generalized in [28]. Another complementary strategy to error mitigation is to optimize the quantum circuits in the compilation process. A variety of architecture-agnostic and architecture-specific tools for circuit optimization have been developed, e.g., an industry standard tool called t|ket⟩ in [29] from Cambridge Quantum Computing. Ref. [30] introduces a new technique to optimize quantum circuits by identifying the amplitudes of computational basis states and removing redundant controlled gates in polynomial time with quantum measurement. This optimization protocol has been applied to the PS simulation in [24], together with the gate-error mitigation method in [28]. Ref. [30] successfully demonstrates that both the circuit optimization and the error mitigation methods can simplify the circuit significantly and improve the performance on NISQ device, depending on the initial states of the circuit corresponding to different initial particles of parton showers.

3 APPLICATIONS OF QUANTUM COMPUTING TO CHARGED PARTICLE PATTERN RECOGNITION

Raw data recorded by detectors is processed by reconstruction algorithms before it can be used for physics analyses. The track reconstruction algorithms used to reconstruct the trajectories of charged particles passing through the tracking detectors are typically the most computationally demanding. The required computing resources for such algorithms scale approximately quadratically with the number of charged particles per event, i.e. the amount of pile up, and therefore will become even more challenging at future colliders. Therefore, new ideas and approaches for track reconstruction algorithms are currently an extremely active field of research to ensure that the physics capabilities of the HL-LHC and beyond can be fully exploited.

Track reconstruction algorithms can be characterized into global and local approaches. Global algorithms process all the data, or hits in the detectors, from an event simultaneously and return a set of tracks. Local algorithms aim to identify the set of hits corresponding to a single track and are run many times to identify the full set of tracks. Examples of global methods include the Hough transform [31, 32] and neural networks [33]. The most widely used local method is the Kalman filter [34–36] and recently there has been extensive exploration into the use of Graph Neural Networks (GNNs) [37]. A number of different track reconstruction algorithms have been explored for quantum computers and, in most cases, the open dataset produced for the tracking machine learning challenge has been used [38, 39]. This dataset will be referred to as the TrackML dataset.

The first track reconstruction algorithm developed for quantum computers is a global track reconstruction algorithm as presented in [40]. The track reconstruction problem is formulated as a QUBO problem and quantum annealers are used to identify the global minimum. This algorithm, as is the case for all quantum algorithms discussed here, should be regarded as a hybrid quantum-classical algorithm because it requires pre- and post-processing on a classical computer.

The algorithm initially groups the hits in the detectors into doublets and then triplets. A QUBO is constructed from the triplets and the goal is to identify which pairs of triplets can be combined to form quadruplets. The weights in the QUBO depend on the compatibility between the properties of the triplets including their curvature and the angles between them, because triplets from the same track are expected to have identical properties. The QUBO is minimized on the quantum annealer by selecting the combinations of triplets compatible with the trajectories of charged particles. However, given the limited number of qubits available on quantum computers today, the QUBO is decomposed into smaller sub-QUBOs that are solved individually using fewer qubits. A software tool, called qbsolve [41], from D-Wave is used to perform this splitting and to recombine the solved sub-QUBOs so that the global minimum can be found. After minimization, a final post-processing step is performed on a classical computer to convert the accepted triplets back to doublets. Any duplicates or doublets with unresolved conflicts with other doublets are removed. The final track candidates are required to have at least five hits to reduce the contribution from random combinations of hits, or fakes.

The algorithm was studied using the TrackML dataset but restricting it to the central region of the detector, or barrel, which has a simpler geometry and less material and hence is a simpler problem for pattern recognition algorithms. In addition, events were filtered to select particular fractions of particles to emulate

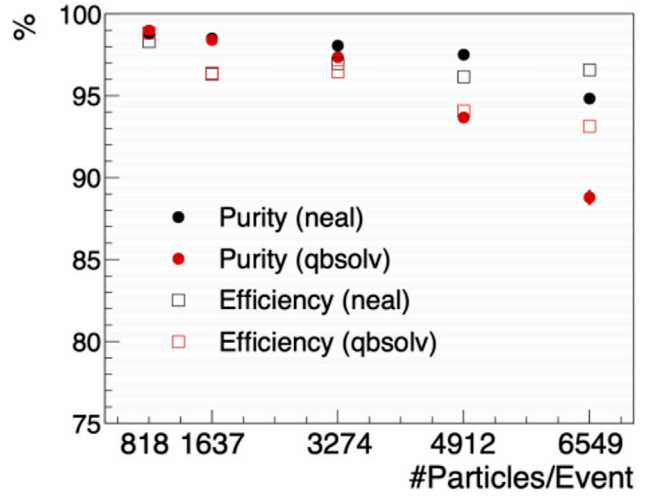


FIGURE 2 | The dependence of the efficiency (open squares) and purity (closed circles) of the quantum annealing pattern recognition algorithm on the number of charged particles per event. Results from the qbsolv solver (red) and the neal solver (black) are compared. From [42].

datasets with different amounts of pile up. This allows the dependence of the performance on the amount of pile up to be studied. The performance was studied using simulations of quantum annealers on Cori, a supercomputer located at the National Energy Research Scientific Computing Center (NERSC), and on quantum annealing hardware from D-Wave. Two different quantum annealers were used: the Ising D-Wave 2X located at Los Alamos National Laboratory and the D-Wave LEAP cloud service, which is an interface provided by D-Wave that allows users to run on a number of different quantum computers.

Key performance metrics for track reconstruction include the efficiency and the purity. The efficiency is the fraction of true particles that are successfully reconstructed and the purity is the fraction of reconstructed tracks that correspond to true particles. Ref. [40] showed that the efficiency is 100% in events with low pile up and it decreased to 90% in events with the level of pile up expected at the HL-LHC. However, while the purity is close to 100% at low pile up, it decreases rapidly with increasing pile up to reach only 50% at HL-LHC multiplicities. This demonstrates that the algorithm is impacted by fake tracks from random combinations of hits. The performance on the quantum hardware and in simulation is found to be consistent, which demonstrates that the algorithm is not significantly impacted by the noise on the quantum annealer.

Ref. [42] improves the performance of this quantum annealing pattern recognition algorithm by modifying the weights in the QUBO to include information about the impact parameters of the triplets. The impact parameters provide a measure of the distance of closest approach of tracks to the location where the proton-proton collision occurred. They are used by [42] to preferentially select tracks produced at the primary interaction point over secondaries produced in decays of primary particles or through interactions in the detector material. The efficiency and purity are shown in **Figure 2** using simulation as a function of the number of particles per event and results are shown for two different solving algorithms: qbsolv and neal. While the efficiency is slightly improved over the results shown in [40], the purity is dramatically improved to 85% or 95% depending on the solver at HL-LHC multiplicities. While this approach is very effective at reducing the fake rate, it is also expected to have a low efficiency to reconstruct tracks from secondary decays, such as B-hadrons and τ-leptons. The performance of neal is superior to qsolve in all cases.

[42] also studied the performance of annealing algorithms for pattern recognition using a digital annealing machine from Fujitsu. This is a quantum-inspired classical computer specifically formulated to solve annealing problems. Comparable physics performance was obtained between the quantum and digital annealers, however the computational time on the digital annealer was found to be far superior to the quantum annealer and to be essentially independent of the amount of pile up in the event.

A similar algorithm for pattern recognition on quantum annealers is presented in [43]. They also used triplets and the objective function depends on the angles between the triplets, a bias term to preferentially select high-momentum tracks and the point of origin of the tracks. There are penalty terms corresponding to bent and poorly oriented tracks.

The TrackML dataset was also used and tracks were reconstructed in both the barrel and the endcap of the detector. To ensure that the problem could be solved on quantum annealers available today, the dataset was split geometrically into 32 sectors and then sub-QUBOs were defined within each sector. Results were obtained both using simulation and a D-Wave 2X machine. In most regions of phase space, the efficiency was found to be approximately 90% and the purity greater than 95%.

Track reconstruction algorithms are also used in trigger detectors to select events which are subsequently processed in more detail for offline reconstruction. For such applications, algorithms need to be run in real time but can tolerate lower precision than offline reconstruction. One approach, which has been considered is the use of a memory bank of patterns of measurements. This allows track reconstruction algorithms to be replaced by looking up patterns in the memory bank. This means that the problem of the amount of processing power needed is transformed into memory requirements.

[44] explored how quantum computers, with their potential for exponential storage capacity, can be used for such problems. While the amount of classical memory required to store patterns depends exponentially on the number of elementary memory units, the amount of quantum memory depends only logarithmically, which means that far less memory is required. To demonstrate this, they presented an implementation of a quantum associative memory protocol (QuAM) in [45] for 2-bit patterns circuit-based quantum computers using the IBM Qiskit framework. They used Trugenberger's algorithm to store the patterns and the generalized Grover's algorithm to retrieve the patterns.

They explored IBM circuit-based quantum computers with up to 14 qubits, however, these did not provide sufficient qubits to solve 2-bit patterns.

An additional approach for charged particle pattern recognition, relying on quantum machine learning, a Quantum Graph Neural Network (QGNN) will be discussed in **Section 4.2**.

4 APPLICATIONS OF QUANTUM MACHINE LEARNING

Recent developments of quantum computing architecture and algorithms make Quantum Machine Learning (QML) a promising early application in NISQ era. QML includes a wide range of research topics, e.g., information theory aspects such as quantum learning complexity, accuracy and asymptotic behavior in a fault-tolerant regime, as well as more near-term aspects such as data encoding, learning circuit models and hybrid architectures with classical calculations. An important aspect of QML for early HEP applications [46] is how effectively one can exploit machine learning to characterize a quantum state generated from input data in a quantum-classical combined setting. The quantum-classical hybrid approach is particularly useful for near-term quantum devices because the quantum part can focus on classically-intractable calculations while the classical part handles the rest of the computation including, e.g., optimization of the learning model. Classical machine learning is used extensively in the analysis of HEP experimental data. The significant growth of the data volume foreseeable for the HL-LHC era and beyond further motivates quantum-empowered machine learning in the workflow of future HEP experiments.

A number of experimental applications of QML to data analysis and event reconstruction in representative HEP experiments have been explored. They are categorized into two approaches based on quantum annealing and quantum circuit models, each having its own advantages and disadvantages in terms of the maturity and applicability. Most QML applications in HEP (discussed below) belong to supervised machine learning, but several applications (mentioned explicitly) use unsupervised learning techniques.

4.1 Quantum Machine Learning With Quantum Annealing

QA-based machine learning (QAML) formulates ML as a QUBO problem (see **Section 3**) and looks for the best classification. This is done by minimizing the cost function that quantifies the difference between the predictions and the true labels. In [47], QAML is applied to the classification of Higgs events where the Higgs boson decays into two photons, one of the key channels used for the Higgs discovery. From eight high-level kinematic features in a set of training events $\{\boldsymbol{x}_\tau, \boldsymbol{y}_\tau\}$ with τ being the event index, 36 *weak* classifiers $c_i(\boldsymbol{x}_\tau)$ are constructed (with i being the classifier index) in such a way that the signal (background) events populate at positive (negative) values within the range between −1 and 1. The term *weak* indicates that these classifiers are constructed from various combinations of arithmetic operations on the kinematic features. A *strong* classifier is then constructed as a linear combination of the weak classifiers with binary coefficients $w_i \in \{0, 1\}$. The objective function is defined as:

$$O(\boldsymbol{w}) = \sum_{i,j} C_{ij} w_i w_j + \sum_i (\lambda - 2C_i) w_i, \tag{1}$$

where $C_{ij} = \sum_\tau c_i(\boldsymbol{x}_\tau) c_j(\boldsymbol{x}_\tau)$, $C_i = \sum_\tau c_i(\boldsymbol{x}_\tau) \boldsymbol{y}_\tau$ with the true label $\boldsymbol{y}_\tau \in \{+1, -1\}$, and λ is a positive parameter to penalize the number of non-zero weight, w_i, terms. The objective function O is finally transformed into the problem Hamiltonian by converting the binary w_i to spin variable $\sigma_z^i \in \{-1, 1\}$ for quantum annealing.

QAML performance for Higgs classification in two-photon events is compared in [47] with two classical ML techniques based on Boosted Decision Tree (BDT) and Deep Neural Network (DNN). In addition, the same problem Hamiltonian is adopted to Simulated Annealing (SA), a classical analogue of QA with energy fluctuations controlled by artificial temperature variables. The results show that the QAML and SA have similar classification performance in terms of the area under receiver operating characteristic (ROC) curve for the true-positive and false-positive rates, which correspond to the signal and background efficiencies, respectively. The annealing-based methods have no clear advantage over the classical ML methods, though a hint of slight advantage is seen at small dataset size. The QAML method has been further extended in [48] to the so-called QAML-Z, that aims to zoom into the energy surface to optimize real-valued coefficients and sequentially apply QA to an augmented set of weak classifiers. The QAML-Z method is applied in [49] to investigate the selection of Supersymmetric top (stop) quark events against standard model background events.

4.2 Quantum Machine Learning With Quantum Circuits

There are two widely-used ML implementations for near-term gate-based quantum computers: Variational Quantum Algorithms (VQA) in [50] and Quantum Support Vector Machines (QSVM) with kernel method in e.g., [51] and [52]. For both approaches, the input classical data $\boldsymbol{x} \in \mathbb{R}^d$ is embedded in a Hilbert space with a unitary operator $U(x_i)$ to create a quantum state $|\phi(x_i)\rangle = U(x_i)|0\rangle^{\otimes n}$, where n is the number of qubits. Since the NISQ devices have only limited resources, simplified problems have been considered for both approaches with the circuit-based machine learning.

VQA uses an ansatz created using unitary operator $V(\boldsymbol{\theta})$ with tunable parameters $\boldsymbol{\theta}$ and measures the produced final state $V(\boldsymbol{\theta})|\phi(x_i)\rangle$ with certain Hermitian operators. A cost function is defined from the measurement outcome and the input ground truth, and the parameter optimization or *training* is performed classically by minimizing the cost function. The minimization of cost function is an important subject in the field, and one of the representative methods exploits the gradients of the cost

function with respect to the $\boldsymbol{\theta}$ parameters, which is so-called the parameter-shift rule [53–56]. This VQA approach has a wide range of applications including classification, regression and optimization.

QSVM performs classification in high-dimensional feature space by constructing the kernel function from the inner products of embedded quantum states $K(x_i, x_j) = |\langle\phi(x_i)|\phi(x_j)\rangle|^2$. The kernel function is sampled by estimating the probability of measuring the $|0\rangle^{\otimes n}$ state from the circuit and is used by the classical SVM to build a hyperplane for data classification. QSVM with a kernel function often becomes more efficient than VQA in the training stage because the parameter tuning is performed fully classically with the obtained kernel function for the SVM. A polynomial number of quantum gates used in VQA and QSVM, acting on a quantum state that spans a Hilbert space exponentially, could bring quantum advantage with near-term NISQ devices in the future.

VQA-based HEP applications have largely focused on the classification of physics process of interests so far. They include the classifications of Higgs boson [57, 58], supersymmetric particles with missing transverse momentum [59], neutrino events [60], new resonances in di-top final states [61] and $B\bar{B}$ production in e^+e^- collisions [62].

[57] studies the classification of $t\bar{t}H$ events with $H \to \gamma\gamma$ decays and $H \to \mu^+\mu^-$ events from the dominant background processes for each channel using VQA. After reducing the kinematic variables into 10 variables using a principal component analysis, [57] encodes the input data as angles of R_Z rotation gates, and produces an ansatz composed of R_Y and R_Z rotation gates with parameters and an entangling layer of controlled-Z gates between adjacent qubits. Binary classification is performed with the VQA framework using both simulation and quantum hardware. The results show that the VQA approach has a discrimination power similar to the classical SVM and BDT, and the performance with quantum device is comparable to that of a quantum simulator.

Ref. [59] explores the classification of the production of SUSY events with two $W \to l\nu$ decays and two neutralinos against a SM background process of two $W \to l\nu$ production. Restricting the input features to 3, 5 and 7 kinematic variables, two different implementations of VQA with different configurations of data encoding and ansatz circuits are investigated using a quantum simulator and quantum hardware. As either the number of input variables or training events in the sample increase, the classical techniques using BDTs and DNNs outperform the VQA classifiers in the simulator. However, the VQA classifiers have comparable performance when the training sample size or the number of input variables is small.

In general, it is interesting to investigate the impact of data encoding on VQA performance, or more specifically if there is any efficient way to encode data from HEP experiments. [62] attempts to address this question in the classification of $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ events with $B \to K^+K^-$ decays from $e^+e^- \to q\bar{q}$ background events. The momenta of the final state particles (input data $\boldsymbol{x}$) are encoded into the quantum states $|\psi(\boldsymbol{x})\rangle$. The kernel $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = |\langle\psi(\boldsymbol{x}_i)|\psi(\boldsymbol{x}_j)\rangle|^2$ is derived by measuring the output state, which feeds into the classical SVM for the classification. In the encoding stage, input particle momenta are encoded by maintaining the kinematic properties of individual particles with a dense-encoding technique. An example data encoding gate proposed by the authors is shown in **Figure 3**. In this encoding, individual particles are encoded first with their momentum variables (p, θ, ϕ) in spherical coordinates, then each particle is entangled with the other in the event through one of their qubits. [62] demonstrates that the classification performance is improved by exploiting this encoding scheme for the classification of $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ events.

Ref. [58] considers the classification of $t\bar{t}H$ with $H \to b\bar{b}$ decays using 67 input features. The authors first perform feature reduction to adapt the large number of input features to the capacity of current quantum computation with techniques based on auto-encoding and feature selection. The VQA and QSVM performance are compared with a DNN and a BDT. The Area Under Curve (AUC) values of the quantum classifiers are similar to those from the DNN and BDT, indicating that the performance is similar.

In the absence of a clear new physics signal in HEP experiments, a data-driven, signal-agnostic search has gained considerable interests over years. Unsupervised machine learning for new physics searches, often realized as anomaly detection, will be an important tool in future colliders. Several pioneering studies with unsupervised QML technique have been done in the contexts of anomaly detection. Ref. [63] attempts to combine an anomaly detection technique with a graph representation of HEP events through quantum computation. In particular, the continuous-variable (CV) model of quantum computer, programmed using photonic quantum device, is exploited to survey graph-represented events to search for $pp \to HZ$ signal. Ref. [63] uses Gaussian boson sampling of CV events as input into a quantum variant of K-means clustering algorithm for anomaly detection. A discrete (qubit-based) QML approach for unsupervised learning is also investigated in [64], where the feasibility of anomaly detection is explored using quantum autoencoders (QAE) based on variational quantum circuits. With the benchmark process of $pp \to H \to t\bar{t}$ for signal, the QAE performance has been compared with that from a classical autoencoder for anomaly detection.

Various classical ML techniques have been applied to the tasks of detector simulation and reconstruction in HEP experiments (see the review in HEP ML Community). Among them, the Generative Adversarial Network (GAN) is extensively studied to simulate calorimeter energy deposits of particle showers as images, aiming for increased speed compared to a full Geant4-based detector simulation. The quantum version of a GAN (QGAN) and its variants have been proposed and investigated for HEP detector simulation [65] and Monte Carlo event generation [66]. Ref. [67] employs the Parameterized Quantum Circuit (PQC) GAN model, which is composed of a quantum Generator and a classical Discriminator to demonstrate a proof-of-concept for the QGAN-based shower simulation. In particular, the GAN model in [67] uses two PQCs to sample two probability distributions, one for the shower images and the other for normalized pixel intensities in a single image. This allows a

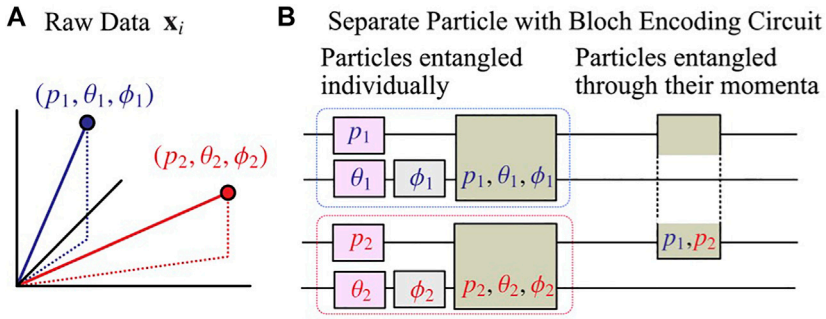


FIGURE 3 | (A) The raw data $\boldsymbol{x}_i$ containing two particles with their momenta denoted by (p_i, θ_i, ϕ_i) in spherical coordinates (i = 1, 2). **(B)** A proposed data encoding composed of individual particle encoding with 2 qubits (each surrounded by a dashed box) and entangling between particles through their momentum bits. From [62].

full training set of images to be captured. Even though the image size is restricted only to 2 × 2 pixels, the proposed Dual-PQC GAN model manages to generate a sample of individual images and their probability distributions consistent with those in the training set.

A technique from machine learning that has been applied to the problem of charged particle pattern recognition are GNNs. Ref. [68, 69] explored how GNNs can be applied to the problem of pattern recognition. They developed a hybrid quantum-classical algorithm, called a Quantum GNN (QGNN) that relies on a series of quantum edge and quantum node networks. After iterating through the series, a final edge network obtains the final segment classification. [69] showed that the QGNN performs similarly to classical methods when the number of hidden dimensions is low. They studied the scaling performance of the QGNN in simulation by varying the number of hidden dimensions and qubits. The QGNN performance improves with additional hidden dimensions, however, some saturation of the best loss is observed. The results were limited to simulation due to the current status of NISQ hardware. The application of this technique to the LUXE experiment is explored in [70].

5 CONCLUSION AND OUTLOOK

Quantum computing has been undergoing explosive progress recently and may have the potential to provide solutions to computing challenges in HEP. There are many potential applications and we have reviewed algorithms for quantum simulation, quantum pattern recognition and quantum machine learning and their challenges.

QFT simulation with quantum computers is highly motivated because universal quantum computers can perform such simulation with exponentially smaller computing resources than classical computers. However, it would require many more physical qubits than are currently available. Several representative studies focusing on a simplified QFT, in particular parton showers and hard scatterings, have been reviewed here. These studies considered simple cases compared to the realistic MC simulations currently employed in HEP experiments, but demonstrate that interesting quantum properties can be captured even with current devices. A full QFT simulation envisioned by Jordan, Lee and Preskill will certainly require many logical qubits free from errors, composed of millions of physical qubits, called Fault-Tolerant Quantum Computers (FTQC). The realization of FTQC is however still decades away. A near-term goal of quantum simulation is to develop quantum algorithms for each QFT simulation step and evaluate the feasibility to realize potential advantage over classical simulation with near-term technologies.

Pattern recognition algorithms are among the most computationally demanding algorithms for HEP. A number of studies have focused on global algorithms using quantum annealers. These are typically hybrid quantum-classical algorithms with pre- and post-processing performed on classical computers and finding the solution on a quantum annealer. All algorithms were able to obtain excellent efficiencies, however some reported significant rates of fake tracks particularly in very busy events. In all cases, the problems needed to be simplified to be able to run on NISQ devices. A study on the use of quantum circuits as associated memory for triggers was presented, which would exploit the exponential storage capacity of quantum computers. In the future, it would be interesting to explore local pattern recognition algorithms, particularly on circuit-based computers. One of the main challenges for such algorithms is that they require large amounts of data to be processed on the quantum device, and the transfer of this data may limit the speed of such algorithms.

Quantum machine learning is a promising early application in the NISQ era, not only for HEP but also for other scientific domains and industries. Significant progress has been made over the past few years in aspects including data encoding, ansatz design and trainability for VQA and kernel-based learning architecture. The QAML technique has been employed for the Higgs-boson classification and demonstrated that the classification performance is comparable to the classical approaches based on BDT and DNN. A number of VQA- and QSVM-based classification studies have been demonstrated in representative data analyses for pp and e^+e^- collisions. The quantum classifiers have comparable performance to the conventional BDT and/or DNN methods, meaning that quantum advantage has not

been demonstrated yet. Anomaly detection using unsupervised QML methods is being explored. The QML approach has also been investigated in the detector simulation as well, with the demonstration of reproducing a calorimeter shower of energy deposits in a simplified setting. QML applications to HEP data analysis will pose several challenges in the future. A large-volume, high-dimensional data produced from HEP experiments needs to be embedded into quantum states, which requires significant computing resources. The encoded states are then processed in VQA by a parameterized quantum circuit to extract underlying properties in the data. Recent studies show that the ansatz needs to have sufficient expressibility for a given problem but should not be too expressive, making it exponentially harder to compute the gradients of cost function with the number of qubits. This infamous problem, so-called barren plateau in the cost function landscape, and the data encoding are currently very active areas of research for QML. It will be crucial to explore them further in the context of HEP-oriented QML architectures towards realizing quantum advantage for QML in HEP.

AUTHOR CONTRIBUTIONS

The main contributions from HG were **sections 1** and **3** and the main contributions from KT were **sections 2** and **4**.

FUNDING

HG has received funding from the Alfred P. Sloan Foundation under grant no. FG-2020-13732. KT has received funding from the U.S.-Japan Science and Technology Cooperation Program in High Energy Physics under the project "Optimization of HEP Quantum Algorithms" (FY2021-23).

REFERENCES

1. Berners-Lee TJ, Cailliau R, Groff JF, Pollermann B. World Wide Web: An Information Infrastructure for High-Energy Physics. *Conf Proc C* (1992) 920: 157–64.
2. Foster I, Kesselman C, Tuecke S. The Anatomy of the Grid - Enabling Scalable Virtual Organizations. *The International Journal of High Performance Computing Applications* (2001). 15 3:200–222. doi:10.1177/109434200101500302
3. Bruening OS, Collier P, Lebrun P, Myers S, Ostojic R, Poole J, et al. *LHC Design Report. Tech. Rep. CERN-2004-003-V-1*. Geneva: CERN Yellow Reports: Monographs (2004). doi:10.5170/CERN-2004-003-V-1
4. Apollinari G, Béjar Alonso I, Brüning O, Fessia P, Lamont M, Rossi L, et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*. Geneva: Tech. Rep. CERN Yellow Reports: Monographs (2017).
5. ATLAS Collaboration. *ATLAS Software and Computing HL-LHC Roadmap. Tech. Rep*. Geneva: CERN (2022).
6. Software CO Computing. *Evolution of the CMS Computing Model towards Phase-2Tech. Rep*. Geneva: CERN (2021).
7. Shadura O, Krikler B, Stewart GA. HL-LHC Computing Review Stage 2, Common Software Projects: Data Science Tools for Analysis. In: J Pivarski, E Rodrigues, K Pedro, editors. *LHCC Review of HL-LHC Computing* (2022).
8. Yazgan E. *HL-LHC Computing Review Stage-2*. Common Software Projects: Event Generators (2021). doi:10.48550/arXiv.2109.14938
9. Abada A. FCC-hh: The Hadron Collider: Future Circular Collider Conceptual Design Report Volume 3. *Eur Phys J ST* (2019) 228:755–1107. doi:10.1140/epjst/e2019-900087-0
10. Preskill J. Quantum Computing in the NISQ Era and beyond. *Quantum* (2018) 2:79. doi:10.22331/q-2018-08-06-79
11. Bruzewicz CD, Chiaverini J, McConnell R, Sage JM. Trapped-ion Quantum Computing: Progress and Challenges. *Appl Phys Rev* (2019) 6:021314. doi:10.1063/1.5088164
12. Kjaergaard M, Schwartz ME, Braumüller J, Krantz P, Wang JI-J, Gustavsson S, et al. Superconducting Qubits: Current State of Play. *Annu Rev Condens Matter Phys* (2020) 11:369–95. doi:10.1146/annurev-conmatphys-031119-050605
13. Preskill J. *Quantum Computing 40 Years Later* (2021). doi:10.48550/arXiv.2106.10522
14. [Dataset] HEP ML Community. *A Living Review of Machine Learning for Particle Physics* (2021). doi:10.48550/arXiv.2102.02770
15. Jordan SP, Lee KSM, Preskill J. Quantum Algorithms for Quantum Field Theories. *Science* (2012) 336:1130–3. doi:10.1126/science.1217069
16. Preskill J. *Simulating Quantum Field Theory with a Quantum Computer* (2018). doi:10.48550/arXiv.1811.10085
17. Argüelles CA, Jones BJP. Neutrino Oscillations in a Quantum Processor. *Phys Rev Res* (2019) 1:033176. doi:10.1103/physrevresearch.1.033176
18. Jha AK, Chatla A, Bamba BA. *Quantum Simulation of Oscillating Neutrinos* (2020). doi:10.48550/arXiv.2010.06458
19. Chang SY. *Quantum Generative Adversarial Networks in a Continuous-Variable Architecture to Simulate High Energy Physics Detectors* (2021). doi:10.48550/arXiv.2101.11132
20. de Jong WA, Metcalf M, Mulligan J, Płoskoń M, Ringer F, Yao X. Quantum Simulation of Open Quantum Systems in Heavy-Ion Collisions. *Phys Rev D* (2021) 104:L051501. doi:10.1103/physrevd.104.l051501
21. Pérez-Salinas A, Cruz-Martinez J, Alhajri AA, Carrazza S. Determining the Proton Content with a Quantum Computer. *Phys Rev D* (2021) 103. doi:10.1103/physrevd.103.034027
22. Bauer CW, Nachman B, Freytsis M. Simulating Collider Physics on Quantum Computers Using Effective Field Theories. *Phys Rev Lett* (2021) 127. doi:10.1103/physrevlett.127.212001
23. Stetina TF, Ciavarella A, Li X, Wiebe N. Simulating Effective Qed on Quantum Computers. *Quantum* (2022) 6:622. doi:10.22331/q-2022-01-18-622
24. Nachman B, Provasoli D, de Jong WA, Bauer CW. Quantum Algorithm for High Energy Physics Simulations. *Phys Rev Lett* (2021) 126:062001. doi:10.1103/PhysRevLett.126.062001
25. Williams S, Malik S, Spannowsky M, Bepari K. *A Quantum Walk Approach to Simulating Parton Showers* (2021). doi:10.48550/arXiv.2109.13975
26. Bepari K, Malik S, Spannowsky M, Williams S. Towards a Quantum Computing Algorithm for Helicity Amplitudes and Parton Showers. *Phys Rev D* (2021) 103:076020. doi:10.1103/PhysRevD.103.076020
27. Endo S, Benjamin SC, Li Y. Practical Quantum Error Mitigation for Near-Future Applications. *Phys Rev X* (2018) 8:031027. doi:10.1103/PhysRevX.8.031027
28. He A, Nachman B, de Jong WA, Bauer CW. Zero-noise Extrapolation for Quantum-Gate Error Mitigation with Identity Insertions. *Phys Rev A* (2020) 102:012426. doi:10.1103/physreva.102.012426
29. Sivarajah S, Dilkes S, Cowtan A, Simmons W, Edgington A, Duncan R. t|ket⟩: a Retargetable Compiler for NISQ Devices. *Quan Sci. Technol.* (2020) 6:014003. doi:10.1088/2058-9565/ab8e92
30. Jang W, Terashi K, Saito M, Bauer CW, Nachman B, Iiyama Y, et al. Quantum Gate Pattern Recognition and Circuit Optimization for Scientific Applications. *EPJ Web Conf* (2021) 251:03023. doi:10.1051/epjconf/202125103023

31. Hough P. Machine Analysis of Bubble Chamber Pictures. *Conf Proc C* (1959) 590914:554–8. Available at: https://cds.cern.ch/record/107795?ln=en
32. Duda RO, Hart PE. Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Commun ACM* (1972) 15:11–5. doi:10.1145/361237.361242
33. Hopfield JJ. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proc Natl Acad Sci U.S.A* (1982) 79:2554–8. doi:10.1073/pnas.79.8.2554
34. Kalman RE. A New Approach to Linear Filtering and Prediction Problems. *J Basic Eng* (1960) 82:35–45. doi:10.1115/1.3662552
35. Billoir P. Track Fitting with Multiple Scattering: A New Method. *Nucl Instr Methods Phys Res* (1984) 225:352–66. doi:10.1016/0167-5087(84)90274-6
36. Frühwirth R. Application of Kalman Filtering to Track and Vertex Fitting. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1987) 262:444–50. doi:10.1016/0168-9002(87)90887-4
37. Duarte J, Vlimant J-R. Graph Neural Networks for Particle Tracking and Reconstruction. *arxiv* (2022) 387–436. doi:10.1142/9789811234033_0012
38. Amrouche S, Basara L, Calafiura P, Estrade V, Farrell S, Ferreira DR, et al. The Tracking Machine Learning challenge: Accuracy Phase. *The NeurIPS* (2020) 18:231–64. doi:10.1007/978-3-030-29135-8_9
39. Amrouche S. and others. *The Tracking Machine Learning challenge : Throughput Phase* (2021). doi:10.48550/arXiv.2105.01160
40. Bapst F, Bhimji W, Calafiura P, Gray H, Lavrijsen W, Linder L, et al. A Pattern Recognition Algorithm for Quantum Annealers. *Comput Softw Big Sci* (2020) 4:1. doi:10.1007/s41781-019-0032-5
41. Booth M, Reinhardt S, Roy A. *Partitioning Optimization Problems for Hybrid Classical/quantum Execution. Tech. Rep*. dWave Technical Report (2017). Available at: https://docs.ocean.dwavesys.com/projects/qbsolv/en/latest/_downloads/bd15a2d8f32e587e9e5997ce9d5512cc/qbsolv_techReport.pdf
42. Saito M, Calafiura P, Gray H, Lavrijsen W, Linder L, Okumura Y, et al. Quantum Annealing Algorithms for Track Pattern Recognition. *EPJ Web Conf* (2020) 245:10006. doi:10.1051/epjconf/202024510006
43. Zlokapa A, Anand A, Vlimant JR, Duarte JM, Job J, Lidar D, et al. *Charged Particle Tracking with Quantum Annealing-Inspired Optimization* (2019). doi:10.48550/arXiv.1908.04475
44. Shapoval I, Calafiura P. Quantum Associative Memory in HEP Track Pattern Recognition. *EPJ Web Conf* (2019) 214:01012. doi:10.1051/epjconf/201921401012
45. Ventura D, Martinez T. Quantum Associative Memory with Exponential Capacity. 1998 IEEE International Joint Conference on Neural Networks Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98CH36227) (1998) 11:509–13. doi:10.1109/IJCNN.1998.682319
46. Guan W, Perdue G, Pesah A, Schuld M, Terashi K, Vallecorsa S, et al. Quantum Machine Learning in High Energy Physics. *Mach Learn Sci Technol* (2021) 2:011003. doi:10.1088/2632-2153/abc17d
47. Mott A, Job J, Vlimant J-R, Lidar D, Spiropulu M. Solving a Higgs Optimization Problem with Quantum Annealing for Machine Learning. *Nature* (2017) 550:375–9. doi:10.1038/nature24047
48. Zlokapa A, Mott A, Job J, Vlimant J-R, Lidar D, Spiropulu M. Quantum Adiabatic Machine Learning by Zooming into a Region of the Energy Surface. *Phys Rev A* (2020) 102:062405. doi:10.1103/PhysRevA.102.062405
49. Bargassa P, Cabos T, Cordeiro Oudot Choi AA, Hessel T, Cavinato S. Quantum Algorithm for the Classification of Supersymmetric Top Quark Events. *Phys Rev D* (2021) 104:096004. doi:10.1103/PhysRevD.104.096004
50. Cerezo M, Arrasmith A, Babbush R, Benjamin SC, Endo S, Fujii K, et al. Variational Quantum Algorithms. *Nat Rev Phys* (2021) 3:625–44. doi:10.1038/s42254-021-00348-9
51. Havlíček V, Córcoles AD, Temme K, Harrow AW, Kandala A, Chow JM, et al. Supervised Learning with Quantum-Enhanced Feature Spaces. *Nature* (2019) 567:209–12. doi:10.1038/s41586-019-0980-2
52. Schuld M, Killoran N. Quantum Machine Learning in Feature hilbert Spaces. *Phys Rev Lett* (2019) 122:040504. doi:10.1103/PhysRevLett.122.040504
53. Gian Giacomo Guerreschi MS. *Practical Optimization for Hybrid Quantum-Classical Algorithms* (2017). doi:10.48550/arXiv.1701.01450
54. Mitarai K, Negoro M, Kitagawa M, Fujii K. Quantum Circuit Learning. *Phys Rev A* (2018) 98:032309. doi:10.1103/PhysRevA.98.032309
55. Schuld M, Bergholm V, Gogolin C, Izaac J, Killoran N. Evaluating Analytic Gradients on Quantum Hardware. *Phys Rev A* (2019) 99:032331. doi:10.1103/PhysRevA.99.032331
56. Mari A, Bromley TR, Killoran N. Estimating the Gradient and Higher-Order Derivatives on Quantum Hardware. *Phys Rev A* (2021) 103:012405. doi:10.1103/PhysRevA.103.012405
57. Wu SL, Chan J, Guan W, Sun S, Wang A, Zhou C, et al. Application of Quantum Machine Learning Using the Quantum Variational Classifier Method to High Energy Physics Analysis at the LHC on IBM Quantum Computer Simulator and Hardware with 10 Qubits. *J Phys G: Nucl Part Phys* (2021) 48:125003. doi:10.1088/1361-6471/ac1391
58. Belis V, González-Castillo S, Reissel C, Vallecorsa S, Combarro EF, Dissertori G, et al. Higgs Analysis with Quantum Classifiers. *EPJ Web Conf* (2021) 251:03070. doi:10.1051/epjconf/202125103070
59. Terashi K, Kaneda M, Kishimoto T, Saito M, Sawada R, Tanaka J. Event Classification with Quantum Machine Learning in High-Energy Physics. *Comput Softw Big Sci* (2021) 5:2. doi:10.1007/s41781-020-00047-7
60. Chen SYC. Quantum Convolutional Neural Networks for High Energy Physics Data Analysis. *Phys. Rev. Res.* 4, 013231. doi:10.1103/PhysRevResearch.4.013231 (2020).
61. Blance A, Spannowsky M. Quantum Machine Learning for Particle Physics Using a Variational Quantum Classifier. *J High Energ Phys* (2021) 2021:212. doi:10.1007/jhep02(2021)212
62. Heredge J, Hill C, Hollenberg L, Sevior M. *Quantum Support Vector Machines for Continuum Suppression in B Meson Decays* (2021). doi:10.48550/arXiv.2103.12257
63. Blance A, Spannowsky M. Unsupervised Event Classification with Graphs on Classical and Photonic Quantum Computers. *J High Energ Phys* (2021) 2021. doi:10.1007/jhep08(2021)170
64. Ngairangbam VS, Spannowsky M, Takeuchi M. *Anomaly Detection in High-Energy Physics Using a Quantum Autoencoder* (2021). doi:10.48550/arXiv.2112.04958
65. Chang SY, Herbert S, Vallecorsa S, Combarro EF, Duncan R. Dual-parameterized Quantum Circuit gan Model in High Energy Physics. *EPJ Web Conf* (2021) 251:03050. doi:10.1051/epjconf/202125103050
66. Bravo-Prieto C. *Style-based Quantum Generative Adversarial Networks for Monte Carlo Events* (2021). doi:10.48550/arXiv.2110.06933
67. Ying C, Yunheng SZ. *Quantum Simulations of the Non-unitary Time Evolution and Applications to Neutral-Kaon Oscillations* (2021).
68. Tüysüz C, Carminati F, Demirköz B, Dobos D, Fracas F, Novotny K, et al. A Quantum Graph Neural Network Approach to Particle Track Reconstruction (2020). doi:10.5281/zenodo.4088474
69. Tüysüz C, Rieger C, Novotny K, Demirköz B, Dobos D, Potamianos K, et al. Hybrid Quantum Classical Graph Neural Networks for Particle Track Reconstruction. *Quan Mach. Intell.* (2021) 3:29. doi:10.1007/s42484-021-00055-9
70. Funcke L, Hartung T, Heinemann B, Jansen K, Kropf A, Kühn S. Studying Quantum Algorithms for Particle Track Reconstruction in the LUXE experiment. In: 20th International Workshop on Advanced Computing and Analysis Techniques in Physics Research: AI Decoded - Towards Sustainable, Diverse, Performant and Effective Scientific Computing (2022).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer GNP declared a past collaboration with one of the authors KT to the handling editor.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Gray and Terashi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

ORIGINAL RESEARCH
published: 09 June 2022
doi: 10.3389/fphy.2022.888078

Future Circular Collider: Integrated Programme and Feasibility Study

Michael Benedikt and Frank Zimmermann*

CERN, Geneva, Switzerland

The Future Circular Collider (FCC) Integrated Project foresees, in a first stage, a high-luminosity high-energy electron-positron collider, serving as Higgs, top and electroweak factory, and, in a second stage, an energy frontier hadron collider, with a centre-of-mass energy of at least 100 TeV. This programme well matches the highest priority future requests issued by the 2020 Update of the European Strategy for Particle Physics. In 2021, with the support of the CERN Council, a five-year FCC Feasibility Study was launched. In this article, we present the FCC integrated project and the preparations for the FCC Feasibility Study.

Keywords: hadron collider, lepton collider, future circular collider, European strategy for particle physics, Higgs factory, electroweak factory, top-quark factory, TeraZ factory

OPEN ACCESS

Edited by:
Alessandro Tricoli, Brookhaven National Laboratory (DOE), United States

Reviewed by:
Nadia Pastrone, Istituto Nazionale Fisica Nucleare, Italy
Jaime Hernández-Sánchez, Benemérita Universidad Autónoma de Puebla, Mexico

****Correspondence:***
Michael Benedikt
michael.benedikt@cern.ch

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *02 March 2022*
Accepted: *22 March 2022*
Published: *09 June 2022*

Citation:
Benedikt M and Zimmermann F (2022) Future Circular Collider: Integrated Programme and Feasibility Study. Front. Phys. 10:888078. doi: 10.3389/fphy.2022.888078

1 FCC INTEGRATED PROJECT

1.1 Overview

The Future Circular Collider (FCC) shall be located in the Lake Geneva basin and linked to the existing CERN facilities [4]. The FCC "integrated programme" is inspired by the successful past Large Electron Positron collider (LEP) and Large Hadron Collider (LHC) projects at CERN. It represents a comprehensive long-term programme maximising physics opportunities. A similar project is under study in China [25, 32].

The first stage of the FCC integrated project is an e^+e^- collider, called FCC-ee, which would serve as Higgs factory, electroweak and top factory at highest luminosities, and run at four different centre-of-mass energies, namely on the Z pole, at the WW threshold, at the ZH production peak, and at the $t\bar{t}$ threshold. In a second stage, the FCC-ee would be followed by a highest-energy proton collider, FCC-hh with a centre-of-mass energy of 100 TeV, that would naturally succeed the LHC at the energy frontier. This hadron collider can also accommodate ion and lepton-hadron collision options, providing for complementary physics. The lepton and hadron colliders would profit from a common civil engineering and also from sharing the technical infrastructures. In particular, the FCC would build on and reuse CERN's existing infrastructure, e.g., the existing chain of hadron accelerators, from Linac4 over PSB, PS and SPS to the LHC, can serve as an injector complex for the FCC-hh.

The technical schedule of the FCC integrated project foresees the start of FCC tunnel construction around the year 2030 — or three years after a possible project approval —, the first e^+e^- collisions at the FCC-ee during the early 2040s, and the first FCC-hh hadron collisions by 2065–70 — see **Figure 1**. In this way, the FCC integrated project would allow for a seamless continuation of High Energy Physics (HEP) after the completion of the High Luminosity LHC (HL-LHC) physics programme.

A comprehensive Conceptual Design Report (CDR) for the FCC was published in 2019 [26–28], describing the physics cases, the design of the lepton and hadron colliders, and the underpinning technologies and infrastructures. According to this design, the FCC-ee is the most sustainable of all the proposed Higgs and electroweak factory proposals, in that it implies the lowest energy

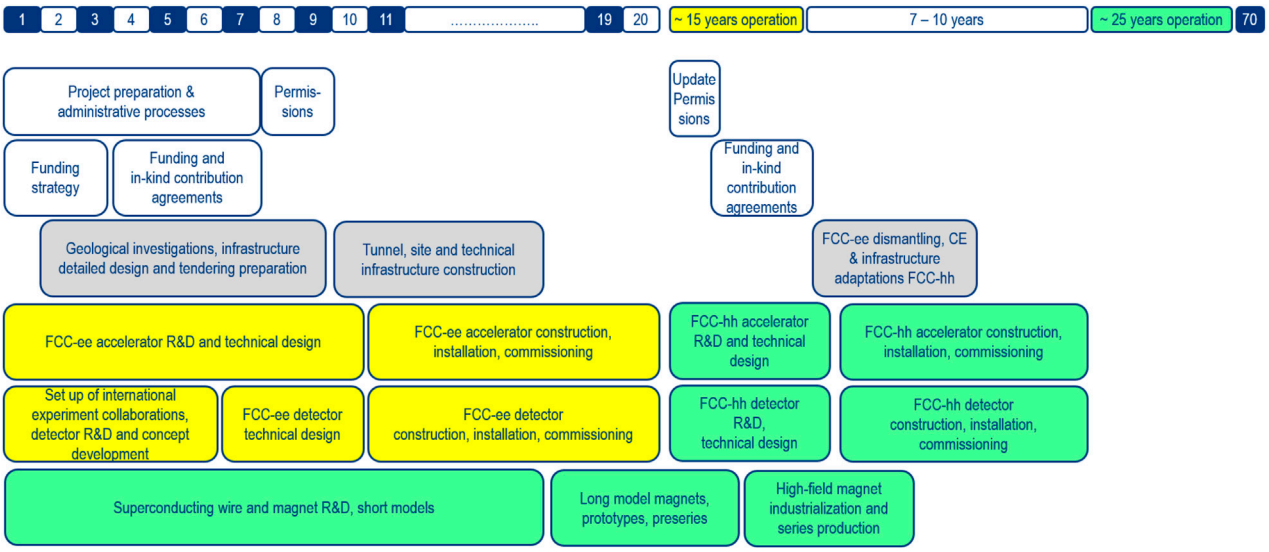


FIGURE 1 | Technical schedule of the FCC integrated project with year 1 equal to 2021 (a similar schedule was presented in Ref. [3]).

consumption for a given value of total integrated luminosity [3], over the collision energy range from 90 to 365 GeV.

1.2 FCC-ee R&D

FCC-ee research and development (R&D) focuses on further improving the overall efficiency, on obtaining the measurement precision required, and on achieving the target performance in terms of beam current and luminosity.

Key FCC-ee R&D items for improved energy efficiency include high-efficiency continuous wave (CW) radiofequency (RF) power sources (klystrons and/or solid state), high-Q SC cavities for the 400–800 MHz range, and possible applications of high-temperature superconductor (HTS) magnets. For ultra high precision centre-of-mass energy measurements, the R&D should cover state-of-art and beyond in terms of spin-polarisation simulations and measurements (inv. Compton, beamstrahlung, etc.). Finally, for high luminosity, high current operation, FCC-ee requires a next generation beam stabilization/feedback system to suppress instabilities arising over a few turns, a robust low-impedance collimation scheme, and a machine tuning system based on artificial intelligence.

1.2.1 SRF Cavity Developments

Since PETRA, TRISTAN and LEP-2, superconducting RF systems are the underpinning technology for modern circular lepton colliders. The FCC-ee baseline foresees the use of single-cell 400 MHz Nb/Cu cavities for high-current low-voltage beam operation at the Z production energy, four-cell 400 MHz Nb/Cu cavities at the W and H (ZH) energies, and a complement of five-cell bulk Nb 800 MHz cavities at 2 K for low-current high-voltage $t\bar{t}$ operation [26]. In the full-energy booster, only multi-cell 400 and 800 MHz cavities will be installed. For the FCC-ee collider, also alternative RF scenarios, with possibly fewer changes between operating points, are being explored, such as novel 600 MHz slotted waveguide elliptical (SWELL) cavities [24].

1.2.2 R&D for the FCC-ee Arcs

Aside from the various RF systems, another major component of the FCC-ee is the regular arc, covering almost 80 km. The arc cells must be cost effective, reliable and easily maintainable. Therefore, as part of the FCC R&D programme it is planned to build a complete arc half-cell mock up including girder, vacuum system with antechamber and pumps, dipole, quadrupole and sextupole magnets, beam-position monitors, cooling and alignment systems, and technical infrastructure interfaces, by the year 2025.

Constructing some of the magnets for the FCC-ee final focus or arcs with advanced high-temperature superconductor (HTS) technology could lower energy consumption and increase operational flexibility. The focus of this HTS R&D will not be on reaching extremely high field, but on operating lower-field SC magnets at temperatures between 40 and 77 K.

1.2.3 Beam Diagnostics

As experience at previous and present colliders has taught us, adequate beam diagnostics is essential for reaching or exceeding design performance. For this reason, the FCC-ee R&D programme foresees the prototyping of key beam diagnostics, like bunch-by-bunch longitudinal charge-density monitors, ultra-low emittance measurements, beam-loss and beamstrahlung monitors, real time monitoring of the collision offsets, a polarimeter for each beam able to measure the 3D polarization vector as well as the beam energy, and fast luminometers.

1.2.4 Polarimetry and Centre-of-mass Energy Calibration

Highly precise centre-of-mass energy calibration at c.m. energies of 91 GeV (Z pole) and 160 GeV (WW threshold), a cornerstone of the precision physics programme of the FCC-ee, relies on using resonant depolarisation of wiggler-pre-polarised pilot bunches

[5]. The operation with polarised pilot bunches requires constant and high precision monitoring of the residual 3-D spin-polarization of the colliding bunches, which—if nonzero—would affect the physics measurements.

1.2.5 FCC-ee Pre-Injector

Concerning the FCC-ee pre-injector, the CDR design foresaw a pre-booster synchrotron. At present, this choice is under scrutiny. As an alternative, and possibly new baseline, it is proposed to extend the energy of the injection linac to 10–20 GeV, for direct injection into the full-energy booster. The S-band linac could be based on state-of-the-art technology as employed for the FERMI upgrade at the ELETTRA synchrotron radiation facility.

It is also envisaged to design, construct and then test with beam a novel positron source plus capture linac, and measure the achievable positron yield, at the PSI SwissFEL facility, with a primary electron energy that can be varied from 0.4 to 6 GeV.

1.2.6 Full Energy Booster

The injection energy for the full-energy booster is defined by the field quality of its low-field magnets. Magnet development and prototyping of booster dipole magnets, along with field measurements, should guide the choice of the injection energy.

1.2.7 Lessons from SuperKEKB and Beam Studies

The SuperKEKB collider, presently being commissioned [19], features many of the key elements of FCC-ee: double ring, large crossing angle, low vertical IP beta function β_y^* (design value ~0.3 mm), short design beam lifetime of a few minutes, top-up injection, and a positron production rate of up to several 10^{12}/s. SuperKEKB has achieved, in both rings, the world's smallest ever β_y^* of 0.8 mm, which also is the lowest value considered for FCC-ee. Profiting from a new "virtual" crab-waist collision scheme, first developed for FCC-ee [21], in December 2021 SuperKEKB reached a world record luminosity of $3.81 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. However, many issues still need to be addressed, such as a vertical emittance blow up due to an unexplained mechanism, the transverse machine impedance, and single-bunch instability threshold, sudden beam losses without any accompanying beam oscillation, insufficient quality of the injected beam, etc.

In view of the SuperKEKB experience, studies of vertical emittance tuning is another important R&D frontier for FCC-ee. This includes simulating realistic beam measurements, constructing optics tuning knobs, especially for the final focus, and developing beam-based alignment procedures for the entire ring. Software development is an important component of this activity. Effects of beam-beam collisions and monitor resolution limits need to be considered, as should be the impact of machine errors and tuning on the dynamic aperture and on the achievable polarisation levels.

Beam studies relevant to FCC-ee—for example on optics correction, vertical emittance tuning, crab-waist collisions, or beam energy calibration—can, and will, also be conducted at INFN-LNF/DAFNE, DESY/PETRA III, and KIT/KARA [14].

1.3 High-Field Magnet R&D

The primary technology of the future circular hadron collider, FCC-hh, is the high-field magnets, Both high-field dipoles and quadrupoles [27] are required, or, possibly, combined-function magnets [12].

For constructing the accelerator magnets of the present LHC, the Tevatron, RHIC, and HERA, wires based on Nb-Ti superconductor were used. However, Nb-Ti magnets are limited to maximum fields of about 8 T, as being operated for the LHC. The HL-LHC will, for the first time in a collider, deploy some tens of dipole and quadrupole magnets with a peak field of 11–12 T, based on a new high-field magnet technology using a Nb_3Sn superconductor. The Nb_3Sn superconductor holds the promise to approximately double the magnetic field, from ~8 T at the LHC, to 16 T for the FCC-hh.

Recently, several important milestones were accomplished in the development of high-field Nb_3Sn magnets. At CERN, a block coil magnet, FRESCA2, with a 100 mm bore, achieved a world-record field of 14.6 T at 1.9 K [29]. In the US, a Nb_3Sn cosine-theta accelerator dipole short-model demonstrator with 60 mm aperture [33], reached a similar field, of 14.5 T at 1.9 K [34].

Forces acting on the magnet coils greatly increase with the strength of the magnetic field, while, at the same time, most higher-field conductors, such as the brittle Nb_3Sn, tend to be more sensitive to pressure. Therefore, force management becomes a key element in the design of future high-field magnets.

Beside the development of optimized magnet design concepts, such as "canted cosine-theta" dipoles [7], higher field can be facilitated by a higher-quality conductor. A Nb_3Sn wire development programme was set up for the FCC [2]. For Nb–Ta–Zr alloys, it could be demonstrated that an internal oxidation of Zr leads to the refinement of Nb_3Sn grains and, thereby, to an increase of the critical current density [6]. The phase evolution of Nb_3Sn wire during heat treatment is equally under study, as part of the FCC conductor development programme in collaboration with TVEL, JASTEC, and KEK [13]. Advanced Nb_3Sn wires with Artificial Pinning Centers (APCs) produced by two different teams reached the target critical current density for FCC, of 1500 A/mm^2 at 16 T [1, 30], which is 50% higher than for the HL-LHC wires. The APCs allow for better performance; they decrease magnetization heat during field ramps, improve the magnet field quality at injection, and reduce the probability of flux jumps [31].

In addition to Nb_3Sn wire, also high-temperature superconductors (HTS) are of interest, since they might allow for higher fields, operation at higher temperature, and, ultimately, perhaps even lower cost. In this context, the FCC conductor programme has been exploring the potential of REBCO coated conductors (CCs). In particular, the critical surfaces for the current density, $J_c(T, B, \theta)$, of coated conductors from six different manufacturers: American Superconductor Co. (US), Bruker HTS GmbH (Germany), Fujikura Ltd. (Japan), SuNAM Co. Ltd. (Korea), SuperOx ZAO (Russia) and SuperPower Inc. (US) have been studied [22].

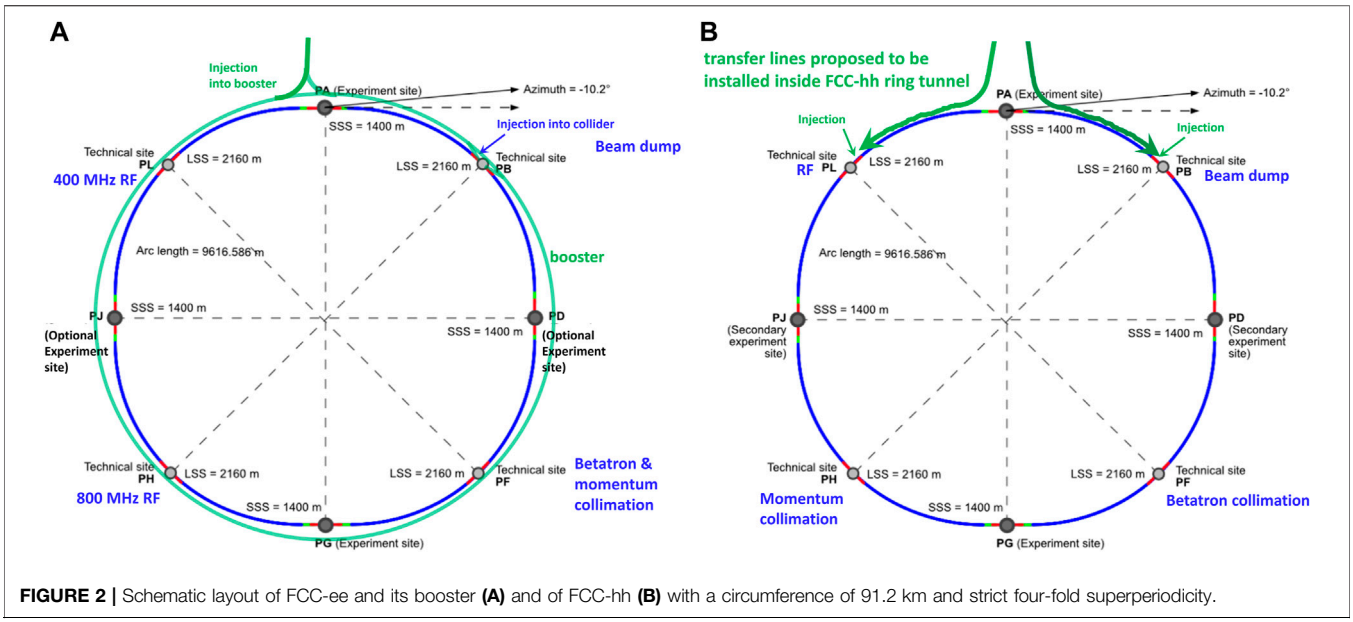


FIGURE 2 | Schematic layout of FCC-ee and its booster **(A)** and of FCC-hh **(B)** with a circumference of 91.2 km and strict four-fold superperiodicity.

Outside the accelerator field, HTS magnet technology could play an important role for fusion research. A number of companies are developing HTS magnets in view of fusion applications. One of these companies is Commonwealth Fusion Systems who in partnership with MIT's Plasma Science and Fusion center are designing SPARC, a compact net fusion energy device [18]. Their magnets are based on second generation ReBCO (Rare-earth barium copper oxide) conductors. Recently they successfully demonstrated a coil with 20 T field [17]. An interesting view on HTS prospects is presented in a Snowmass 2020 Letter of Interest [16], according to which the actual material and process costs of HTS tapes are a small fraction of their current commercial value and that there is a historical link between manufactured volume and price [15].

2 FCC FEASIBILITY STUDY

2.1 European Strategy Update 2020 and Feasibility Study Launch

The 2020 Update of the European Strategy for Particle Physics (ESPPU) [10] states that "*An electron-positron Higgs factory is the highest-priority next collider. For the longer term, the European particle physics community has the ambition to operate a proton-proton collider at the highest achievable energy.*" and "*Europe, together with its international partners, should investigate the technical and financial feasibility of a future hadron collider at CERN with a centre-of-mass energy of at least 100 TeV and with an electron-positron Higgs and electroweak factory as a possible first stage. Such a feasibility study of the colliders and related infrastructure should be established as a global endeavour and be completed on the timescale of the next Strategy update.*" Responding to this key request from the ESPPU, in the summer of 2021, the five-year Future Circular Collider Feasibility Study was launched [8, 9].

2.2 Collider Design Optimisation

2.2.1 Placement and Revised Layout

The 2019 FCC CDR describes the baseline FCC design with a circumference of 97.75 km, 12 surface sites, and two primary collision points. In 2021, a further design optimisation has resulted in an optimised placement of much lower risk, with a circumference of 91.2 km and only 8 surface sites, and which would be compatible with either 2 or 4 collision points. Consequently, adaptations of the CDR design and re-optimisation of the machine parameters are underway, taking into account not only the new placement, but also, for FCC-ee, the possibly larger number of interaction points, and the mitigation of complex "combined" effects, e.g. the interplay of transverse and longitudinal impedance with the beam-beam interaction. **Figure 2** sketches the layouts and possible straight-section functions for both electron-positron and hadron collider.

2.2.2 Parameter Update

Preliminary FCC-ee parameters for the cases of either two or four IPs are shown in **Table 1**, updated parameters for FCC-hh in **Table 2**. In the FCC-ee CDR [26], the operation at the Z and W assumed a 60°/60° phase advance per arc cell. The mitigation of the combined impedance and beam-beam effects requires a larger momentum compaction factor than in the CDR [23]. This has resulted in a "long" 90° cell, of twice the cell length used for the H and $t\bar{t}$ operation [20]. The beam parameters, in particular the emittances, bunch length, lifetime, and luminosity still need to be validated in strong-strong beam-beam simulations and in weak-strong simulations including errors and optics corrections. The

TABLE 1 | Preliminary key parameters of FCC-ee [20], as evolved from the CDR parameters, now with a shorter circumference of 91.2 km, and a new arc optics for Z and W running. Luminosity values are given per interaction point (IP), for scenarios with either 2 (left) or 4 IPs (right). Both the natural bunch lengths due to synchrotron radiation (SR) and their collision values including beamstrahlung (BS) are shown. The FCC-ee considers a combination of 400 MHz radiofrequency systems (at the first three energies, up to 2 × 2 GV) and 800 MHz (additional cavities for $t\bar{t}$ operation), with respective voltage strengths as indicated. The beam lifetime shown represents the combined effect of the luminosity-related radiative Bhabha scattering and beamstrahlung, the latter relevant only for ZH and $t\bar{t}$ running (beam energies of 120 and 182.5 GeV).

Running mode	Z	W	ZH	$t\bar{t}$	Z	W	ZH	$t\bar{t}$
Number of IPs	2				4			
Beam energy (GeV)	45.6	80	120	182.5	45.6	80	120	182.5
Bunches/beam	11600	1120	380	44	8800	1120	336	42
Beam current [mA]	1400	135	26.7	5.0	1400	135	26.7	5.0
Luminosity/IP [10^{34} cm^{-2} s^{-1}]	181	19.5	7.5	1.31	181	17.3	7.2	1.25
Energy loss/turn [GeV]	0.039	0.37	1.87	10.0	0.039	0.37	1.87	10.0
Synchr. Rad. Power [MW]	100				100			
RF Voltage 400/800 MHz [GV]	0.12/0	1.0/0	2.48/0	4./7.67	0.12/0	1.0/0	2.48/0	4./7.67
Rms bunch length (SR) [mm]	4.35	3.55	2.50	1.67	4.32	3.55	2.50	1.67
Rms bunch length (+BS) [mm]	11.4	7.02	3.75	2.16	15.2	7.02	4.45	2.54
Rms hor. emittance $\varepsilon_{x,y}$ [nm]	0.71	2.17	0.64	1.49	0.71	2.17	0.64	1.49
Rms vert. emittance $\varepsilon_{x,y}$ [pm]	1.42	4.34	1.29	2.98	1.42	4.34	1.29	2.98
Longit. damping time [turns]	1170	216	64.5	18.5	1170	216	64.5	18.5
Vertical IP beta β_y' [mm]	0.8	1.0	1.0	1.6	0.8	1.0	1.0	1.6
Beam lifetime (BS + lum.) [min.]	38	36	10	12	19	20	7	10

TABLE 2 | Key parameters of FCC-hh compared with the HL-LHC and LHC.

	FCC-hh	HL-LHC	LHC
Centre-of-mass energy (TeV)	100	14	14
Dipole field [T]	16–17	8.33	8.33
Circumference [km]	91.2	26.7	26.7
Beam current [A]	0.5	1.1	0.58
Bunch Intensity [10^{11}]	1 1	2.2	1.15
Bunch spacing [ns]	25 25	25	25
Synchr. radiation power [kW]	5400	15	7
SR power/length [W/m/aperture]	32.1	0.33	0.17
Longit. emit. damping time [h]	0.45	12.9	12.9
IP beta function $\beta_{x,y}^*$ [m]	1.1 0.3	0.15 (min.)	0.55
Normalized rms emittance [μm]	2.2	2.5	3.75
Peak luminosity [10^{34} cm^{-2}s^{-1}]	5 30	5 (lev.)	1
Events/bunch crossing	170 1000	132	27
Stored energy beam [GJ]	7.8	0.7	0.36

luminosity values per IP are slightly higher for two IPs than for four. Therefore, the beam lifetime due to radiative Bhabha scattering (inversely proportional to the total luminosity) is about a factor two higher, which allows a more aggressive choice for the beamstrahlung-induced lifetime in $t\bar{t}$ operation.

2.2.3 Monochromatisation

In addition to the 4 baseline running modes on the Z pole, at the WW threshold, at the (Z) H production peak, and above the $t\bar{t}$ threshold, listed in **Table 1**, another optional operation mode, presently under investigation for FCC-ee, is the direct *s*-channel Higgs production, $e^+e^- \rightarrow H$, at a centre-of-mass energy of 125 GeV. Here, a monochromatization scheme should reduce the effective collision energy spread in order for the latter to become comparable to the width of the Higgs [11].

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

MB is the FCC Study Leader, FZ his deputy. They have jointly managed the FCC conceptional design study, since 2014, and are also coordinating the new Feasibility Study. Together, they have defined the structure and contents of this article, and wrote the text.

FUNDING

This work was supported, in part, by the European Union's Horizon 2020 Research and Innovation programme under grant no. 951754 (FCCIS).

ACKNOWLEDGMENTS

We extend our warm thanks to the global FCC collaboration for the results reported in this article. In particular, we wish to thank A. Ballarino, A. Blondel, M. Boscolo, G. Dissertori, M. Giovannozzi, J. Gutleber, K. Hanke, P. Janot, R. Kersevan, M. Koratzinos, R. Losito, V. Mertens, M. Migliorati, K. Oide, T. Raubenheimer, and D. Shatilov. Part of the contents of this manuscript has been presented, along with other similar proposals, at the European Physical Society Conference on High Energy Physics (EPS-HEP2021) [32].

REFERENCES

1. Balachandran S, Tarantini C, Lee PJ, Kametani F, Su Y-F, Walker B, et al. Beneficial Influence of Hf and Zr Additions to Nb4at.%Ta on the Vortex Pinning of Nb_3Sn with and without an O Source. *Superconductor Sci Technology* (2019) 32:044006.
2. Ballarino A, Hopkins SC, Bordini B, Richter D, Tommasini D, Bottura L, et al. The CERN FCC Conductor Development Program: A Worldwide Effort for the Future Generation of High-Field Magnets. *IEEE Trans Appl Supercond* (2019) 29:1–9. doi:10.1109/TASC.2019.2896469
3. Benedikt M, Blondel A, Janot P, Mangano M, Zimmermann F. Future Circular Colliders Succeeding the LHC. *Nat Phys* (2020) 16:402–7. doi:10.1038/s41567-020-0856-2
4. Benedikt M, Zimmermann F. The Physics and Technology of the Future Circular Collider. *Nat Rev Phys* (2019) 1:238–40. 3 p. doi:10.1038/s42254-019-0048-0
5. Blondel A, Janot P, Wenninger J, Abmann R, Aumon S, Azzuri P, et al. *Polarization and Centre-of-Mass Energy Calibration at FCC-ee* (2019). p. 107. *arXiv 1909.12245*. doi:10.48550/arXiv.1909.12245
6. Buta F, Bonura M, Matera D, Bovone G, Ballarino A, Hopkins SC, et al. Very High Upper Critical fields and Enhanced Critical Current Densities in Nb_3Sn Superconductors Based on Nb-Ta-Zr Alloys and Internal Oxidation. *J Phys Mater* (2021) 4:025003. doi:10.1088/2515-7639/abe662
7. Caspi S, Borgnolutti F, Brouwer L, Cheng D, Dietderich D, Felice H, et al. Canted–Cosine–Theta Magnet (CCT)—a Concept for High Field Accelerator Magnets. *IEEE Trans Appl superconductivity* (2013) 24:1–4.
8. CERN Council. *Main Deliverables and Timeline of the FCC Feasibility Study* (2021). Restricted CERN Council - Two-Hundred-and-Third Session. CERN/SPC/1161
9. CERN Council. *Organisational Structure of the FCC Feasibility Study* (2021). Restricted CERN Council - Two-Hundred-and-Third Session. CERN/SPC/1155/Rev.2
10. European Strategy Group. *2020 Update of the European Strategy for Particle Physics (Brochure)*. CERN-ESU-015 (2020).
11. Faus-Golfe A, Valdivia Garcia M, Zimmermann F. The Challenge of Monochromatization, Direct s-Channel Higgs Production: $e^+e^- \rightarrow H$. *Eur Phys J Plus* (2022) 137. doi:10.1140/epjp/s13360-021-02151-y
12. Giovannozzi M, Todesco E. Combined-function Optics for Circular High-Energy Hadron Colliders. *Eur Phys J Plus* (2022) 137. doi:10.1140/epjp/s13360-022-02583-0
13. Hopkins SC, Baskys A, Ballarino A, Lachmann J, Leineweber A. Phase Evolution during Heat Treatment of Nb_3Sn Wires under Development for the FCC Study. *IEEE Trans Appl Supercond* (2021) 31:1–6. doi:10.1109/TASC.2021.3063675
14. Keintzel J, Benedikt M, Tomas R. Experimental Beam Tests for FCC-ee. *PoS, Proc EPS-HEP2021* (2021) 6:877. doi:10.22323/1.398.0877
15. Koratzinos M. *Private Communication* (2021).
16. Matias V, Moeckly B. *Price of 2G HTS Wire: Plenty of Room to Go Down*. 2020 Snowmass Letter Of Interest. USA; CA (2020).
17. MIT. *MIT News* (2021). Available at: https://news.mit.edu/2021/MIT-CFS-major-advance-toward-fusion-energy-0908. (Accessed November 15, 2021).
18. MIT. *SPARC* (2021). Available at: https://www.psfc.mit.edu/sparc. (Accessed November 15, 2021).
19. Ohnishi Y. Status and Perspectives of the SuperKEKB Project. *PoS, Proc EPS-HEP* (2021) 2021:471–8. doi:10.1393/ncc/i2012-11156-2
20. Oide K. *Private Communication* (2021).
21. Oide K, Aiba M, Aumon S, Benedikt M, Blondel A, Bogomyagkov A, et al. Design of beam optics for the future circular collider e+e– collider rings. *Phys Rev Accel Beams* (2016) 19:111005. doi:10.1103/PhysRevAccelBeams.19.111005
22. Senatore C, Barth C, Bonura M, Kulich M, Mondonico G. Field and Temperature Scaling of the Critical Current Density in Commercial REBCO Coated Conductors. *Supercond Sci Technol* (2015) 29:014002. doi:10.1088/0953-2048/29/1/014002
23. Shatilov D. *Private Communication* (2021).
24. Syratchev I, Peauger F, Karpov I, Brunner O. A Superconducting Slotted Waveguide Elliptical Cavity for FCC-ee. *Zenodo* (2021) 031953. doi:10.5281/zenodo.5031953
25. The CEPC Study Group. *CEPC Conceptual Design Report: Volume 1 - Accelerator* (2018). *arXiv* 1809.00285.
26. The FCC Collaboration. FCC-ee: The Lepton Collider. *Eur Phys J Spec Top* (2019) 228.
27. The FCC Collaboration. FCC-hh: The Hadron Collider. *Eur Phys J Spec Top* (2019) 228.
28. The FCC Collaboration. FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1. *Eur Phys J C* (2019) 79.
29. Willering G, Arnestad H, Bajko M, Bajas H, Bortot L, Bottura L, et al. Tests of the FRESCA2 100 Mm Bore Nb3Sn Block-Coil Magnet to a Record Field of 14.6 T. *IEEE Trans Appl Supercond* (2019) 29:1–6. doi:10.1109/TASC.2019.2907280
30. Xu X, Peng X, Rochester J, Sumption M, Tomsic M. *Achievement of FCC Specification in Critical Current Density for Nb_3Sn Superconductors with Artificial Pinning Centers* (2019). arXiv preprint arXiv:1903.08121. doi:10.48550/arXiv.1903.08121
31. Xu X, Sumption M, Peng X, Collings EW. Refinement of Nb3Sn Grain Size by the Generation of ZrO_2 Precipitates in Nb_3Sn Wires. *Appl Phys Lett* (2014) 104:082602. doi:10.1063/1.4866865
32. Zimmermann F. Towards a Circular Higgs and Electroweak Factory. In: Proceedings of Science The European Physical Society Conference on High Energy Physics (EPS-HEP2021); 26–30 July 2021 (2022). Online conference, jointly organized by Universität Hamburg and the research center DESY; Copyright owned by the author(s) under the Creative Commons License CC BY-NC-ND 4.0.
33. Zlobin A, Barzi EZ, Carmichael JR, Chlachidze G, Dimarco J, Kashikhim VV, et al. First Measurements of Magnet Quench Performance and Field Quality of the 15 T Nb_3Sn Dipole Demonstrator MDPCT1. In: 2019 North American Particle Accelerator Conf.(NA-PAC'19); September 2-6, 2019; Lansing, MI, USA. Geneva, Switzerland: JACoW (2019).
34. Zlobin A, Novitski I, Barzi E, Baldini M, Carmichael J, Caspi S, et al. Reassembly and Test of High-Field Nb_3Sn Dipole Demonstrator MDPCT1. *IEEE Trans Appl Supercond* (2021) 31:1. doi:10.1109/TASC.2021.3057571

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Benedikt and Zimmermann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

REVIEW
published: 14 June 2022
doi: 10.3389/fphy.2022.913510

Detector Simulation Challenges for Future Accelerator Experiments

John Apostolakis[1], Marilena Bandieramonte[2], Sunanda Banerjee[3], Nazar Bartosik[4], Gloria Corti[1], Gabriele Cosmo[1], V. Daniel Elvira[5], Thomas Evans[6], Andrei Gheata[1], Simone Pagan Griso[7], Vladimir Ivantchenko[1], Christopher Jones[5], Markus Klute[8], Charles Leggett[7], Ben Morgan[9]*, Tadej Novak[10], Kevin Pedro[5] and Harald Paganetti[11]*

[1]European Organization for Nuclear Research (CERN), Organisation Européenne pour la Recherche Nucléaire (CERN), Geneva, Switzerland, [2]Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA, United States, [3]Department of Physics, University of Wisconsin, Madison, WI, United States, [4]Sezione di Torino, Istituto Nazionale di Fisica Nucleare (INFN), Torino, Italy, [5]Fermi National Accelerator Laboratory (FNAL), Batavia, IL, United States, [6]Oak Ridge National Laboratory, Oak Ridge, TN, United States, [7]Physics Division, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA, United States, [8]Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, [9]Department of Physics, The University of Warwick, Coventry, United Kingdom, [10]Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany, [11]Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, United States

OPEN ACCESS

Edited by:
Patrizia Azzi,
National Institute of Nuclear Physics of Padova, Italy

Reviewed by:
Daniel Jeans,
High Energy Accelerator Research Organization, Japan
Valentin Volkl,
European Organization for Nuclear Research (CERN), Switzerland
Sylvie Braibant,
University of Bologna, Italy

***Correspondence:**
V. Daniel Elvira
daniel@fnal.gov
Ben Morgan
ben.morgan@warwick.ac.uk

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *05 April 2022*
Accepted: *16 May 2022*
Published: *14 June 2022*

Citation:
Apostolakis J, Bandieramonte M, Banerjee S, Bartosik N, Corti G, Cosmo G, Elvira VD, Evans T, Gheata A, Pagan Griso S, Ivantchenko V, Jones C, Klute M, Leggett C, Morgan B, Novak T, Pedro K and Paganetti H (2022) Detector Simulation Challenges for Future Accelerator Experiments. Front. Phys. 10:913510. doi: 10.3389/fphy.2022.913510

Detector simulation is a key component for studies on prospective future high-energy colliders, the design, optimization, testing and operation of particle physics experiments, and the analysis of the data collected to perform physics measurements. This review starts from the current state of the art technology applied to detector simulation in high-energy physics and elaborates on the evolution of software tools developed to address the challenges posed by future accelerator programs beyond the HL-LHC era, into the 2030–2050 period. New accelerator, detector, and computing technologies set the stage for an exercise in how detector simulation will serve the needs of the high-energy physics programs of the mid 21st century, and its potential impact on other research domains.

Keywords: high energy physics, particle physics, radiation, simulation, Monte Carlo, software, computing, high performance computing

1 INTRODUCTION

Simulation is an essential tool to design, build, and commission the sophisticated accelerator facilities and particle detectors utilized in experimental high energy physics (HEP). In this context, simulation refers to a software workflow consisting of a chain of modules that starts with particle generation, for example, final state particles from a proton-proton collision. A second module simulates the passage of these particles through the detector geometry and electromagnetic fields, as well as the physics interactions with its materials. The output contains information about times, positions, and energy deposits of the particles when they traverse the readout-sensitive components of the detector. In most modern experiments, this module is based on the Geant4 software toolkit [1–3] but other packages such as FLUKA [4, 5] and MARS [6] are also widely used, depending on the application. A third module generates the electronic signals from the readout components in response to the simulated interactions, outputting this data in the same format as the real detector system. As such, the datasets generated through simulation may be input to the same algorithms used to reconstruct physics observables from real data. Simulation is thus not only vital in designing HEP experiments, it also plays a fundamental role in the interpretation, validation, and analysis of the large and complex datasets collected by experiments to produce physics results, and its impact here should not be underestimated [7].

With many unanswered questions remaining in particle physics and the end of the Large Hadron Collider (LHC) program expected in the late 2030's, plans and ideas for the next big facilities of the 2030s–2050s are gaining momentum. As these facilities intend to explore ever higher energy scales and luminosities, the scale of simulated data samples needed to design the detectors and their software, and analyze the physics results will correspondingly grow. Simulation codes will thus face challenges in scaling both their throughput and accuracy to meet these sample size requirements with finite but ever evolving computational facilities [8]. The LHC era has already seen a significant evolution of simulation methods from "full" detailed history-based algorithms to a hybrid of full and "fast" parameterized or machine-learning based algorithms for the most computationally expensive parts of detectors [9]. A hybrid simulation strategy, using a combination of full and fast techniques will play a major role for future collider experiments, but full simulation will still be required to develop and validate the fast algorithms, as well as to support searches and analyses of rare processes. The goal of this article is to discuss how detector simulation codes may evolve to meet these challenges in the context of the second and third elements of the above simulation chain, that is the modeling of the detector, excluding the generation of initial particles. An overview of the computational challenges here may be found in [8].

Section 2 presents the design parameters of future accelerators and detectors relevant to their simulation such as colliding particle types, beam parameters, and backgrounds. Challenges in the description and implementation of complex detector geometries and particle navigation through rapidly varying magnetic fields and detector elements of different shapes and materials are discussed in **Section 3**, while the physics models needed to describe the passage of particles through the detector material at the energy ranges associated with the colliders under consideration will be discussed in **Section 4**. Beam backgrounds from particle decay or multiple hard collisions are another important topic of discussion, particularly in the case of beams with particles that decay or emit synchrotron radiation, and will be discussed in **Section 5**. **Section 6** focuses on readout modeling in the context of the opportunities and challenges posed by new detector technology, including novel materials and new generation electronics. **Section 7** looks forward to the computing landscape anticipated in the era of future colliders, and how these technologies could help improve the physics and computing performance of detector simulation software, and even shape their future evolution. **Section 8** will discuss the evolution of simulation software toolkits, including how they might adjust to new computing platforms, experiment software frameworks, programming languages, and the potential success of speculative ideas, as well as the features that would be needed to satisfy the requirements of future collider physics programs. For decades, HEP has collaborated with other communities, such as medical and nuclear physics, and space science, on detector simulation codes, resulting in valuable sharing of research and resources. **Section 9** will present examples of application of detector simulation tools originating in HEP, in particular to the medical field, and how the challenges for future HEP simulation may overlap.

This article is one of the first reviews on the role and potential evolution of detector simulation in far future HEP collider physics programs. We hope it contributes to highlight its strategic importance both for HEP and other fields, as well as the need to preserve and grow its priceless community of developers and experts.

2 FUTURE ACCELERATORS AND DETECTORS IN NUMBERS

There are several designs for future particle accelerators, each with its strengths and challenges. This chapter focuses on the accelerator and detector design parameters and issues relevant for software modeling. In particular, we survey a number of the most mature proposals, including the high luminosity LHC (HL-LHC), the high energy LHC (HE-LHC), the Large Hadron-electron Collider (LHeC) and its high luminosity upgrade (HL-LHeC), the Future Circular Collider (FCC) program of ee (electron-positron), hh (hadron-hadron), and eh (electron-hadron) colliders, the Circular Electron Positron Collider (CEPC), the Muon Collider, the International Linear Collider (ILC), the Compact Linear Collider (CLIC), and the Cool Copper Collider (CCC). **Table 1** summarizes the parameters of these proposed future accelerators, including design values for maximal energy, peak luminosity, and integrated luminosity, and references for each proposal. There are other potential future colliders that are still being designed, including the Super Proton-Proton Collider (SPPC) [10], an electron-muon collider [11], a muon-proton collider [12], and a muon-ion collider [13].

Modern particle physics accelerators operate with bunched beams and reach peak luminosities higher than $1–2 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, exceeding the initial LHC design specification. The luminosity for future hadron colliders, such as the High Luminosity LHC (HL-LHC), is limited by the maximum number of simultaneous proton-proton collisions, or pileup, under which the detectors can operate effectively. For circular lepton colliders at higher energies, the luminosity is limited by the beamstrahlung (deflection-induced synchrotron radiation), and "top-up" or "top-off" schemes to inject additional particles during beam circulation are expected to be necessary to extend beam lifetime [14]. For linear machines, design parameters like the beam size, beam power, beam currents, and repetition rates drive the peak luminosity.

Proton-proton collisions offer the greatest energy reach, but they are limited by construction costs and the availability of high-field magnets. The largest proposed center-of-mass collision energy comes from the FCC-hh at 100 TeV. Lepton colliders can also push the energy Frontier to multiple TeV. The muon collider requires R&D in order to reduce the transverse and longitudinal beam emittance via cooling and to accelerate to collision energies all within the muon's 2.2 μs lifetime [15]. However, it offers an exciting path to collision energies up to a few tens of TeV by suppressing synchrotron radiation relative to electrons. The beam-induced background (BIB) created by beam muons decaying in flight places new and unique demands on simulation [16]. Wakefield acceleration also offers a possibility

TABLE 1 | The parameters of various future accelerators. * Muon colliders face beam-induced backgrounds, which have different properties from pileup at ee or pp colliders.

Collider	Particles	$\sqrt{s}$	Peak lumi. [10^{34} cm^{-2} s^{-1}]	Peak Pileup	Int. lumi. [ab^{-1}]
HL-LHC [99, 100]	pp	14 TeV	7.5	200	3–4
HE-LHC [100, 101]	pp	27 TeV	16	500	15
LHeC [102, 103]	ep	1.3 TeV	0.5–2.4	0.1	1
HE-LHeC [102, 103]	ep	1.77 TeV	1.5	0.1	2
FCC-ee [104–106]	ee	88–365 GeV	1.5–230	0	1.5–150
FCC-hh [104, 105, 107]	pp	100 TeV	30	1,000	20
FCC-eh [104, 105, 107]	ep	3.5 TeV	1.5	1	2
CEPC [108, 109]	ee	90–240 GeV	32–3	0	2.6–16
Muon Collider [110]	$\mu\mu$	3–14 TeV	1.8–40	*	1–20
ILC [111, 112]	ee	250–500 GeV	2.7–3.6	0	1–4
CLIC [113, 114]	ee	0.38–3 TeV	1.5–6	0	1–5
CCC [115]	ee	250–550 GeV	1.3–2.4	0	2–4

TABLE 2 | Summary of technologies and applications for future projects.

Technology	Tracker	Calorimeter	Muon Detector	Pid
Solid state	Planar, 3D MAPS, LGAD, CMOS	Si, GaAs		LGAD
Gas	TPC, DC	RPC, MPGD	RPC, MPGD DT, MWPC	TPC, DC MRPC
Scintillator	SciFi, SiPM	Tiles, fibers, crystals	Panels	
Noble liquid		LAr		
Cherenkov		Quartz fibers		RICH, TOF TOP, DIRC

for reaching high energies more compactly in the further future [17].

The proposed detector technologies for the next generations of experiments at colliders are growing in breadth, as indicated by the summary in **Table 2**. These increases in technological variety are driven by both physics goals and experimental conditions. In addition, new detectors will be increasingly complex and granular. The interplay between instrumentation and computing is therefore increasingly important, as detectors become more challenging to simulate. One example is the upcoming High Granularity Calorimeter (HGCAL) at the Compact Muon Solenoid (CMS) experiment [18]. With roughly six million channels, it will be the most granular calorimeter built to date. This massively increases the geometry complexity, leading to a ~40–60% increase in the time to simulate the detector [19]; in addition, the increased precision of the detector is expected to require correspondingly more precise physical models, which may further double the simulation time in existing software [20]. The incorporation of precision timing information may also place more demands on the accuracy of the simulation.

The HL-LHC is the nearest-future collider surveyed here, and most further-future colliders aim at higher precision measurements or present even more difficult environments. Therefore, detector complexity should be expected to continue to increase, in order to facilitate the physics programs and measurements for these new colliders. More than ever before, increasingly energetic and potentially heavier particles will interact with the detector materials, and massive increases in accumulated luminosity will enable physicists to explore the tails of relevant kinematic distributions very precisely. New technologies will pose their own challenges, such as the muon collider BIB, or new materials whose electromagnetic and nuclear interactions may not be fully characterized. This motivates the continued development of detector simulation software, to ensure its computational performance and physical accuracy keep up with the bold next steps of experimental high energy physics.

3 GEOMETRY DESCRIPTION AND NAVIGATION

Geometric modeling is a core component of particle transport simulation. It describes both the material properties of detector components, which condition the particle interactions, and their geometric boundary limits. Particles are transported through these geometries in small spatial steps, requiring fast and accurate computation of distances and finding the geometry location after crossing volume boundaries. This task uses a significant fraction of total simulation time even for the current LHC experiments [8], making performance a general concern for the evolution of geometry modeling tools. As discussed in **Section 2**, future detectors will have higher granularity and, in same cases, will need to handle higher interaction rates than at the LHC, requiring the geometry modeling and navigation software to increase the throughput

of the above calculations given this increased complexity. Providing the navigation precision necessary to achieve the required physics accuracy will likely be challenged by the presence of very thin detectors placed far away from the interaction point.

The detector geometry description of a HEP experiment goes through several processing steps between the initial computer-aided designs (CAD) [21] to the in-memory representation used by the simulation. These transformations primarily reduce the complexity and level of detail available in the CAD model to increase computing performance without compromising the required physics accuracy. To start with, the detector design study phase is particularly important for future collider experiments. Essential detector parameters concerning solid angle acceptances, material composition or engineering constraints need to be optimized in a tight cycle involving full simulation. DD4hep [22] has emerged as a commonly used detector description front-end for future accelerator experiments, providing an internal model representation independent on the geometry modeler back-end, and interfaces for importing the geometry representation from many sources. This not only facilitates the design optimization cycle, but also the handling of multiple geometry versions and the integration of important detector conditions such as alignment, which affect the geometry during the experiment operation.

Although the geometry models at the core of today's HEP detector simulation were designed in the 60's, Geant geometry implementations [1, 23] have enjoyed continuous success over many generations of CPU architectures because of a number of features that reduce both the memory footprint and algorithmic complexity. Multiple volume placements, replication using regular patterns, and hierarchies of non-overlapping 'container' volumes enable efficient simulation of very complex setups comprising tens of millions of components. However, creating the model description for such setups is often a long and arduous process, and the resulting geometry is very difficult to update and optimize.

The most popular 3D models used in simulation nowadays are based on primitive solid representations such as boxes, tubes, or trapezoids, supporting arbitrarily complex Boolean combinations using these building blocks. Different simulation packages use different constructive solid geometry (CSG) flavours [24], providing a number of features and model constraints to enhance the descriptive power and computation efficiency. However, performance can be highly degraded by overuse of some of these features, such as creating unbalanced hierarchies of volumes or creating overly complex Boolean solids. Using such inefficient constructs in high occupancy detector regions near the interaction point generally leads to significant performance degradation.

The current geometry implementations have a very limited adaptive capability for optimizing such inefficient constructs, mainly due to the high complexity of the model building blocks. The geometry queries can only be decomposed to the granularity of solid primitives, so user-defined constructs cannot be internally simplified. This calls for investigating surface models as alternatives to today's geometry representations. Adopting boundary representation (BREP) models [25] composed of first and second-order algebraic surfaces, would allow decomposing navigation tasks into simple surface queries. An appropriate choice of the BREP model flavor allowing surface queries to be independent could greatly favor the highly-parallel workflows of the future.

Developing automatic conversion tools from CAD surface-based models to the Geant4 simulation geometry proved to be too challenging in the past. DD4hep provides a conversion path for complex surfaces into tessellated bodies usable directly in simulation [26]. A conversion procedure relying exclusively on tessellation would however introduce important memory and performance overheads during simulation. Supporting surface representations directly in the simulation geometry would make such conversions more efficient. This would also provide a simpler transition from the engineering designs to the simulation geometry, having fewer intermediate representations. It would also make it easier to implement transparent build-time optimizations for inefficient user constructs.

Successive upgrades to adapt to new computing paradigms such as object-oriented or parallel design have not touched the main modelling concepts described above, which served their purpose for decades of CPU evolution but are quickly becoming a limiting factor for computing hardware with acceleration. Recent R&D studies [27, 28] have shown that today's state-of-the-art Geant-derived geometry codes such as VecGeom [29] represent a bottleneck for vectorized or massively parallel workflows. Deep polymorphic code stacks, low branch predictability, and incoherent memory access are some of the most important reasons for performance degradation when instruction execution coherence is a hardware constraint. This is intrinsic to the model being used, combining in the same query algorithms of very different complexity, called in an unpredictable manner and unfriendly to compiler optimizations. These studies also indicate the need to simplify the geometry models being used, highly reducing or eliminating unnecessary abstractions.

Performance optimization is particularly important for common geometry navigation tasks such as collision detection of the simulated particle trajectories with the geometry setup, and relocation after crossing volume boundaries. Navigation helpers are using techniques such as *voxelization* [30] or bounding volume hierarchies (BVH) [31] to achieve logarithmic complexity in setups having several millions objects. Adopting efficient optimization strategies will be more relevant for the more complex detectors of the future.

The same problem of collision detection is addressed by graphics systems, in particular, ray-tracing (RT) engines such as NVIDIA OptiX [32] that make use of dedicated hardware acceleration. Adapting HEP detector simulations to use such engine was implemented in the Opticks library [33], and demonstrated speedups of more than two orders of magnitude compared to CPU-based Geant4 simulations of optical photon transport in large liquid-Argon detectors. This required adapting the complete optical photon simulation workflow to GPUs, but also simplifying and transforming the geometry description to match OptiX requirements. Generalizing this approach for future

HEP detector simulation would require a major re-engineering effort, in particular for the geometry description. How exactly RT technology evolves will likely have a big impact on the solutions adopted for detector geometry modeling. As the use of RT acceleration proliferates in the gaming industry, APIs supported by dedicated languages and libraries will most probably be made publicly available. Combined with larger on-chip caches, future low-latency graphics chips may allow externalizing geometry as an accelerated service for simulation. Such service could become an important booster, but would be conditional to the simplification of the geometry description and added support for batched multi-track workflows.

Evolution in computing technology will most probably present game-changing opportunities to improve simulation software, as described in **Section 7**. For example, tensor cores [34] provide a large density of Flops, although at a cost in precision. Geometry step calculations cannot make use of half-precision floating point (*FP16*) directly because rounding errors would become too important and affect both physics precision and transportation over large distances in the detector. Some optimizations may however be delegated to a *FP16*-based navigation system using ML inference to, for instance, prioritize candidate searches. Single-precision *FP32*-based geometry distance computation should be given more weight in the context of the evolution of reduced-precision accelerator-based hardware, because the option to reduce precision fulfils physics requirements in most cases. Furthermore, it would provide a significant performance boost due to a smaller number of memory operations for such architectures. Recent studies report performance gains as large as 40% for certain GPU-based simulation workflows [28], making R&D in this area a good investment, as long as memory operations remain the dominant bottleneck, even if chips evolve to provide higher Flops at *FP32* precision or better. The precision reduction option is, however, not suitable for e.g., micron-thin sensors, where the propagation rounding errors become comparable to the sensor thickness. Addressing this will require supporting different precision settings depending on the detector region.

4 PHYSICS PROCESSES AND MODELS

As mentioned in **Section 1**, Geant4 has emerged as the primary tool to model particle physics detectors. Geant4 offers a comprehensive list of physics models [35] combined with the continuous deployment of new features and improved functionality, as well as rigorous code verification and physics validation against experimental data.

4.1 Current Status

During the first two periods of data taking in 2010–2018, the LHC experiments produced, reconstructed, stored, transferred, and analyzed tens of billions of simulated events. The physics quality of these Geant4-based Monte Carlo samples produced at unprecedented speed was one of the critical elements enabling these experiments to deliver physics measurements with greater precision and faster than in previous hadron colliders [7, 36]. Future accelerator programs will, however, require the implementation of additional physics processes and continuous improvements to the accuracy of existing ones. A quick review of the current status of physics models in Geant4 will precede a discussion of future needs.

Physics processes in Geant4 are subdivided over several domains, the most relevant for HEP being particle decay, electromagnetic (EM) interactions, hadronic processes, and optical photon transport. The precision of the modeling has to be such that it does not become a limiting factor to the potential offered by detector technology. EM physics interactions of $e^-/e^+/\gamma$ with the detector material, producing EM showers in calorimeters, consume a large fraction of the computing resources at the LHC experiments. Reproducing the response, resolution, and shower shape at a level of a few per mille requires modeling particle showers down to keV levels, which contain a large number of low-energy secondary particles that need to be produced and transported through magnetic fields. This level of accuracy is required in order to distinguish EM particles from hadronic jets, and to efficiently identify overlapping showers. Highly accurate models for energy deposition in thin calorimeter layers are also essential for reconstruction of charged particles and muons. Simulation of tracking devices requires accurate modeling of multiple scattering and backscattering at low and high energy, coupled with very low energy delta electrons. Geant4 delivers on all these requirements by modeling EM processes for all particle types in the 1 keV to 100 TeV energy range. The accuracy of Geant4 EM showers is verified by the CMS [37] and ATLAS [38] experiments.

Geant4 models physics processes for leptons, long-lived hadrons, and hadronic resonances. Simulation of particle decay follows recent PDG data. The decay of *b*-, *c*-quark hadrons and τ-leptons is outsourced to external physics generators via predefined decay mechanisms.

Simulation of optical photon production and transport is also provided by Geant4. The main accuracy limitation arises from the large compute time required to model the large number of photons and the many reflections that may occur in within the detector. Various methods to speed-up optical photon transport are available, depending on tolerance to physics approximations.

Hadron-nuclear interaction physics models are needed to simulate hadronic jets in calorimeters, hadronic processes in thin layers of tracking devices, and for simulating shower leakage to the muon chambers. Geant4 hadronic physics is based on theory models and tuned on thin target data [3]. This approach guarantees a more reliable predictive power than that offered by parametric models for a wide range of materials, particle types and energy ranges for which data measurements are not available. Parameter tuning and model extensions are necessary to describe all particle interactions at all energies [2]. Geant4 has adopted the approach of combining several models that fit the data best in different energy ranges, since it is unrealistic to expect that one single model would do the job over the full kinematic range of interest. This is done by providing several sets of predefined "physics lists", which are combinations of EM and hadronic processes and models.

Experiments need to identify the most suitable for their own physics program by performing the necessary physics validation studies and possibly applying calibration corrections [37–40].

4.2 Future Needs

The large data volumes to be collected by the HL-LHC experiments will enable experiments to reduce statistical uncertainties, therefore demanding more accurate simulation to help reduce systematic uncertainties in background estimation and calibration procedures. The next generation of HEP detectors to be commissioned at the LHC and designed to operate in future lepton and hadron colliders will have finer granularity and incorporate novel materials, requiring simulation physics models with improved accuracy and precision, as well as a broader kinematic coverage. Materials and magnetic fields will also need to be described in more detail to keep systematic uncertainties small. Moreover, new technologies [41–43] will allow detectors to sample particle showers with a high time resolution of the order of tens of picoseconds, which will need to be matched in simulation. Consequently, the simulation community has launched an ambitious R&D effort to upgrade physics models to improve accuracy and speed, re-implementing them from the ground up when necessary (e.g., GeantV [27], Adept [28], Celeritas [44]). Special attention will be needed to extend accurate physics simulation to the $O(100)TeV$ domain, including new processes and models required to support the future collider programs.

Achieving an optimal balance between accuracy and software performance will be particularly challenging in the case of EM physics, given that the corresponding software module is one of the largest consumers of compute power [36]. Reviews of EM physics model assumptions, approximations and limitations, including those for hadrons and ions will be needed to support the HL-LHC and Future Collider (FCC) programs. The Geant4 description of multiple scattering [45] of charged particles provides predictions in good agreement with data collected at the LHC. Nevertheless, the higher spatial resolution in new detectors [41, 46–48], may require even higher accuracy to reproduce measured track and vertex resolutions. Excellent modeling of single-particle scattering and backscattering across Geant4 volume boundaries for low energy electrons are critical for accurate descriptions of shower shapes in calorimeters, such as CMS's high granularity hadronic calorimeter. At the very high energies present at the FCC, nuclear size effects must be taken into account, and elastic scattering models must be extended to include nuclear form factors in the highest energy range. The description of form factors may affect EM processes at high energies in such a way that it affects shower shapes and high energy muons. A theoretical description of the Landau-Pomeranchuk-Migdal (LPM) effect, significant at high energy, is included in the models describing the bremsstrahlung and pair-production processes in Geant4. For the latter, introducing LPM leads to differences in cross-sections at very high energies that will need to be understood when data become available. A relativistic pair-production model is essential for simulation accuracy at the FCC. Rare EM processes like γ conversion to muon and hadron pairs also becomes important at very high energies and will have to be added. This is also essential to properly model beam background effects in the collision region of a Higgs Factory. In the cases of the FCC and dark matter search experiments, the description of pair production will need to be extended to include the emission of a nearby orbital electron (triplet production) and to take into account nuclear recoil effects. Finally, γ radiative corrections in EM physics may affect significantly the accuracy of measurements at Higgs factories and will need to be added to the models. All these rare processes must be added to the simulation to improve the accuracy in the tails of the physics distributions, where backgrounds become important. These corrections must be included such that they are invoked only as needed, thus not increasing the computing cost of EM modeling. At the FCC collision energy, the closeness of tracking devices to the interaction points will also require widening the range of physics models of short lived particles. This will be particularly important for high-precision heavy flavor measurements, as non-negligible fractions of beauty and charm hadrons will survive long enough to intercept beam pipes and the first detector layers. Describing the interaction of such particles with matter may already be required at the HL-LHC program because of a reduction of the distance between the trackers and the interaction point [41, 46]. A review of how detector simulation interfaces to dedicated decay generators during particle transport may be necessary.

In hadronic interactions, more than one model is needed to describe QCD physics processes accurately over the whole energy range. Typically, a hadronic interaction is initiated when a high energy hadron collides with a nucleon in the nucleus of a given material. This is followed by the propagation of the secondary particles produced through the nucleus, the subsequent de-excitation of the remnant nucleus and particle evaporation, until the nucleus reaches the ground state. Different sets of models map naturally to these phases depending on the initial energy of the collision: a parton string model for energy above few GeV, an intra-nuclear cascade model below that threshold. Pre-compound and de-excitation models are used to simulate the last steps in the evolution of the interaction. A reliable description of showers in hadronic calorimeters requires accurate descriptions of all these processes.

Geant4 offers two main physics lists to describe hadronic physics in high energy collider experiments. The main difference between the two consists in the choice of the model describing the initiating quark-parton phase mentioned above, either a quark-gluon string model, or a Fritiof model [3]. Having more than one model allows to estimate the systematic uncertainties arising from the approximations they use. Unfortunately, neither of them is accurate enough to describe the hadronic interactions at multi-TeV energies occurring at the FCC. New processes will need to be implemented in the hadronic physics simulation suite to address this higher energy domain, taking inspiration from those available in the EPOS generator [49], used by the cosmic ray and heavy ion physics communities.

Another element essential for the simulation of hadronic physics is precise calculations of interaction cross-sections. At the highest energies, Geant4 uses a general approach based on the Glauber theory [50], while at lower energies cross-sections are

evaluated from tables obtained from the Particle Data Group [51]. This approach profits from the latest thin-target experiment measurements and provides cross-sections for any type of projectile particle. The precision of cross-section calculations for different types of particles will need to be improved as more particle types become relevant to particle flow reconstruction in granular calorimeters.

A correct description of particle multiplicity within hadronic showers is also needed to model the physics performance of highly granular calorimeters (e.g. CMS [18]), and is also essential to simulate high-precision tracking devices (e.g. LHCb spectrometer). The parameters describing hadronic models must be tuned to describe all available thin target test beam data, and the models expanded to provide coverage to as many beam particles and target nuclei as possible. For flavor physics, it is important to take into account the differences in hadronic cross-sections between particle and anti-particle projectiles.

5 BEAM BACKGROUNDS AND PILEUP

The main categories of beam backgrounds at *ee* colliders are machine and luminosity induced [52]. The former is due to accelerator operation and includes Synchrotron Radiation (SR) and beam gas interactions. The latter arises from the interaction of the two beams close to the interaction point of the experiment.

The SR that may affect the detector comes from the bending and focusing magnets closest to it. While detectors will be shielded, a significant fraction of photons may still scatter in the interaction region and be detected. This is expected to be one of the dominant sources of backgrounds in the FCC-ee detector [53]. Beam gas effects are a result of collisions between the beam and residual hydrogen, oxygen and carbon gasses in the beam pipe inside the interaction region.

The luminosity induced background is generated from the electromagnetic force between the two approaching bunches, which leads to the production of hard bremstrahlung photons. These may interact with the beam and an effect similar to e^+e^- pair creation can occur, or they scatter with each other which can result in hadrons, and potentially jets, in the detector. Stray electrons due to scattering between beam particles in the same bunch can also hit the detector.

The main background at *pp* colliders are the large number of inelastic proton–proton collisions that occur simultaneously with the hard-scatter process, collectively known as pileup. This usually results in a number of soft jets coinciding with the collision. The number of interactions per crossing at the future colliders is expected to exceed 1,000, compared to no more than 200 at the end of the HL-LHC era. An additional source of luminosity induced background is the cavern background. Neutrons may propagate through the experimental cavern for a few seconds before they are thermalized, thus producing a neutron-photon gas. This gas produces a constant background, consisting of low-energy electrons and protons from spallation.

Machine induced backgrounds at *pp* colliders are similar to the *ee* ones [38]. Besides the beam gas, the beam halo is a background resulting from interactions between the beam and upstream accelerator elements. In general, pile-up dominates over the beam gas and beam halo.

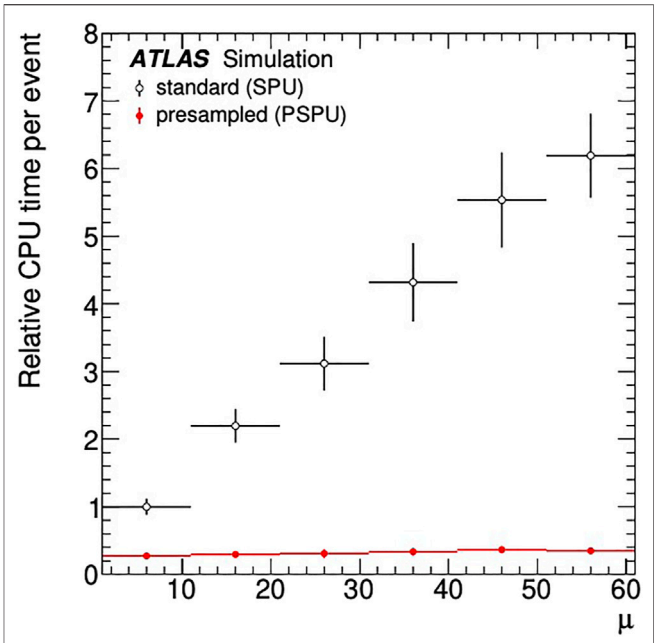


FIGURE 1 | Comparison of the average CPU time per event in the standard ATLAS pileup digitization (black open circles) and the pre-sampled pileup digitization (red filled circles) as a function of the number of *pp* collisions per bunch crossing (μ). The CPU time is normalized to the time taken for the standard pileup for the lowest μ bin. Taken from Ref. [55].

Muon colliders are special in that the accelerated particles are not stable. Decays of primary muons and the interaction of their decay products with the collider and detector components [54] constitute the main source of beam background. Compared to *ee* colliders this represents an additional source of background resulting in a large number of low momentum particles that may not be stopped by shielding and enter the interaction region of the detector. Additionally, this type of background needs to be simulated with higher precision outside of the interaction region.

An important consideration is the detector response and readout time compared to the time between collisions, which is often longer. In-time and out-of-time pile-up should be considered separately. In-time pileup arises from additional collisions that coincide with the hard-scatter one, while out-of-time pile-up comes from collisions occurring at bunch crossings different from the hard-scatter one, although affecting the readout implicitly.

5.1 Bottlenecks in Computational Performance

The biggest bottleneck in the time it takes to model pileup in a *pp* collider is the number of interactions per bunch crossing. As seen in black circles in **Figure 1**, the CPU time requirement has a very steep dependence on this parameter, which needs to match data-taking conditions. The second issue can be the slow response time of the detectors, requiring a large number of out-of-time bunch

crossings to be simulated. This can be solved by only simulating the detectors when needed, as not all have the same sensitive time range. Improvements in detector technologies that will be used in future experiments may make these times small enough not to cause a significant overhead.

Traditionally each in-time or out-of-time interaction is sampled individually and taken into account at the digitisation step, when detector digital responses are emulated. Experiments pre-sample pile-up events and reuse them between different samples to reduce computational time [55, 56]. While the pre-sampling itself still has the same CPU limitations, using those pileup events barely depends on the amount of pileup (red circles in **Figure 1**), but could cause larger stress on storage. Thresholds to analogue signals are applied at digitization to reduce the amount of saved digits significantly, at the cost of reduced precision when two digital channels are merged. Thus pre-sampling thresholds need to be tuned for each individual detector, and computing resources can only be saved by reusing pre-sampled events, where a compromise between CPU savings and increased storage needs to be made in a way that maintains optimal physics performance.

Another option to fully avoid the CPU bottleneck of pileup pre-sampling is to use pileup events from data. The main bottlenecks here are non-constant detector conditions and alignment. Re-initializing the simulated geometry adds overheads which may be mitigated by averaging conditions over long periods. However, this solution will come at the cost of reproducing data less precisely. Furthermore detector readout only provides digital information above some thresholds which are usually tuned for primary collisions and thus relatively high. This reduces precision when merging the information with the simulated hard-scatter event.

While other types of background are much lower at *pp* colliders and their simulation can usually be skipped, this is not the case for *ee* colliders. Some of those backgrounds, e.g., beam gas effects, synchrotron radiation and intra-beam scattering, happen outside the detector cavern. They are simulated by the accelerator team as they also affect beam operations. To avoid re-simulating the same type of background, the simulation can be shared with the experiment as a list of particles that enter the interaction region [57], though this is still a large number of low-momentum particles to simulate. Experiments thus also use randomly-triggered collision events for the background estimation, while also being affected by the threshold effects.

5.2 Optimal Strategy for Future Colliders

During the development stage of the future experiments, detailed simulation of all types of beam backgrounds is of utmost importance. Simulation provides estimates of the physics impact of backgrounds and helps to optimize the detector design to minimize them as much as possible [58]. Some backgrounds can be parametrized or even completely neglected. One such example is that of cavern background neutrons at hadron colliders. In most cases their contribution is orders of magnitude smaller than that of pileup, although outer muon chambers would require a detailed description, if high precision is required. As low momentum neutron simulation is very slow, it can be performed only once and used to derive parametrized detector responses, which can then be injected at the digitization stage.

As discussed earlier in this section, separate simulation of beam backgrounds and pre-digitization saves computing resources and has a negligible impact on physics performance when reused randomly between samples. With the increased background rates expected in future colliders, iterative mixing and merging of background contributions will become an essential technique. Detector readout thresholds must be set sufficiently low to allow merging of digital signals multiple times with negligible degradation of accuracy. This would allow iterative pileup pre-sampling, where multiple events with a low number of interactions could be merged to give an event with a high number of interactions. It would also allow to merge different types of backgrounds that would be prepared independently. Furthermore, a special set of lower background thresholds could be setup in the actual detector to enable the use of real data events as background sources. The latter would yield a reduced performance degradation as compared to current detectors.

Most of all the beam background simulation strategy depends on physics accuracy requirements. As mentioned in **Section 1**, current experiments are moving towards a more frequent use of fast simulation methods, either based on parametrized detector responses or on machine learning technologies. The latter could be used to choose the precision of the simulation algorithm depending on the event properties, or to fully generate the background on the fly. Regardless of the choice of the strategy used to simulate large volumes of physics samples, a detailed modeling as provided by full simulation will always be needed, if nothing else to derive and tune the faster methods.

6 ELECTRONIC SIGNAL MODELING

The ambitious physics program at future accelerator-based experiments requires detectors which can perform very accurate measurements and handle high occupancy at the same time. To achieve these goals, it is of paramount importance to collect as much information from each individual detector channel as possible, including the three spatial coordinates, time and energy.

For simplicity, this section focuses on two main classes of detectors that pose the most challenges from a computational point of view: tracking detectors and calorimeters. Those are the ones that usually use the largest number of electronic readout channels, thus their behavior needs to be simulated in detail.

New generation calorimeters are designed as tracking devices as well as providers of energy deposition information in the form of the five-dimensional measurement referred to in the first paragraph. These extended capabilities beyond traditional calorimetric observables present challenges to the simulation effort, since modeling must achieve accurate descriptions of all these observables simultaneously. Additionally, calorimeters will often operate in a high-occupancy environment in which sensor

and electronics performance degrade fast as a consequence of radiation damage.

The digitization step of simulation takes as input the Geant4-generated analogue signals from the detector. The first step of the digitization process accumulates this input and groups it for individual read out elements. This is done in a number of time slots which define the integration time for the detector. Beyond this step, modeling is highly detector dependent. It is driven by detailed descriptions of readout electronics including the noise component, cross-talk, and the readout logic which involves the shaping of the signal and the digitization of the pulse. Finally, a digit is recorded when the signal is above a predefined threshold.

6.1 Tracking Detectors

Various types of tracking detectors are currently employed in HEP experiments at colliders [51], with the most widely used being silicon, gaseous (RPC, MDT, Micromegas, etc), transition-radiation, and scintillation detectors. Of these, silicon-based detectors are among the most challenging and computationally expensive to simulate, given the large number of channels and observables involved.

Silicon detectors give rise to electron-hole pairs which are collected with a certain efficiency, amplified, digitized, and recorded. When biased by a voltage difference, the response of the sensor to the passage of ionizing particles is characterized by its charge collection efficiency (*CCE*) and its leakage current (I_{leak}). As the sensors are operated well above their full depletion voltage, the *CCE* is expected to be high. The current digitization models for silicon detectors use either parametric or bottom-up approaches. For parametric approaches, the overall simulated energy deposit is split across readout channels using a purely parametric function based either on detailed simulations or data; for bottom-up approaches the energy deposit is used to generate multiple electron-hole pairs that are then propagated through a detailed simulation of the electric field and used to compute the expected signal generated at the electrodes. Several models are employed for how the overall deposited energy is split. They range from simple models performing an equal-splitting along the expected trajectory to more complex models [59], each giving different increasing levels of accuracy at the price of being computationally more expensive.

Exposure to radiation induces displacements in the lattice and ionization damage, liberating charge carriers. These effects contribute to a reduction of the *CCE* and increase in the I_{leak}. The increase in instantaneous luminosity projected at the HL-LHC collider challenged experiments to implement simulation models able to predict the reduced *CCE* expected in the presence of radiation damage. A detailed simulation of the electric field is used with more refined models describing the probability of charge-trapping and reduced *CCE* [60–62]. Those models tend to be heavy on computing resources, prompting parametric simulation approaches to be developed as well.

Detector designs for future colliders differ substantially depending on the type of environment they will have to withstand. Detectors at moderate to high-energy e^+e^- colliders will see a clean event and moderate rates of radiation. For such detectors, a detailed simulation strategy is crucial for high precision physics measurements; however, the demand for large simulated samples makes a hybrid approach including parametrizations most likely. Silicon-based tracking detectors are also the technology of choice at muon colliders. The radiation environment within this machine poses unique challenges due to the high level of beam-induced backgrounds (BIBs). Real-time selection of what measurements are most likely to come from the interaction point rather than from BIBs is likely to rely on detailed shape analyses of the neighboring pixels that give signals as well as possible correlation across closely-spaced layers [63]. A hybrid approach will likely be needed, consisting of a detailed simulation of the detector layers where the raw signal multiplicity is the highest and needs to be reduced, together with a fast simulation approach for the rest of the tracking detector. For detectors at future hadron colliders, the extreme radiation environment near the interaction point will make it mandatory to implement radiation damage effects in the simulation. For this, a parametrized approach would also be the most realistic path to keep computational costs under control.

6.2 Calorimeters

Calorimeters may be broadly classified as of two types. In homogeneous calorimeters, the entire volume is sensitive and contribute a signal through the generation of light from scintillation or Cerenkov emission. These photons are collected, amplified, digitized and recorded. In sampling calorimeters, the material that produces the particle shower is distinct from the material that measures the deposited energy. Particles traversing sampling calorimeters lose energy through the process of ionization and atomic de-excitation. The charge of the resulting products (electrons and ions) is subsequently collected, amplified, digitized and recorded. In homogeneous calorimeters, modeling photon transport to the photo-transducers is CPU intensive and typically implemented as a parametrization tuned to predictions obtained from a specialized simulation package [33, 64]. Nowadays, simulation of optical photons is offloaded to GPUs to mitigate computing costs, taking advantage of the high levels of parallelism achievable for electromagnetically interacting particles' transport. The photon transmission coefficient is affected by radiation damage due to formation of color centers in the medium, thus an assumption is made on the distribution of color centers in the medium. The light output, $L(d)$, after receiving a radiation dose d, is described by an exponential function that depends on the dose:

$$L(d) = L_0 \exp(\mu \cdot d),$$

where the parameter μ is a property of the material and depends on the dose rate. The radiation damage parametrizations are typically calibrated from data coming out of a monitoring system. The radiation dose and the neutron fluence (flux over time) are estimated using an independent simulation of the detector setup.

The next step in the simulation chain for calorimeters is the treatment of the photo transducer, the most commonly used type

being silicon photo-multipliers. These devices also suffer time-dependent effects related to the radiation exposure: decrease of photo-statistics (fewer photons reaching the device) and increase of the noise coming from dark currents. The noise increases with the square root of the fluence, which in turn is proportional to the sensor's area. Signal simulation in silicon photo-multipliers involves: emulation of photo-statistics using a Poisson distribution, description of the distribution of the photo electrons according to pulse shape, adjustmentment of the signal arrival time, as well as the modeling of the dark current (thermal emission of photo-electrons), the cross-talk among the channels induced in the neighbors of the fired pixels, the pixel recovery time after being fired, and the saturation effect for large signals when several photo-electrons fall on the same pixel. An exponential function describes accurately the re-charge of the pixel as a function of time, while cross-talk can be modeled using a branching Poisson process. The Borel distribution [65, 66] analytically computes the probability of neighboring cells to fire.

Finally, the simulation of the readout electronics includes: the readout gain, adjusted to get an acceptable signal to noise ratio throughout the life time of the detector; the electronics noise, with contributions from the leakage current in the detector, the resistors shunting the input to the readout chip, and the implementation of the so-called common mode-subtraction; and the ADC pulse shape, which decides the fraction of charge leaked to the neighboring bunches. Zero suppression is also modeled, keeping only the digits which cross a threshold in the time bunch corresponding to sample of interest.

In future colliders, simulation of silicon-based calorimeters will face similar challenges than those described in the previous section for tracking devices. Parametrizations of time consuming photon transport may be replaced with detailed modeling and processed on computing devices with hardware accelerators. Radiation damage will be more pronounced in high-background environments such as high-energy hadron colliders and muon colliders, introducing a time-dependent component all through the signal simulation chain which will need to be measured from data and modeled in detail.

7 COMPUTING

Non-traditional, heterogeneous, architectures, such as GPUs, have recently begun to dominate the design of new High Performance Computing centers, and are also showing increasing prevalence in data centers and cloud computing resources. Transitioning HEP software to run on modern systems is proving to be a slow and challenging process, as described in **Section 7.3**. However, in the timescale of future colliders, this evolution in the computing landscape offers tremendous opportunity to HEP experiments. The predicted increase in compute power, the capability to offload different tasks to specialized hardware in hybrid systems, the option to run inference as a service in remote locations in the context of a machine learning approach, open the field of HEP simulation to a world where simulation data could grow several times in size, while preserving or improving physics models and detector descriptions.

7.1 Projection of Hardware Architecture Evolution

For example, the U.S. Department of Energy (USDOE) will be setting up three new GPU-accelerated, exascale platforms in 2023–2024 at the Oak Ridge Leadership Computing Facility (OLCF [67]), Argonne Leadership Computing Facility (ALCF [68]), and Lawrence Livermore National Laboratory. Additionally, the National Energy Research Scientific Computing Center (NERSC [69]) is deploying an NVIDIA-based GPU system for basic scientific research. **Figure 2** shows peak performance in Flops for machines deployed at the OLCF between 2012 and 2023. In addition to the projected ~ 55 × increase in computing performance from 2012 to 2022, the percent of peak provided by GPUs has increased from ~ 91% to greater than 98% over that period. This situation is reflected in computing centers around the world such as Piz Daint in Swizterland [70], Leonardo in Italy [71], and Karolina in Czechia [72] that heavily use NVIDIA GPUs, LUMI in Finland [73] that will use AMD GPUs, and MareNostrum 4 in Spain [74] that uses both NVIDIA and AMD GPUs. Japan's Fugaku [75], the current leader of the Top 500 supercomputers list [76], has a novel architecture with very wide registers that behave very much like a GPU. We see similar heterogeneous computing center designs in smaller institutional clusters and grid computing sites. Thus, in order to take advantage of the massive increases in computing capability provided at the HPC centers, optimizing existing and future simulation codes for GPUs is essential. The other HPCs at the head of the current Top500 List which do not explicitly use GPUs, such as Fugaku, have hybrid architectures that have very wide vector processors that offer much the same functionality as traditional GPUs.

The primary driver of this evolution is the power requirements of high-performance computing. **Figure 3** shows power consumption for OLCF machines from 2012 to 2022. Here, we see that for a 3× increase in total power consumption there is a 17 fold increase in Flops per MW.

It is difficult to predict the exact nature of the hardware landscape beyond 5 years or so, but undoubtedly we will see evolutionary changes of current hardware rather than revolutionary ones—a failed product can now cost billions of dollars due to design and fabrication costs. Core counts will continue to go up, as transistor feature sizes decrease, with increasing use of multi-chip and 3D stacked solutions needed to avoid overly large silicon sizes. It is also likely that vendors will devote larger sections of silicon to specialized functions, such as we see with Tensor and Ray Tracing cores in current GPUs. FPGA and ASIC vendors are now offering specialized component layouts for domain specific applications, and this level of customization will likely increase. We are also beginning to see the combination of multiple different types of cores, such as high and low power CPUs and FPGAs in the same silicon die or chiplet array, leading to more integrated heterogeneous architectures with faster communication channels between the various components and much quicker offload speeds.

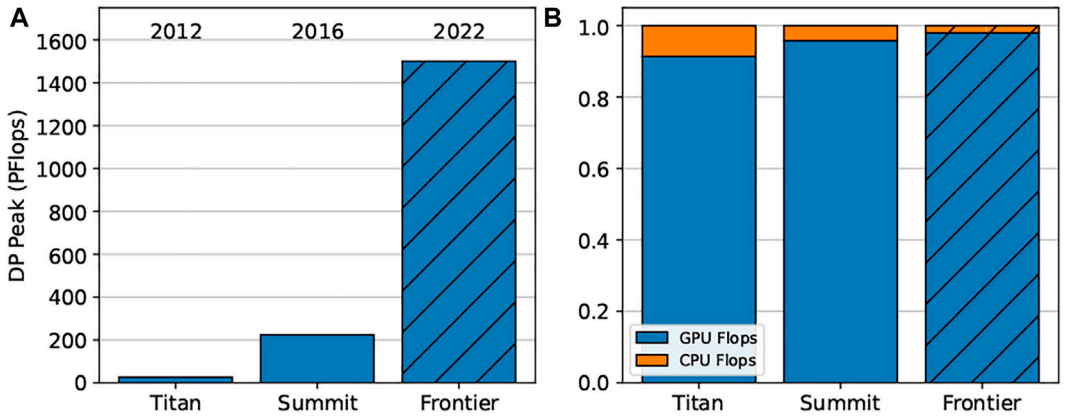


FIGURE 2 | Peak performance in Flops **(A)** and fraction of Flops provided by GPU and CPU **(B)** for GPU-accelerated systems deployed at the Oak Ridge Leadership Computing Facility (OLCF). The peak performance for Frontier is projected.

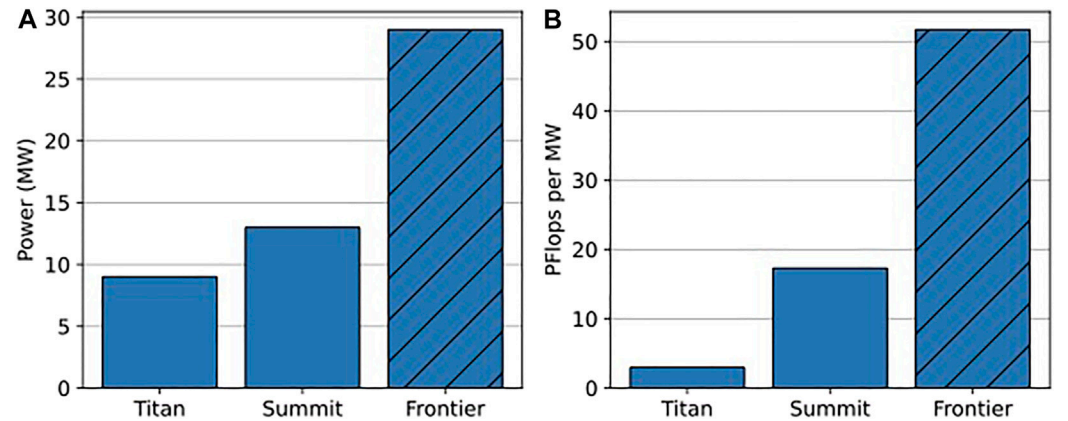


FIGURE 3 | Power consumption **(A)** and Flops per MW **(B)** for GPU-accelerated systems deployed at the Oak Ridge Leadership Computing Facility (OLCF). The power requirements for Frontier are projected.

7.2 Description of Heterogeneous Architectures

Heterogeneous architectures such as GPUs and FPGAs are fundamentally different from traditional CPU architectures. CPUs typically possess a small number of complicated cores that excel at branch prediction and instruction prefetching. They have multiple levels of large, fast caches, and typically have very low access latencies. GPUs, on the other hand, have a very large number of simple cores (hundreds of thousands for modern GPUs), that do not handle branch mis-predictions gracefully. GPU cores that are grouped in a block must operate in lockstep, all processing the same instruction. Branch mis-predictions and thread divergence will cause a stall, greatly decreasing throughput. GPUs often have much more silicon devoted to lower and mixed precision operations than they do for double precision calculations, which are heavily used in High Energy Physics. GPUs are optimized for Single Instruction Multiple Data (SIMD) style of operations, where sequential threads or cores access sequential memory locations—randomized memory access causes significant performance degradation. Finally, GPUs have very high access latencies compared to CPUs—it can take tens of microseconds to offload a kernel from a host to a GPU. The combination of massive parallelism, memory access patterns, and high latencies of GPUs require a fundamentally different programming model than that of CPUs.

The architecture of FPGAs is considerably simpler than that of more general purpose GPUs and CPUs, consisting of discrete sets of simple logic and I/O blocks linked by programmable interconnects. Programming an FPGA consists of mapping the program flow of the code onto the logic layout of the device and activating the appropriate interconnects. The concept of directly encoding operations into hardware has gained traction over the last 5 years, and current compute GPUs have significant operations encoded directly into the hardware including mixed-precision matrix-matrix multiplication (tensor cores) and ray-tracing for AI and graphics applications, respectively.

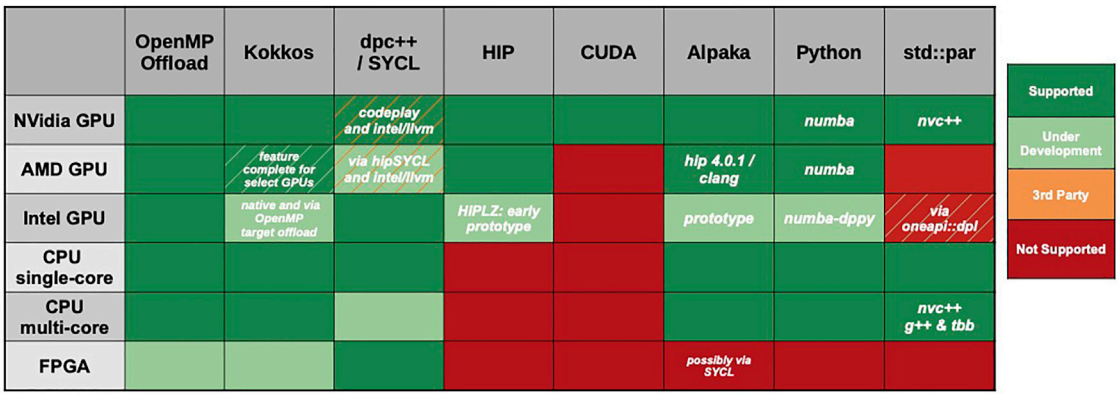


FIGURE 4 | Portability solutions for heterogeneous architectures.

FPGAs potentially offer greater promise in this regard because they can be encoded for domain-specific operations, whereas tensor cores have limited utility outside of the deep-learning AI space. This strength is also a weakness when it come to deploying integrated FPGA hardware in large compute centers as developing code on FPGAs is considerably more challenging than on GPUs because the programming languages are not as flexible and the compilation times are several orders of magnitudes slower, making the programming cycle much more difficult. Thus, no major FPGA-based large systems are currently in development, and we suspect that FPGA usage will remain restricted to local deployments for the near-to-medium time frame.

7.3 Challenges for Software Developers

All of the GPU manufacturers support programming only with their own software stack. NVIDIA uses CUDA, AMD promotes HIP, and Intel employs oneAPI. Other heterogeneous architectures such as FPGAs also use unique programming languages such as Verilog and HLS. The vast majority of current HEP software is written in C++, and supported by physicists who are usually not professional developers. Typical HEP workflows encompass millions of lines of code, with hundreds to thousands of kernels, none of which dominate the computation. In order to target the current diverse range of GPUs and FPGAs, we would have to rewrite a very large fraction of the HEP software stack in multiple languages. Given the limited available workforce, and the extremely challenging nature of validating code that executes differently on multiple architectures, experiments would have to make very difficult choices as to which hardware they could target, ignoring large amounts of available computing power. Fortunately, we have seen a number of portability solutions start to emerge recently, such as Kokkos, Raja, Alpaka, and SYCL, which are able to target more than one hardware backend (see **Figure 4**). Furthermore, hardware vendors have seen the benefits of cross platform compatibility, and are working to develop standards which they are trying to incorporate into the C++ standard. Ideally, a single language or API that could target both CPUs and all available heterogeneous architectures would be the preferred solution.

Currently, mapping computational physics and data codes to GPU architectures requires significant effort and profiling. Most HEP code bases are not easily vectorizable or parallelizable, and many HEP applications are characterized by random memory access patterns. They tend to follow sequential paradigms, with many conditional branch points, which make them challenging to adapt to GPUs. Even tasks such as particle transport, which in high luminosity environments such as the HL-LHC seemingly offer very high levels of parallelism, are in fact very difficult to run efficiently on GPUs due to rapid thread divergence cause by non-homogeneous geometrical and magnetic field constraints.

One avenue that offers some hope for easier adoption of GPUs is the use of Machine Learning (ML) techniques to solve physics problems. We are seeing increasing acceptance of ML algorithms for pattern recognition and feature discrimination tasks in HEP, as well as for more novel tasks such as generative models for energy depositions in calorimeter simulations. ML backends for all GPU and other heterogeneous architectures already exist, and are often supported directly by the hardware manufacturers, which greatly eases the burden for HEP developers.

8 SOFTWARE TOOLKITS

The evolution of simulation software toolkits will depend greatly on the hardware, whose evolution on the timescale of 10 years is uncertain as discussed in **Section 7**. Today's leading toolkit, the Geant4 toolkit [3] used by most large experiments' detector simulation, and also the particle transport tools FLUKA [5] and MARS15 [6] used in the assessment of radiation effects, are large, complex, and have evolved over 30 years of CPU-centric computation.

8.1 Computing Hardware Accelerator Usage

Whether current simulation toolkits can be adapted to profit adequately from a variety of computing hardware accelerators,

principally GPUs, or whether new accelerator-centric codes can be created and then interfaced into existing toolkits is a key research question. The profitability of the conversion also involves the effort required for the development of the production level code, and the cost to create GPU-capable applications. The latter is under active exploration.

The research into GPU usage is inspired by efforts in related particle transport applications in HEP and other fields. As discussed in Secs. 3 and 6, the Opticks project [33] offloads simulation of optical photons to NVidia GPUs and demonstrates methods to deal with complex specialised geometries on these devices, specifically ones that have many repetitive structures. MPEXS, a CUDA-based application for medical physics [77] using Geant4-derived physics models, also demonstrated efficient use of GPU resources for regular 'voxelised' geometries. However, the general problem of modeling a large range of energies for particles combined with the full complexities of modern detector geometries has not been tackled yet. Solving these general problems is the domain of two ongoing R&D efforts, the Celeritas project [44] and the AdePT prototype [28]. Both are starting by creating CUDA-based proof-of-concept implementations of electromagnetic physics, and particularly showering, in complex detector geometries on GPUs. Key goals of the projects include identifying and solving major performance bottlenecks, and providing a first template for efficiently extracting energy deposits, track passage data, and similar user-defined data. Initially, both are targeting the simulation of electron, photon, and positron showers in complex geometrical structures currently described by deep hierarchies with many repetitions of volumes at different levels. They have identified the need for a geometry modeller adapted for GPUs and accelerators, and sufficiently capable to handle these complex structures (see **Section 3**). They are in the process of defining and developing solutions for such a geometry modeller.

The limitations of the bandwidth and latency for communication between the CPU and accelerator are important constraints in the utilization of GPUs and other accelerators for particle transport simulation, and for the overall application. Minimising the amount of data exchanged, such as input particles and output hits, between the CPU and accelerator, is an important design constraint for GPU-based particle transport. The types of detectors for which it is suitable may depend on this. The contention for this resource may also constrain the overall application which integrates the particle transport and showering with event generation, generation of signal, and further reconstruction.

Existing prototypes such as AdePT and Celeritas strongly focus on keeping computation inside the accelerator, and moving back to the CPU only the absolute minimum of data and work. When only a selected region of a geometry is accelerated, a particle which escapes that region must be returned - as must particle tracks which undergo (rare) interactions not currently simulated in GPU code, e.g., photo-nuclear interactions. Of course the largest and critical data transferred out of the accelerator are the experiment hit records (or processed signal sum values) and other user information such as truth information.

Early phase exploration of the potential of FPGAs for particle transport is being conducted for medical physics simulation [78]. Yet the challenges involved appear more daunting, due to the need to compile a complex tool into hardware. It seems likely that this approach would be investigated only after implementations are built using 'simpler' building blocks on GPUs. Potentially these will profit from leveraging implementations created for portable programming frameworks.

Based on current trends, except situations where ultimate performance is required for time critical applications, we expect the established vendor-specific libraries (CUDA, Hip, DPC++) to be slowly supplanted by the emerging portable programming paradigms (Kokkos, Alpaka, SYCL), and within a few years a convergence to be established on standard-supported languages and libraries such as C++'s standard library std:par execution policy. With the importance of portability between hardware of different vendors, it is critical to identify and invest in cross-vendor solutions, and potentially paradigms that can be used to investigate alternative hardware platforms, as mentioned above for FPGAs.

8.2 Opportunities for Parallelism

We expect applications and future toolkits will need to expose multiple levels of parallelism in order to manage resources and to coordinate with other computations, such as reconstruction and event generation. Such levels could entail parallel processing of different events as well as parallel processing of multiple algorithms or even multiple particles within an event. A detector simulation toolkit cannot assume that it controls all resources, but must cooperate with other ongoing tasks in the experiment application. At this point, it is unclear how to accomplish this cooperation efficiently.

Seeking to obtain massive parallelism of thousands or tens of thousands of active particles is challenging to develop in detector simulation. The GeantV project [27] explored the potential of SIMD-CPU based parallelism by marshalling similar work ('event-based' in the parlance of neutron simulation), e.g., waiting till many particles entered a particular volume before propagating the particles through that volume. The project's conclusion was that the speedup potential was modest - between 1.2 and 2.0.

It seems clear that the ability to execute many concurrent, independent kernels on recent GPUs is of crucial interest to HEP, as it avoids the need for very fine grained parallelism at the thread level, which was the goal of the GeantV project. Given the difficulty of taking advantage of the full available parallelism of modern GPUs by a single kernel, being able to execute many kernels doing different tasks will be invaluable.

8.3 Parametrized Simulation

In parallel with the need for a full, detailed simulation capability to meet the physics requirements of the future colliders, the focus is growing on developing techniques that replace the most CPU-intensive components of the simulation with faster methods (so called "fast simulation" techniques), while maintaining an adequate physics accuracy. This category includes optimization/biasing techniques that aim at tuning parameters

concerning simulation constituents such as geometry or physics models and which are strictly experiment specific, as well as the possibility of parametrizing part of the simulation (i.e. electromagnetic shower development in calorimeters), by combining different machine learning techniques. R&D efforts are ongoing in all the major LHC experiments to apply cutting-edge techniques in generative modelling with deep learning approaches, e.g., GANs, VAEs and normalizing flows, targeting the description of electromagnetic showers.

We expect the bulk production of Monte Carlo simulation data to be performed with a combination of detailed and parametrized simulation techniques. To this end, enabling the possibility to combine fast and full simulation tools in a flexible way is of crucial importance. Along these lines, we expect Geant4 to evolve coherently by providing tools allowing integration of ML techniques with an efficient and smooth interleaving of different types of simulation.

8.4 Future of Geant4

Due to its versatility, the large number of physics modeling options, and the investment of many experiments including the LHC experiments, we expect an evolved Geant4 to be a key component of detector simulation for both the ongoing and the near future experiments well into the 2030s. Over the next decade, we expect Geant4's capabilities to evolve to include options for parameterized simulation using machine learning, and acceleration for specific configurations (geometry, particles and interactions) on selected hardware, both of which should significantly increase simulation throughput. These enhanced capabilities will however come with significant constraints, due to the effort required to adapt user code to the accelerator/heterogeneous computing paradigm. Furthermore, there is a need to demonstrate that substantial speedup or throughput improvements can be obtained before such an investment in adaptation of user applications can be undertaken. Full utilization of accelerators may not be required as offloading some work to accelerators should free up CPU cores to do additional work at the same time thereby improving throughput. In addition, some HPC sites may require applications to make some use of GPUs in order to run at the site. Therefore, some minimum GPU utilization by simulation may make it possible for experiments to run on such HPC resources thereby reducing the total time it takes to do large scale simulation workflows.

9 APPLICATIONS OF HEP TOOLS TO MEDICAL PHYSICS AND OTHER FIELDS

After the initial developments of Monte Carlo (MC) methods for the Manhattan project, the tools became available to the wider research community after declassification in the 1950's. One of the early adapters of MC methods were physicists in radiation therapy. Researchers were eager to predict the dose in patients more accurately as well as designing and simulating detectors for quality assurance and radiation protection. The simulations were done mainly using in-house developed codes, with some low energy codes modeling photons up to 20 MeV developed or transferred from basic physics applications [35, 79]. Use of MC tools from the HEP domain mainly started with heavy charged particle therapy, first using protons and Helium ions and later employing heavier ions such as Carbon ions. Early research here was also done with in-house codes mostly studying scattering in inhomogeneous media. In the early 1990's more and more high-energy physicists entered the field of medical physics and brought their expertise and codes with them. Thus started the use of general-purpose MC codes in radiation therapy that were initially developed and designed for high energy physics applications, such as Geant4 and Fluka. Fruitful collaborations were also established with the space physics field, with HEP-developed toolkits applied to particle detector design as well as the similar areas of dosimetry and radiation damage [80].

9.1 Beam Line Design and Shielding Calculations

Beam line design and shielding calculations are done prior to installing a treatment device. These applications of MC are no different to the HEP use case except for the beam energies studied. Beam line transport would be done by the machine manufacturers and is often based on specialized codes such as, for instance, Beam Delivery Simulation (BDSIM) [35]. On the other hand, shielding calculations aim at a conservative estimate with limited required accuracy and would use mostly analytical methods.

Shielding calculations are also critical in both manned and unmanned space missions to determine the radiation environment for humans [81] and instrumentation, as well as detector backgrounds [82].

9.2 Detector Design Studies

Nuclear and HEP physics hardware developments are frequently finding applications in radiation therapy and space missions due to similar requirements concerning sensors and real-time data processing. Detectors are less complex compared to HEP but the components used in simulations are very similar. Differences are in the particles of interest as well as the energy region of interest. As in HEP, MC simulations are a powerful tool to optimize detectors and treatment devices [83, 84]. In fact, for radiation therapy or diagnostic imaging, MC are not only being employed by researchers but also by vendors to optimize their equipment.

9.3 Dose Calculation

Predicting the dose in patients is arguably the most important task in radiation therapy and has therefore been the most active MC topic [85]. It has similar importance in space physics for predicting dose rates for astronauts and in materials/electronics [81, 86].

Despite its accuracy, MC dose calculation has not found widespread use in treatment planning in medicine. However, vendors of commercial planning systems have now developed very fast Monte Carlo codes for treatment planning where millions of histories in thick targets need to be simulated in minutes or seconds in a very complex geometry, i.e. the patient as imaged with CT [87]. Therefore, these specialized codes have

replaced multi-purpose MC codes that are often less efficient. Multi-purpose codes are however being used as a gold standard for measurements that are not feasible in humans. In addition, they are often used to commission treatment planning and delivery workflows. As we are dealing with biological samples such as patients, scoring functionality often goes beyond about what is typically used in HEP such as scoring phase spaces on irregular shaped surfaces or dealing with time-dependent geometries.

9.4 Diagnostic Medical Imaging

MC has long been used in the design of imaging systems such as positron emission tomography (PET) or computed tomography (CT) [88]. Like in therapy, HEP codes are being applied either directly or tailored to imaging applications, i.e. for low energy applications [89]. Time of flight as well as optical simulations are done using MC. In recent years MC is more and more used to also understand interactions in patients. As radiation therapy is pursuing image-guided therapy, imaging devices are also incorporated in treatment machines resulting in problems that are being studied using MC such as the interaction between magnetic resonance imaging (MRI) and radiation therapy, either conventional (photon based) or magnetically scanned proton treatments.

9.5 Simulation Requirements for Non-HEP Applications

9.5.1 Physics Models and Data for Energy Ranges of Interest

Medical and many space applications typically fall not under high-energy but low-energy physics. HEP tools might therefore not simulate some effects accurately or their standard settings are not applicable for low energies and have to be adjusted and potentially even separately validated [90]. Measurements of fragmentation cross-sections and attenuation curves are needed for MC applications in clinical environments. Most cross-sections and codes are indeed not very accurate for applications outside HEP because materials and energy regions of interest are very different. In fact, cross-sections needed for medical physics applications go mostly back to experiments done in the 1970's and are no longer of interest to the basic physics community. For instance, considerable uncertainties in nuclear interaction cross-sections in biological targets are particularly apparent in the simulation of isotope productions [91]. Furthermore, the interest of high-energy physics is mainly in thin targets whereas medical physics needs accurate representations of thick target physics to determine energy loss in patients or devices including Coulomb scattering and nuclear halo.

9.5.2 Computational Efficiency (Variance Reduction)

In the future we may see two types of MC tools in medical physics, i.e., high-efficiency MC algorithms focusing solely on dose calculation for treatment planning and multi-purpose codes from high energy physics for research and development. The latter can and will be used more and more to replace difficult or cumbersome experiments such as detector design studies for dosimetry and imaging. Nevertheless, thick target simulations are often time consuming and variance reduction techniques have been developed in medical physics [92] that may also be applicable for high-energy physics applications, as discussed in **Section 8**, with cross-fertilization of the two fields.

9.6 Future Role of MC Tools Outside of HEP

The main application of high-energy physics tools to other domains will continue to be in detector design, quality assurance and dose calculation. Furthermore, not only researchers in medical and space physics but also manufacturers of therapy and detector equipment are employing MC methods to develop new devices. Whilst these fields may not in general have the extreme requirements on performance and throughput as the future experiments discussed in **Section 2**, the improvements necessary here for HEP will benefit other user communities. By delivering higher accuracy physics with a smaller computational resource for a given sample size, a commensurate reduction in the costs to research time, money, and environmental impact will be possible.

It is important that collaborations between the many communities utilizing simulation codes are maintained to ensure sharing of requirements and methodologies to mutual benefit. Medical physics increasingly overlaps with radiation biology, where research promises a higher clinical impact than pure physics studies, a paradigm shift that became apparent in the last decade. Monte Carlo codes will thus be applied also in the field of radiation biology and radiation biochemistry [93]. Multiple efforts have already started, most notably the extensions of Geant4 (Geant4-DNA) and TOPAS (TOPAS-nBio) [94, 95]. These extensions require codes to evolve particularly when it comes to physics in small nanometer volumes and computational efficiency when using very small step sizes, which may have commonalities with the geometry developments discussed in **Section 3**. **Figure 5** shows an example of the geometries of typical size and complexity of molecular structures that are targeted by these simulations. The toolkit/API design of codes such as Geant4 have been critical in allowing such extensions, as well as allowing development of a wide range of applications for generic use cases [89, 96–98]. It is vitally important that HEP MC codes continue to use this software architecture to allow such innovation and extension. With simulation geometries, energy regions, materials, particle tracking and scoring that may be very different from HEP applications, continued exchange of ideas from other user communities will be invaluable in maintaining and developing HEP simulation codes.

10 SUMMARY AND CONCLUSION

Detector simulation codes such as Geant4, FLUKA, and MARS have played a central part in the development and operation of the current generation of HEP experiments and in the analysis and interpretation of their physics results. This critical role will continue as physicists design and plan the next generation of

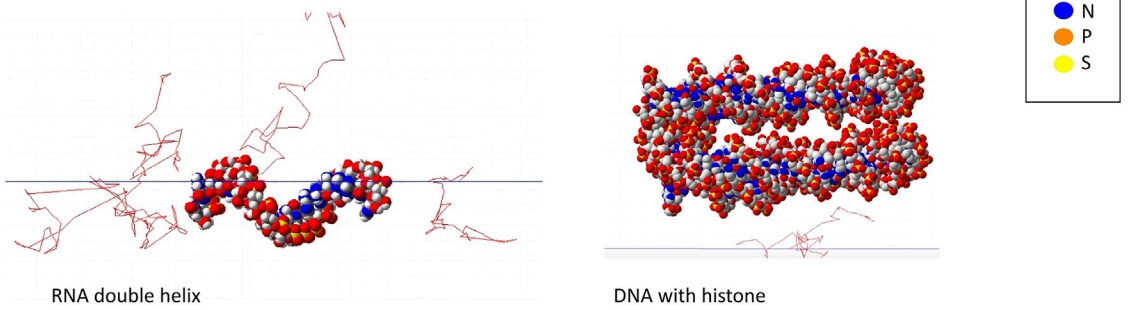


FIGURE 5 | Molecules from the protein data bank read into TOPAS-nBio with a proton track (blue) and secondary electrons (red). Two nucleic acids are shown; an RNA strand and a nucleosome.

collider facilities to operate during the mid-21st century. These experiments, like their predecessors, will push the boundaries of accelerator and detector technology to explore and improve our knowledge of fundamental physics. While simulation codes have already been significantly upgraded through the LHC era to take full advantage of technologies including multi-core CPUs and machine learning, further evolution will be needed for this software to run on future computing architectures and deliver the large and accurate data samples demanded by future collider programs.

The primary challenges for detector simulation posed by future accelerators and detector designs are driven by the increased beam luminosities and energies combined with the high granularity (in space and time) of the proposed detectors. Higher luminosity naturally means that simulations will need to deliver larger sample sizes to reduce statistical uncertainties in, for example, background estimations, driving an overall need to increase performance and hence throughput. Corresponding increases in the accuracy and precision of models for electromagnetic and hadronic physics processes will thus be required to reduce systematic uncertainties, and to extend their domain of validity to cover higher beam energies and novel materials. Beam backgrounds will also increase in line with luminosity, and are a especially important area to model during the design phase of experiments to optimize physics and instrumental backgrounds therefore improving the precision of physics measurements and extending the reach of new particle searches. Higher granularity detector systems will challenge current codes for describing their geometries with the increased number of volumes, as well as propagating particles over large distances while retaining precision of their intersections with small or thin detector elements. R&D programs are already underway to explore directions for evolving this critical area of simulation. They are exploring techniques and hardware used in the computer graphics industry for ray-tracing and Computer-Aided Design (CAD), a particularly promising direction of research. Both high luminosity and detector granularity impact the final simulation step of digitization. The increased number of detector readout channels generates a higher computational load, especially for bottom-up models of signal creation, while the more intense radiation environment will require time-dependent effects measured from data to be modelled.

None of these components of the overall simulation toolkit exist in isolation. For example, the accuracy of energy depositions in a fine grained tracking calorimeter will be dependent on the interplay between the physics models and navigation of particles through the geometry elements under the influence of a magnetic field. Balancing physics accuracy against computing performance will be an important aspect for experimentalists and simulation code developers to consider. It is clear that employing a hybrid of full and fast parametrized or ML-based techniques is a realistic strategy for simulating detectors. Fast simulation may well find application in a broader range of cases than at the present time, either as a full generative step, or to optimize inputs to, or choice of, full Monte Carlo algorithms. Complete, high throughput, "full" simulation workflows will nonetheless be required to develop, validate, and tune "fast" methods, as well as to retrain or optimize them in response to changes in experiment conditions or physics program.

While the debate here is driven by the requirements of future HEP collider programs, simulation software evolves in the context of changes in a broader landscape of developments in hardware and software for High Performance Computing in academia and industry. The ever rapid pace of technology development limits predictions of how this may impact HEP over the next 5 years, let alone the 2040–2050 timescale for experiments in future collider facilities, but even the current evolutionary trends in GPU, FPGA and other new architectures offer many exciting opportunities for greater computational power at lower monetary and environmental cost. Equally, a significant challenge for HEP simulation will be in evolving existing interfaces and algorithms to effectively utilize this diverse range of emerging architectures. Software portability tools to assist targeting multiple hardware backends are developing rapidly, and experience in their use is building within the HEP community. HEP-originated simulation codes have permeated to other fields requiring modeling of radiation

transport, especially in medical, bio-, and space physics. The collaborations established through this wide range of use cases have lead to many mutually beneficial developments and impact in both research and industry, and this can be expected, and should be encouraged, to continue. Though there are differences in energy ranges and detector complexity, increased physics accuracy and computational efficiency and throughput will be to the benefit of all. Furthermore, new or novel commonalities may be found, for example in modeling and navigating complex geometries whether that be a future collider detector or a DNA molecule.

Predicting the future for any technological or scientific endeavour can only offer a blurred snapshot of reality, but it is clear at least that the HEP community will continue to require accurate and computationally efficient detector simulation codes to develop and utilize its next generation of facilities. Developing software that meets these requirements presents a major, yet exciting, challenge that will foster collaboration across fundamental physics, high performance computing and computer science, medical, bio- and space physics, both in academia and industry. It is this depth and breadth of expertise across domains that will support and drive innovation in HEP simulation, making this human resource the most important to nurture and grow to enable the realization of HEP physics programs at future colliders during the second half of the 21st century.

AUTHOR CONTRIBUTIONS

VDE and BM are the main editors and the corresponding authors, contributing throughout the paper and more specifically to the introduction and summary sections. MK and KP contributed throughout the paper and more specifically to the future accelerators and detectors in numbers section. GaC and AG contributed throughout the paper and more specifically to the geometry description and navigation section. GlC and VI contributed throughout the paper and more specifically to the physics processes and models section. NB and TN contributed throughout the paper and more specifically to the beam backgrounds and pileup section. SB and SG contributed throughout the paper and more specifically to the electronic signal modeling section. TE and CL contributed throughout the paper and more specifically to the computing section. JA, MB, and CJ contributed throughout the paper and more specifically to the software toolkits section. HP contributed throughout the paper and more specifically to the Applications of HEP tools to other fields section.

FUNDING

MB is supported by the University of Pittsburgh under PHY1624739 with the US National Science Foundation. VE, CJ, and KP are supported by Fermilab, operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy. TE is supported by the Exascale Computing Project (ECP), project number 17-SC-20-SC. Work for this paper was supported by Oak Ridge National Laboratory (ORNL), which is managed and operated by UT-Battelle, LLC, for the U.S. Department of Energy (DOE) under Contract No. DEAC05-00OR22725. CL is supported by the DOE HEP Center for Computational Excellence at Lawrence Berkeley National Laboratory under B&R KA2401045 with the U.S. Department of Energy. SPG is supported by the U.S. Department of Energy, Office of Science under contract DE-AC02-05CH11231. BM is supported by STFC in the United Kingdom. HP is in part funded by the U.S. National Institute of Health: U24 CA215123 (the TOPAS Tool for Particle Simulation, a Monte Carlo Simulation Tool for Physics, Biology and Clinical Research).

ACKNOWLEDGMENTS

We gratefully acknowledge the international high-energy and nuclear physics simulation community, as well as our colleagues from many other research fields who actively participate in the development, validation, and support of detector simulation software tools. Their effort is essential for particle physicists around the world to develop procedures and produce measurements of outstanding quality.

REFERENCES

1. Agostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, Arce P, et al. Geant4—a Simulation Toolkit. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2003) 506:250–303. doi:10.1016/S0168-9002(03)01368-8
2. Allison J, Amako K, Apostolakis J, Araujo H, Arce Dubois P, Asai M, et al. Geant4 Developments and Applications. *IEEE Trans Nucl Sci* (2006) 53:270–8. doi:10.1109/TNS.2006.869826
3. Allison J, Amako K, Apostolakis J, Arce P, Asai M, Aso T, et al. Recent Developments in Geant4. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2016) 835:186–225. doi:10.1016/j.nima.2016.06.125
4. Ferrari A, Sala PR, Fassò A, Ranft J. FLUKA: A Multi-Particle Transport Code (Program Version 2005). In: *CERN Yellow Reports: Monographs*. Geneva: CERN (2005). doi:10.5170/CERN-2005-010
5. Böhlen TT, Cerutti F, Chin MPW, Fassò A, Ferrari A, Ortega PG, et al. The FLUKA Code: Developments and Challenges for High Energy and Medical Applications. *Nucl Data Sheets* (2014) 120:211–4. doi:10.1016/j.nds.2014.07.049
6. [Dataset] Mokhov N. *MARS15, Version 00* (2016). Available from: https://www.osti.gov/biblio/1282121 (Accessed April 5, 2022).
7. Daniel Elvira V. Impact of Detector Simulation in Particle Physics Collider Experiments. *Phys Rep* (2017) 695:1–54. doi:10.1016/j.physrep.2017.06.002
8. The HEP Software Foundation. A Roadmap for HEP Software and Computing R&D for the 2020s. *Comput Softw Big Sci* (2019) 3:7. doi:10.1007/s41781-018-0018-8
9. Aad G, Abbott B, Abbott DC, Abed Abud A, Abeling K, Abhayasinghe DK, et al. AtlFast3: the Next Generation of Fast Simulation in ATLAS. *Comput Softw Big Sci* (2022) 6:7. doi:10.1007/s41781-021-00079-7
10. Tang J, Scott Berg J, Chai W, Chen F, Chen N, Chou W, et al. Concept for a Future Super Proton-Proton Collider. *arXiv* (2015).

11. Lu M, Levin AM, Li C, Agapitos A, Li Q, Meng F, et al. The Physics Case for an Electron-Muon Collider. *Adv High Energ Phys* (2021) 2021:1–6. doi:10.1155/2021/6693618
12. Kaya U, Ketenoglu B, Sultansoy S, Zimmermann F. Luminosity and Physics Considerations on HL-LHC and HE-LHC Based Mu-P Colliders. *arXiv* (2019).
13. Acosta D, Li W. A Muon-Ion Collider at BNL: The Future QCD Frontier and Path to a New Energy Frontier of μ+μ− Colliders. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2022) 1027:166334. doi:10.1016/j.nima.2022.166334
14. Talman R. Scaling Behavior of Circular Colliders Dominated by Synchrotron Radiation. *Int J Mod Phys A* (2015) 30:1544003. doi:10.1142/S0217751X15440030
15. Delahaye JP, Diemoz M, Long K, Mansoulié B, Pastrone N, Rivkin L, et al. Muon Colliders. *arXiv* (2019).
16. Bartosik N, Andreetto P, Buonincontri L, Casarsa M, Gianelle A, Griso SP, et al. Full Detector Simulation with Unprecedented Background Occupancy at a Muon Collider. *Comput Softw Big Sci* (2021) 5:21. doi:10.1007/s41781-021-00067-x
17. D'Arcy R, Chappell J, Beinortaite J, Diederichs S, Boyle G, Foster B, et al. Recovery Time of a Plasma-wakefield Accelerator. *Nature* (2022) 603:58–62. doi:10.1038/s41586-021-04348-8
18. CMS Collaboration. The Phase-2 Upgrade of the CMS Endcap Calorimeter. In: *Tech. Rep. CERN-LHCC-2017-023, CMS-TDR-019*. Geneva: CERN (2017). doi:10.17181/CERN.IV8M.1JY2
19. Pedro K. Current and Future Performance of the CMS Simulation. *EPJ Web Conf* (2019) 214:02036. doi:10.1051/epjconf/201921402036
20. Pedro K. Integration and Performance of New Technologies in the CMS Simulation. *EPJ Web Conf* (2020) 245:02020. doi:10.1051/epjconf/202024502020
21. Hoschek J, Lasser D. *Fundamentals of Computer Aided Geometric Design*. Natick, MA, USA: AK Peters, Ltd. (1993).
22. [Dataset] Frank M, Gaede F, Petric M, Sailer A. *Aidasoft/dd4hep* (2018). Available from: http://dd4hep.cern.ch/ (Accessed April 5, 2022). doi:10.5281/zenodo.592244
23. Brun R, Bruyant F, Maire M, McPherson AC, Zanarini P. *GEANT 3: User's Guide Geant 3.10, Geant 3.11; Rev. Version*. Geneva: CERN (1987).
24. Foley J. *Computer Graphics : Principles and Practice*. Reading, Mass: Addison-Wesley (1995).
25. Stroud I. *Boundary Representation Modelling Techniques*. Berlin, Germany: Springer Science & Business Media (2006).
26. Frank M, Gaede F, Petrič M, Sailer A. Cad Support and New Developments in Dd4hep. *EPJ Web Conf* (2021) 251:03015. doi:10.1051/epjconf/202125103015
27. Amadio G, Ananya A, Apostolakis J, Bandieramonte M, Banerjee S, Bhattacharyya A, et al. GeantV. *Comput Softw Big Sci* (2021) 5:3. doi:10.1007/s41781-020-00048-6
28. [Dataset] AdePT code repository. *AdePT Code Repository* (2021). Available from: https://github.com/apt-sim/AdePT (Accessed April 5, 2022).
29. Apostolakis J, Bandieramonte M, Bitzes G, Brun R, Canal P, Carminati F, et al. Towards a High Performance Geometry Library for Particle-Detector Simulations. *J Phys Conf Ser* (2015) 608:012023. doi:10.1088/1742-6596/608/1/012023
30. Foley J. *Computer Graphics : Principles and Practice*. Reading, Mass: Addison-Wesley (1990).
31. Larsson T, Akenine-Möller T. A Dynamic Bounding Volume Hierarchy for Generalized Collision Detection. *Comput Graphics* (2006) 30:450–9. doi:10.1016/j.cag.2006.02.011
32. Parker SG, Bigler J, Dietrich A, Friedrich H, Hoberock J, Luebke D, et al. OptiX. *ACM Trans Graph* (2010) 29:1–13. doi:10.1145/1778765.1778803
33. Simon C B. Opticks : GPU Optical Photon Simulation for Particle Physics Using NVIDIA OptiX. *J Phys Conf Ser* (2017) 898:042001. doi:10.1088/1742-6596/898/4/042001
34. Markidis S, Chien SWD, Laure E, Peng IB, Vetter JS. NVIDIA Tensor Core Programmability, Performance & Precision. *CoRR abs/1803.04014* (2018).
35. Nevay LJ, Boogert ST, Snuverink J, Abramov A, Deacon LC, Garcia-Morales H, et al. BDSIM: An Accelerator Tracking Code with Particle-Matter Interactions. *Comput Phys Commun* (2020) 252:107200. doi:10.1016/j.cpc.2020.107200
36. Apostolakis JSFJ, Asai M, Banerjee S, Bianchi R, Canal P, Cenci R, et al. HEP Software Foundation Community White Paper Working Group - Detector Simulation. *arXiv:1803.04165* (2018). doi:10.2172/1437300
37. The CMS CollaborationSirunyan AM, Tumasyan A, Adam W, Andrejkovic JW, Bergauer T, et al. Measurements of Higgs Boson Production Cross Sections and Couplings in the Diphoton Decay Channel at $\sqrt{s}$ = 13 TeV. *JHEP* (2021) 07:027. doi:10.1007/JHEP07(2021)027
38. The ATLAS CollaborationAad G, Abbott B, Abdallah J, Abdelalim AA, Abdesselam A, Abdinov O, et al. The ATLAS Simulation Infrastructure. *Eur Phys J C* (2010) 70:823–74. doi:10.1140/epjc/s10052-010-1429-9
39. Aad G, et al. Measurements of Differential Cross-Sections in Top-Quark Pair Events with a High Transverse Momentum Top Quark and Limits on beyond the Standard Model Contributions to Top-Quark Pair Production with the ATLAS Detector at $\sqrt{s}$ = 13 TeV. *arXiv* (2022).
40. Sirunyan AM, Tumasyan A, Adam W, Ambrogi F, Asilar E, Bergauer T, et al. Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV Using the CMS Detector. *JINST* (2019) 14:P07004. doi:10.1088/1748-0221/14/07/P07004
41. LHCb Collaboration. Framework TDR for the LHCb Upgrade II - Opportunities in Flavour Physics, and beyond. In: *The HL-LHC Era. Tech. Rep. CERN-LHCC-2021-012, LHCB-TDR-023*. Geneva: CERN (2021).
42. Butler JN, Tabarelli de Fatis T. A MIP Timing Detector for the CMS Phase-2 Upgrade. In: *Tech. Rep. CERN-LHCC-2019-003, CMS-TDR-020*. Geneva: CERN (2019).
43. ATLAS Collaboration. Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade. In: *Tech. Rep. CERN-LHCC-2020-007, ATLAS-TDR-031*. Geneva: CERN (2020).
44. Johnson SR, Tognini SC, Canal P, Evans T, Jun SY, Lima G, et al. Novel Features and GPU Performance Analysis for EM Particle Transport in the Celeritas Code. *EPJ Web Conf* (2021) 251:03030. doi:10.1051/epjconf/202125103030
45. Howard A, Ivanchenko V, Novak M, Ribon A. Status of Geant4 Simulations of Calorimeters. *J Inst* (2020) 15:C05073. doi:10.1088/1748-0221/15/05/C05073
46. LHCb Collaboration. LHCb VELO Upgrade Technical Design Report. In: *Tech. Rep. CERN-LHCC-2013-021, LHCB-TDR-013*. Geneva: CERN (2013).
47. The CMS CollaborationSirunyan AM, Tumasyan A, Adam W, Ambrogi F, Asilar E, Bergauer T, et al. The Phase-2 Upgrade of the CMS Tracker. In: *Tech. Rep. CERN-LHCC-2017-009, CMS-TDR-014*. Geneva: CERN (2017).
48. ATLAS Collaboration. Technical Design Report for the ATLAS Inner Tracker Pixel Detector. In: *Tech. Rep. CERN-LHCC-2017-021, ATLAS-TDR-030*. Geneva: CERN (2017).
49. Drescher HJ, Hladik M, Ostapchenko S, Pierog T, Werner K. Parton-based Gribov-Regge Theory. *Phys Rep* (2001) 350:93–289. doi:10.1016/S0370-1573(00)00122-8
50. Grichine VM. Integral Cross-Sections of Light Nuclei in the Glauber-Gribov Representation. *Nucl Instr Methods Phys Res Section B: Beam Interactions Mater Atoms* (2018) 427:60–2. doi:10.1016/j.nimb.2018.04.044
51. Particle Data GroupZyla PA, Barnett RM, Beringer J, Dahl O, Dwyer DA, Groom DE, et al. Review of Particle Physics. *PTEP* (2020) 2020:083C01. doi:10.1093/ptep/ptaa104
52. Natochii A, Vahsen SE, Nakayama H, Ishibashi T, Terui S. Improved Simulation of Beam Backgrounds and Collimation at SuperKEKB. *Phys Rev Accel Beams* (2021) 24:081001. doi:10.1103/PhysRevAccelBeams.24.081001
53. Voutsinas G, Bacchetta N, Boscolo M, Janot P, Kolano A, Perez E, et al. Luminosity- and Beam- Induced Backgrounds for the FCC-Ee Interaction Region Design. In: 8th International Particle Accelerator Conference (2017). doi:10.18429/JACoW-IPAC2017-WEPIK004
54. Collamati F, Curatolo C, Lucchesi D, Mereghetti A, Mokhov N, Palmer M, et al. Advanced Assessment of Beam-Induced Background at a Muon Collider. *J Inst* (2021) 16:P11009. doi:10.1088/1748-0221/16/11/P11009
55. Aad G, Abbott B, Abbott DC, Abed Abud A, Abeling K, Abhayasinghe DK, et al. Emulating the Impact of Additional Proton-Proton Interactions in the ATLAS Simulation by Pre-sampling Sets of Inelastic Monte Carlo Events. *Comput Softw Big Sci* (2022) 6:3. doi:10.1007/s41781-021-00062-2

56. Lange DJ, Hildreth M, Ivantchenko VN, Osborne I. Upgrades for the CMS Simulation. *J Phys Conf Ser* (2015) 608:012056. doi:10.1088/1742-6596/608/1/012056
57. Kou E, Urquijo P, Altmannshofer W, Beaujean F, Bell G, Beneke M, et al. The Belle II Physics Book. *PTEP* (2019) 2019:123C01. [Erratum: PTEP 2020, 029201 (2020)]. doi:10.1093/ptep/ptz106
58. Dannheim D, Sailer A. Beam-Induced Backgrounds in the CLIC Detectors. In: *Tech. Rep. LCD-Note-2011-021*. Geneva: CERN (2012).
59. Bichsel H. Straggling in Thin Silicon Detectors. *Rev Mod Phys* (1988) 60: 663–99. doi:10.1103/RevModPhys.60.663
60. Aaboud M, Aad G, Abbott B, Abbott DC, Abdinov O, Abhayasinghe DK, et al. Modelling Radiation Damage to Pixel Sensors in the ATLAS Detector. *JINST* (2019) 14:P06012. doi:10.1088/1748-0221/14/06/P06012
61. Chiochia V, Swartz M, Bortoletto D, Cremaldi L, Cucciarelli S, Dorokhov A, et al. Simulation of Heavily Irradiated Silicon Pixel Sensors and Comparison with Test Beam Measurements. *IEEE Trans Nucl Sci* (2005) 52:1067–75. doi:10.1109/tns.2005.852748
62. Swartz M, Fehling D, Giurgiu G, Maksimovic P, Chiochia V. A New Technique for the Reconstruction, Validation, and Simulation of Hits in the CMS Pixel Detector. *PoS* (2008) VERTEX2007:035. doi:10.22323/1.057.0035
63. Muon Collider Collaboration. *Expected Detector Performance at the Muon Collider* (2021). to appear.
64. Gentit FX. Litrani: a General Purpose Monte-Carlo Program Simulating Light Propagation in Isotropic or Anisotropic Media. In: *Tech. Rep. CMS-NOTE-2001-044*. Geneva: CERN (2001).
65. Vinogradov S. Analytical Models of Probability Distribution and Excess Noise Factor of Solid State Photomultiplier Signals with Crosstalk. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2012) 695:247–51. doi:10.1016/j.nima.2011.11.086
66. Gallego L, Rosado J, Blanco F, Arqueros F. Modeling Crosstalk in Silicon Photomultipliers. *J Inst* (2013) 8:P05010. doi:10.1088/1748-0221/8/05/P05010
67. Oak Ridge National Laboratory. *Oak Ridge National Laboratory: Leadership Computing Facility* (2022). https://www.olcf.ornl.gov (Accessed April 5, 2022).
68. Argonne Leadership Computing Facility. *Argonne Leadership Computing Facility* (2022). [Online]. Available from: https://alcf.anl.gov (Accessed April 5, 2022).
69. National Energy Research Scientific Computing Center. *National Energy Research Scientific Computing Center* (2022). [Online]. Available from: https://www.nersc.gov (Accessed April 5, 2022).
70. PIZ DAINT. *PIZ DAINT* (2022). [Online]. Available from: https://www.cscs.ch/computers/piz-daint (Accessed April 5, 2022).
71. Our supercomputer davinci-1. *Our Supercomputer Davinci-1* (2021). [Online]. Available from: https://www.leonardo.com/en/innovation-technology/davinci-1-supercomputer (Accessed April 5, 2022).
72. KAROLINA. *KAROLINA* (2020). [Online]. Available from: https://www.it4i.cz/en/infrastructure/karolina (Accessed April 5, 2022).
73. LUMI. *LUMI* (2022). [Online]. Available from: https://lumi-supercomputer.eu (Accessed April 5, 2022).
74. Barcelona Supercomputing Center. *Barcelona Supercomputing Center* (2022). [Online]. Available from: https://www.bsc.es/marenostrum/marenostrum (Accessed April 5, 2022).
75. Supercomputer Fugaku. *Supercomputer Fugaku* (2021). [Online]. Available from: https://www.fujitsu.com/global/about/innovation/fugaku (Accessed April 5, 2022).
76. TOP500 LIST. *TOP500 LIST—NOVEMBER 2021* (2021). [Online]. Available from: https://www.top500.org/lists/top500/list/2021/11/ (Accessed April 5, 2022).
77. Okada S, Murakami K, Incerti S, Amako K, Sasaki T. MPEXS-DNA, a New GPU-Based Monte Carlo Simulator for Track Structures and Radiation Chemistry at Subcellular Scale. *Med Phys* (2019) 46:1483–500. doi:10.1002/mp.13370
78. Khusainov B, Kerrigan EC, Constantinides GA. Automatic Software and Computing Hardware Co-design for Predictive Control. *CoRR abs/1710.08802* (2017).
79. Rogers DWO. Fifty Years of Monte Carlo Simulations for Medical Physics. *Phys Med Biol* (2006) 51:R287–R301. doi:10.1088/0031-9155/51/13/R17
80. Antin G, Nieminen P, Daly E. Applications of GEANT4 for the ESA Space Programme. In: Astroparticle, Particle and Space Physics, Detectors and Medical Physics Applications - Proceedings of the 9th Conference (2006). p. 511–8. doi:10.1142/9789812773678_0085
81. Chavy-Macdonald M-A, Menicucci A, Santin G, Evans H, Jiggens PTA, Nieminen P, et al. High-accuracy Simulations of the ISS Radiation Environment and Applications to Interplanetary Manned Missions. *IEEE Trans Nucl Sci* (2013) 60:2427–34. doi:10.1109/TNS.2013.2273375
82. Eraerds T, Antonelli V, Davis C, Hall D, Hetherington O, Holland A, et al. Enhanced Simulations on the Athena/Wide Field Imager Instrumental Background. *J Astron Telesc Instrum Syst* (2021) 7. doi:10.1117/1.JATIS.7.3.034001
83. Park H, Paganetti H, Schuemann J, Jia X, Min CH. Monte Carlo Methods for Device Simulations in Radiation Therapy. *Phys Med Biol* (2021) 66:18TR01. doi:10.1088/1361-6560/ac1d1f
84. Baumann K-S, Kaupa S, Bach C, Engenhart-Cabillic R, Zink K. Monte Carlo Calculation of Perturbation Correction Factors for Air-Filled Ionization chambers in Clinical Proton Beams Using TOPAS/GEANT. *Z für Medizinische Physik* (2021) 31:175–91. doi:10.1016/j.zemedi.2020.08.004
85. Paganetti H. Range Uncertainties in Proton Therapy and the Role of Monte Carlo Simulations. *Int J Radiat Oncol Biol Phys* (2013) 86:396. doi:10.1016/j.ijrobp.2013.02.017
86. Matthiä D, Hassler DM, de Wet W, Ehresmann B, Firan A, Flores-McLaughlin J, et al. The Radiation Environment on the Surface of Mars - Summary of Model Calculations and Comparison to RAD Data. *Life Sci Space Res* (2017) 14:18–28. doi:10.1016/j.lssr.2017.06.003
87. Fracchiolla F, Engwall E, Janson M, Tamm F, Lorentini S, Fellin F, et al. Clinical Validation of a GPU-Based Monte Carlo Dose Engine of a Commercial Treatment Planning System for Pencil Beam Scanning Proton Therapy. *Physica Med* (2021) 88:226–34. doi:10.1016/j.ejmp.2021.07.012
88. Costa PR, Nersissian DY, Umisedo NK, Gonzales AHL, Fernández-Varea JM. A Comprehensive Monte Carlo Study of CT Dose Metrics Proposed by the AAPM Reports 111 and 200. *Med Phys* (2022) 49:201–18. doi:10.1002/mp.15306
89. Sarrut D, Bała M, Bardiès M, Bert J, Chauvin M, Chatzipapas K, et al. Advanced Monte Carlo Simulations of Emission Tomography Imaging Systems with GATE. *Phys Med Biol* (2021) 66:10TR03. doi:10.1088/1361-6560/abf276
90. Zacharatou Jarlskog C, Paganetti H. Physics Settings for Using the Geant4 Toolkit in Proton Therapy. *IEEE Trans Nucl Sci* (2008) 55:1018–25. doi:10.1109/TNS.2008.922816
91. España S, Zhu X, Daartz J, El Fakhri G, Bortfeld T, Paganetti H. The Reliability of Proton-Nuclear Interaction Cross-Section Data to Predict Proton-Induced PET Images in Proton Therapy. *Phys Med Biol* (2011) 56: 2687–98. doi:10.1088/0031-9155/56/9/003
92. Ramos-Méndez J, Schuemann J, Incerti S, Paganetti H, Schulte R, Faddegon B. Flagged Uniform Particle Splitting for Variance Reduction in Proton and Carbon Ion Track-Structure Simulations. *Phys Med Biol* (2017) 62:5908–25. doi:10.1088/1361-6560/aa7831
93. Plante I. A Review of Simulation Codes and Approaches for Radiation Chemistry. *Phys Med Biol* (2021) 66:03TR02. doi:10.1088/1361-6560/abbd19
94. Bernal MA, Bordage MC, Brown JMC, Davídková M, Delage E, El Bitar Z, et al. Track Structure Modeling in Liquid Water: A Review of the Geant4-DNA Very Low Energy Extension of the Geant4 Monte Carlo Simulation Toolkit. *Physica Med* (2015) 31:861–74. doi:10.1016/j.ejmp.2015.10.087
95. Schuemann J, McNamara AL, Ramos-Méndez J, Perl J, Held KD, Paganetti H, et al. TOPAS-nBio: An Extension to the TOPAS Simulation Toolkit for Cellular and Sub-cellular Radiobiology. *Radiat Res* (2018) 191:125–38. doi:10.1667/RR15226.1
96. Faddegon B, Ramos-Méndez J, Schuemann J, McNamara A, Shin J, Perl J, et al. The TOPAS Tool for Particle Simulation, a Monte Carlo Simulation

Tool for Physics, Biology and Clinical Research. *Physica Med* (2020) 72: 114–21. doi:10.1016/j.ejmp.2020.03.019
97. Santin G, Ivanchenko V, Evans H, Nieminen P, Daly E. GRAS: A General-Purpose 3-d Modular Simulation Tool for Space Environment Effects Analysis. *IEEE Trans Nucl Sci* (2005) 52:2294–9. doi:10.1109/TNS.2005.860749
98. Lei F, Truscott RR, Dyer CS, Quaghebeur B, Heynderickx D, Nieminen R, et al. MULASSIS: A Geant4-Based Multilayered Shielding Simulation Tool. *IEEE Trans Nucl Sci* (2002) 49:2788–93. doi:10.1109/TNS.2002.805351
99. Apollinari G, Béjar AI, Brüning O, Fessia P, Lamont M, Rossi L, et al. High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1. In: *CERN Yellow Reports: Monographs*. Geneva: CERN (2017). doi:10.2172/1767028
100. Cid Vidal X, D'Onofrio M, Fox PJ, Torre R, Ulmer KA. Report from Working Group 3: Beyond the Standard Model Physics at the HL-LHC and HE-LHC. *CERN Yellow Rep Monogr* (2019) 7:585–865. doi:10.23731/CYRM-2019-007.585
101. Abada A, Abbrescia M, Abdus Salam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. HE-LHC: The High-Energy Large Hadron Collider: Future Circular Collider Conceptual Design Report Volume 4. *Eur Phys J ST* (2019) 228:1109–382. doi:10.1140/epjst/e2019-900088-6
102. Bordry F, Benedikt M, Brüning O, Jowett J, Rossi L, Schulte D, et al. Machine Parameters and Projected Luminosity Performance of Proposed Future Colliders at CERN. *arXiv* (2018).
103. Agostini P, Aksakal H, Alekhin S, Allport PP, Andari N, Andre KDJ, et al. The Large Hadron-Electron Collider at the HL-LHC. *J Phys G* (2021) 48:110501. doi:10.1088/1361-6471/abf3ba
104. Benedikt M, Blondel A, Janot P, Mangano M, Zimmermann F. Future Circular Colliders Succeeding the LHC. *Nat Phys* (2020) 16:402–7. doi:10.1038/s41567-020-0856-2
105. Abada A, Abbrescia M, Abdus Salam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1. *Eur Phys J C* (2019) 79:474. doi:10.1140/epjc/s10052-019-6904-3
106. Abada A, Abbrescia M, Abdus Salam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. FCC-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2. *Eur Phys J ST* (2019) 228:261–623. doi:10.1140/epjst/e2019-900045-4
107. Abada A, Abbrescia M, Abdus Salam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. FCC-hh: The Hadron Collider: Future Circular Collider Conceptual Design Report Volume 3. *Eur Phys J ST* (2019) 228:755–1107. doi:10.1140/epjst/e2019-900087-0
108. CEPC Conceptual Design Report. *CEPC Conceptual Design Report: Volume 1 - Accelerator* (2018).
109. Abbrescia M, Ahmad M, Ai X, Albergo S, Abid Aleem M, Alexeev M, et al. *CEPC Conceptual Design Report: Volume 2 - Physics & Detector* (2018).
110. Schulte D. The International Muon Collider Collaboration. *JACoW* (2021) IPAC2021:THPAB017. doi:10.18429/JACoW-IPAC2021-THPAB017
111. Aarons G, Abe T, Abernathy J, Ablikim M, Abramowicz H, Adey D, et al. *ILC Reference Design Report Volume 1 - Executive Summary* (2007).
112. Aryshev A, Behnke T, Berggren M, Brau J, Craig N, Freitas A, et al. *The International Linear Collider: Report to Snowmass 2021* (2022).
113. Robson A, Burrows PN, Catalan Lasheras N, Linssen L, Petric M, Schulte D, et al. *The Compact Linear e^+e^- Collider (CLIC): Accelerator and Detector* (2018).
114. Roloff P, Franceschini R, Schnoor U, Wulzer A. *The Compact Linear e^+e^- Collider (CLIC): Physics Potential* (2018).
115. Bai M, Barklow T, Bartoldus R, Breidenbach M, Grenier P, Huang Z, et al. *C^3: A "Cool" Route to the Higgs Boson and beyond* (2021).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer VV declared a shared affiliation with the authors JA, GC, GC, AG, VI to the handling editor at the time of review.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Apostolakis, Bandieramonte, Banerjee, Bartosik, Corti, Cosmo, Elvira, Evans, Gheata, Pagan Griso, Ivantchenko, Jones, Klute, Leggett, Morgan, Novak, Pedro and Paganetti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

REVIEW
published: 16 June 2022
doi: 10.3389/fphy.2022.872691

Multi Gigabit Wireless Data Transfer in Detectors at Future Colliders

R. Brenner[1]*, *C. Dehos*[2] *and E. Locci*[3]

[1]Uppsala University, Uppsala, Sweden, [2]CEA, Leti, Université Grenoble-Alpes, Grenoble, France, [3]Gangneung-Wonju National University, Gangneung, South Korea

The WADAPT (Wireless Allowing Data And Power Transmission) consortium has been formed to identify the specific needs of different projects that might benefit from wireless communication technologies with the objective of providing a common platform for research and development in order to optimize effectiveness and cost. Wireless technologies have developed extremely fast over the last decade and are now mature enough to be a promising alternative to cables and optical links, with a possibility of revolutionizing detector design. Although wireless readout has the qualities and properties to be used in many collider detectors, this article focuses on the transmission of large amount of data from vertex detectors at high rate, low power budget and in potential high radiation environment. For vertex detectors, the 60 GHz band has proven to be adequate and commercial products are already available, providing 6 Gbps data links. This technology allows efficient partitioning of detectors in topological regions of interest, with the possibility of adding intelligence on the detector to perform four-dimensional reconstruction of the tracks and vertices online, in order to attach the tracks to their vertex with great efficiency even in difficult experimental conditions, and conveniently substitutes a mass of materials (cables and connectors). Early transceiver module products have been successfully tested for signal confinement, crosstalk, electromagnetic immunity and resistance to radiation. In the long run, emerging 140 GHz bands could also be used for higher data rates (>100 Gbps) at future high energy and luminosity hadron colliders.

Keywords: wireless, data transfer, WADAPT, vertex detector, collider

OPEN ACCESS

Edited by:
Petra Merkel,
Fermi National Accelerator Laboratory (DOE), United States

Reviewed by:
David Christian,
Fermi National Accelerator Laboratory (DOE), United States
Jingbo Ye,
Southern Methodist University, United States

***Correspondence:**
R. Brenner
Richard.Brenner@cern.ch

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *09 February 2022*
Accepted: *21 April 2022*
Published: *16 June 2022*

Citation:
Brenner R, Dehos C and Locci E (2022) Multi Gigabit Wireless Data Transfer in Detectors at Future Colliders. Front. Phys. 10:872691. doi: 10.3389/fphy.2022.872691

INTRODUCTION

The WADAPT consortium [1] chose to firstly investigate the relevance of wireless techniques for data transfer within vertex detectors, where commercially-available products in the 60 GHz band [2] would be adequate. The present ATLAS [3] and CMS [4] experiments have been designed for nominal operation at luminosity of 2×10^{34} $cm^{-2}s^{-1}$ and 60 interactions per beam crossing. Pileup lead to a number of mismeasured or misidentified tracks, and add extra energy to calorimeter measurements. Pileup confuses the trigger and also the offline reconstruction and interpretation of events. It also increases the execution time for the reconstruction of events in the High Level Trigger and the offline analysis.

The High Luminosity LHC (HL-LHC) is potentially able to deliver a luminosity up to $7.5 \times 10^{34}cm^{-2}s^{-1}$, increasing the interaction rate and collision pileup beyond existing or envisioned data readout technologies. The detector systems in the experiment will currently need a trigger to sustain a pileup up to 200 interactions per 25 ns beam crossing. The first event selection and data reduction is typically done by a fast trigger decision within 3–6 microseconds after a collision where 40 million

events per second are currently reduced to about 100 kHz and at HL-LHC this will be increased by an order of magnitude. In present experiments this decision is based on information from all detector systems except the tracker. For HL-LHC the CMS experiment has developed a data reduction method with pT filtering in the tracker which opens for the use of tracking data in the trigger [5].

The tracker is the most granular detector system in the experiment that is used to associate charged particles with the correct interaction vertices. The tracking detector would ideally read out all information for every single bunch crossing, perform reconstruction in real time, and then apply trigger criteria to filter events or if possible directly process the data in an event filter farm. The challenge is however to read out the very granular detector in time for the fast trigger decision or before data in the pipelined front-end electronics gets overwritten. The data rate needed to read out all hit clusters of 1-2 pixels in the next generation pixel detectors is between 50 and 100 Tbps[1]. Data from tracking detectors at the two big LHC experiments are today transferred with optical links. The main benefit with optical links compared to wire links is that they are electrically decoupled and free from crosstalk. The weakness of optical links is the size of the connectors, their alignment constrains, and the sensitivity to mechanical damage. Wire/copper links suffer from bandwidth limitation at higher data rates. Furthermore, the tracker is built in a modular way that does not follow the topology of the physics event. The readout is bound to the modularity of the tracker, which is not optimal for a fast trigger decision based on event topology. The advantage of wireless technology is that it is not constrained to the mechanical modularity of the tracker. The technology offers already today data rate comparable with present optical links and no connectors are needed. The next generation of hadron colliders, which might be the envisioned, FCC, would even be more challenging with a peak luminosity of $30 \times 10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, leading to a pileup up to 1,000 interactions per 25 ns beam crossing, and more granular detectors. Although the current electron-positron colliders do not introduce stringent requirements on the trigger capabilities, the experiments would certainly benefit from the wireless technology. For both types of collider experiments, minimizing the amount of material in the region of the tracking detectors will reduce multiple scattering and nuclear interactions that degrade the precision on the measurement of track momentum and interaction vertices, and in addition will reduce the number of fake hits arising from secondaries. Both CMS and ATLAS upgrades aim at reducing passive material by embedding serial powering [7] or by using DC-DC converters with reduced mass [8].

Well-chosen detector technology and geometry, combined with wireless techniques might help in reducing the amount of cables and optimizing their path, and thus minimizing the geometrical inefficiency. The total or even partial removal of cables and connectors will result in cost reductions, simplified installation and repair, and reductions in detector dead material. These two last aspects are especially important in tracking detectors and they may become particularly important in case of limited access or/and hostile environment.

For a full exploitation of the advantages of integrating a wireless point-to-point readout, new architectures will be necessary.

MILLIMETER WAVE WIRELESS TECHNOLOGY

Data Rates and Frequency Band

According to the Shannon-Hartley theorem, the physical data rate at which information can be transmitted over a communication channel is limited by the channel bandwidth and the signal-to-noise ratio[2]. The fast development of the wireless technologies is linked to the improvements made in the electronics, especially the semiconductor devices and the new simulation capacity [9–15]. This technological evolution is expected to continue in the coming years. Some of the main strategies to improve the performance of radio communication systems are the following:

Over the last decade the spectral efficiency, meaning the data rate of the physical layer that could be transmitted in a given radio frequency bandwidth, has been increased by a factor 100. This efficiency is obtained by the use of higher order modulation coding scheme and multiple data streams (MIMO). The use of these advanced techniques requires intensive digital signal processing, increasing the complexity of the chipset design, the size, as well as the power consumption. The recent progress in terms of computing capability and energy per operation brought by the advanced FinFET and FDSOI technologies comes along with this increase in signal processing complexity, containing the power consumption, heat dissipation and manufacturing cost.

According to the Shannon-Hartley theorem, another strategy to increase the throughput is to increase the spectrum bandwidth. Since the radio frequency frontend and antenna have a bandwidth limited in a proportion of the carrier frequency (typically a maximum of 15–20%), then larger signal bandwidths require transmission at higher frequency, which became feasible, once again, with the progresses made on the electronics.

According to the Free Space Path Loss equation, the attenuation of a 60 GHz transmission at 1 m is about 68 dB, which is 21.6 dB higher than a 5 GHz transmission. Consequently at constant transmit power and using a given modulation scheme, the communication range will be 12 times shorter. This shorter signal range has the benefit that it will mitigate co-channel interference making a high-density channel-reuse design possible. The higher frequencies lead to smaller sizes of the RF passive components, including antennae, which can be easily integrated into electronic systems. The application-related antenna must be carefully chosen according to many parameters such as directivity/gain, angle aperture,

[1]The latest description of the pixel detector layout, including the number of staves per barrel and endcap layer, the number of sensors per stave and the estimated data rates per chip can be found in [6].

[2]Here the signal to noise ratio is defined as the average received signal over the bandwidth, and the noise is defined as the average noise or interference power over the bandwidth.

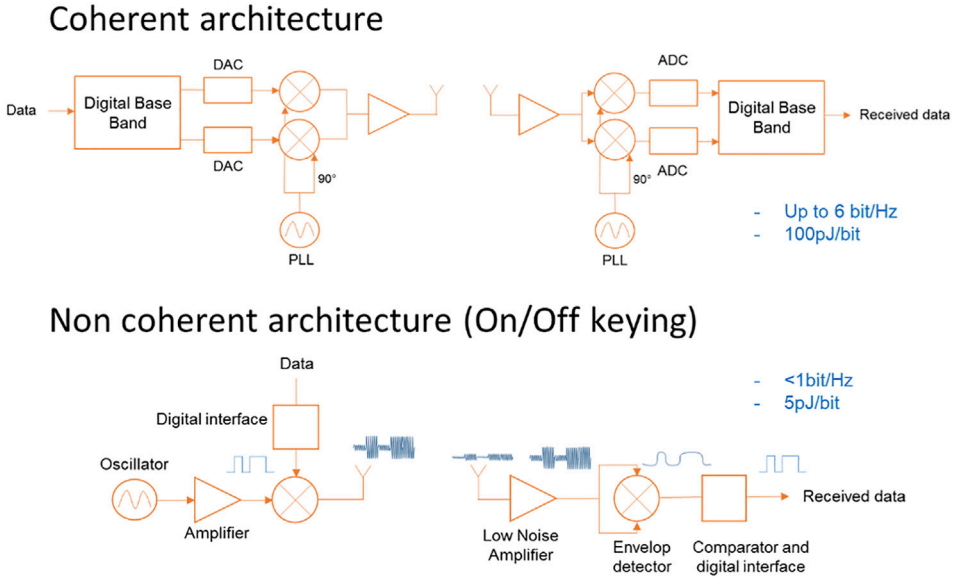


FIGURE 1 | Coherent **(A)** and non-coherent **(B)** RF architecture schematics.

bandwidth, etc. A wide range of antenna topologies can be produced on chip, in package or on advanced printed circuits boards. The use of RF lenses to increase the gain and the directivity of the antenna is also a possibility at this frequency. Improvements lie in carefully selecting the communication protocol. For a radio communication channel to work, a large amount of the transmission time is devoted to transmitter and receiver synchronisation, medium access, medium sensing, legacy standard-protection mechanism or transmission error compensation. The distribution of a system clock in particle physics detectors will facilitate the transmitter-receiver synchronisation.

Transceiver Architectures

Two main transceiver architectures may be considered at mmw frequencies (**Figure 1**).

Coherent architecture requires precise frequency synthesis (using Phase Locked Loop), and time and frequency synchronization mechanism achieved by digital base-band circuit [16]. This architecture allows complex amplitude and phase modulations, achieving high spectral efficiency, but leads to high power consumption.

On the contrary non-coherent architecture allows amplitude modulation only (On Off Keying, Amplitude Shift Keying or Pulse Amplitude Modulation). The receiver can implement a simple envelope detector, a comparator and digital interface, without any frequency reference or digital base band processing. The quadratic envelope detector provides negative gain, so that the overall receiver noise figure increases and affects the receiver sensitivity. For this reason, this architecture is more dedicated to short range and low power communication.

In the last decade there has been tremendous advances in silicon technologies that have made it possible to build high performance RF transceivers operating in the millimeter-wave (mmw) band. Non- coherent architecture brings high data rate (up to 10 Gbps at 60 GHz) and low power (<10 pJ/bit) for short-range applications [17]. Recent advance in antenna and packaging design allows the integration of the mmw antenna on chip [17] or in package [18] with good efficiency. Commercial products are now available for mass market applications with 3 main competitors (Lattice (SiBeam), Molex (Keyssa) and STMicroelectronics (**Figure 2**), proposing different integration schemes function of the application. Strong efforts have also been put on the compatibility with high/low speed digital protocols (USB3, PCIe, Gig Ethernet, UART) for a seamless cable replacement. In parallel a HEP dedicated 60 GHz Integrated Chip was under development in Heidelberg, using 130 nm SiGe BICMOS technology [19] in order to assess the feasibility and performance of the wireless link and establish solid foundation for designing the final reading system. The first prototype was designed to handle a data rate of 4.5 Gbps over a link distance of 1 m. Estimated power consumption for a first full prototype readout is about 150 mW.

Research is progressing [20] towards very high data rate wireless communication systems, using channel aggregation or channel bonding in W or D-band (110–170 GHz), where huge bandwidths (35 GHz aggregated) are available for short range communications. Channel bonding is particularly relevant for the aggregation of data from multiple detectors. A total of 16 channels of 7 Gbps each can be aggregated in a D-band transceiver, reaching 100 Gbps. The data can be sent over a few tens of meter range using directive planar antennas, such as transmit arrays [21], or over light plastic waveguide material [22].

TOWARDS WIRELESS INTEGRATION IN A DETECTOR

Possible wireless applications encompass HEP experiments at high-energy colliders, neutrino physics experiments, astroparticle-physics

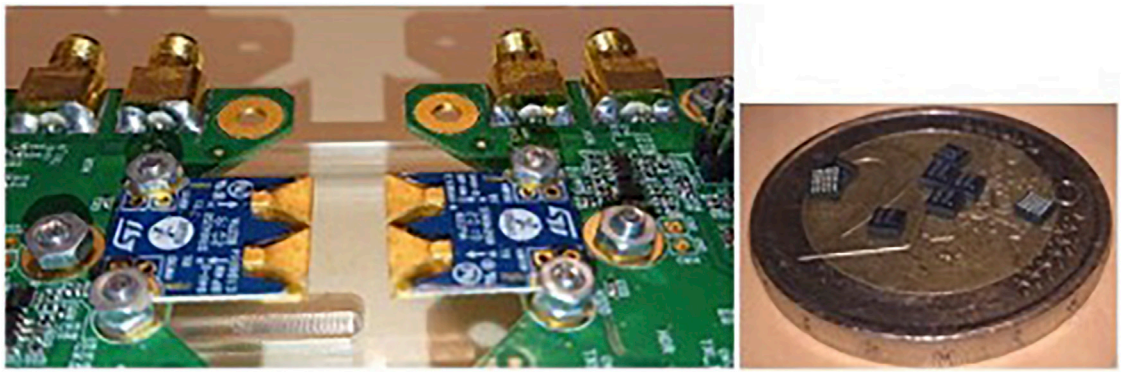


FIGURE 2 | ST Microelecronics ST60 contactless connectivity transceiver in BGA.

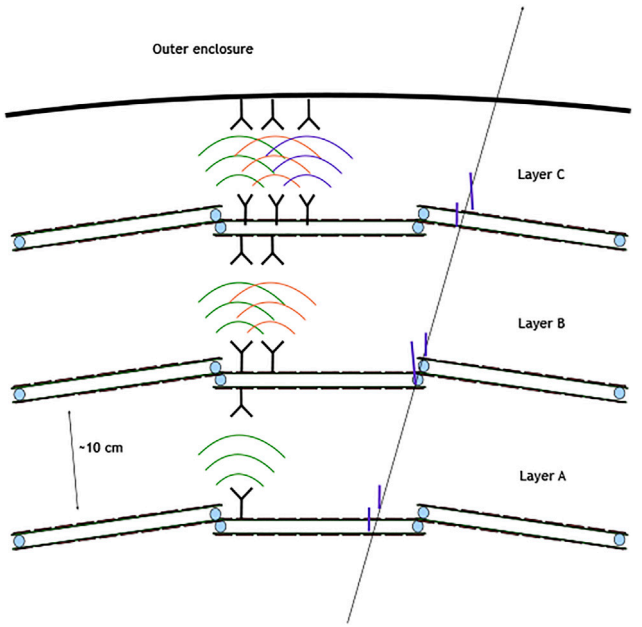


FIGURE 3 | Proposal of a radial readout for the tracking detector of the ATLAS experiment.

experiments and low energy experiments at the intensity frontier for a large range of particle physics detectors. The first step taken by the WADAPT consortium is to bring the proof of concept in the context of data transmission through a typical vertex detector.

Present and future collider experiments require more and more bandwidth for the handling of ever increasing data rates. It is especially true for highly granular tracking detectors operated at high beam luminosities. The performance is currently limited because of insufficient readout bandwidth. An example is the ATLAS Phase II New Inner Tracker Pixel Detector with five-barrel layers, four end-cap rings and a silicon strip detector with four layers and six end-cap disks, in 1 m radius. The readout should operate at up to 4 MHz (25 μs) L0 rate. At downlink a broadcast trigger and control signal will be sent at 160 Mbps while 10k up-links will transmit data at a rate of 5 Gbps. The data readout problematic, based on wired transmission strategy, is addressed in [23]. Different cabling schemes are evaluated with the goal of 5 Gbps data transmission per cable, while reducing as much as possible the material within the small detector volume. The proposed alternative for future colliders is to replace the innermost dense wired network by a multi-hop point-to-point wireless network, streaming the data radially from the centre to outer enclosure (**Figure 3**) [24].

The worldwide license-free 60 GHz band (57–66 GHz) [25] is very attractive in order to achieve high data rate transfer.

These features, the high path loss, high material penetration loss, narrow beam-width, Line-Of-Sight (LOS) mode of propagation, and operation in a controlled environment, makes the 60 GHz band optimal for short range operation as in the considered detector, where the signal propagates over a 10 cm distance between layers. In addition, the use of the high carrier frequency translates into small form factor for passives (antenna), which will reduce the material budget. This provides an extremely desirable frequency re-use that can handle a large number of transceivers in a small area as in the HEP detectors and other detector systems. Signal is routed through each layer from the bottom receiver to the transmitter at top side. This multi-hop concept requires that the wave front stay confined between the two considered layers, and does not crosstalks with neighbours.

Feasibility studies regarding the integration of the 60 GHz wireless technologies in silicon tracking detector were performed [26], based on a commercial coherent transceiver from Hittite [27]. Several aspects relevant for the implementation of 60 GHz links were studied: transmission losses, interference effects, absorbing materials and the influence of the antenna design and directivity, power consumption. The results of these studies are reported in **Sections 3.1–3.4**. More recently, tests have been conducted using non-coherent transceiver prototypes from STMicroelectronics [28], especially for latency, connectivity and radiation hardness.

Transmission and Reflection Tests

The data throughput of a wireless read-out system, as depicted in **Figure 3**, is maximized when links with maximum bandwidth are

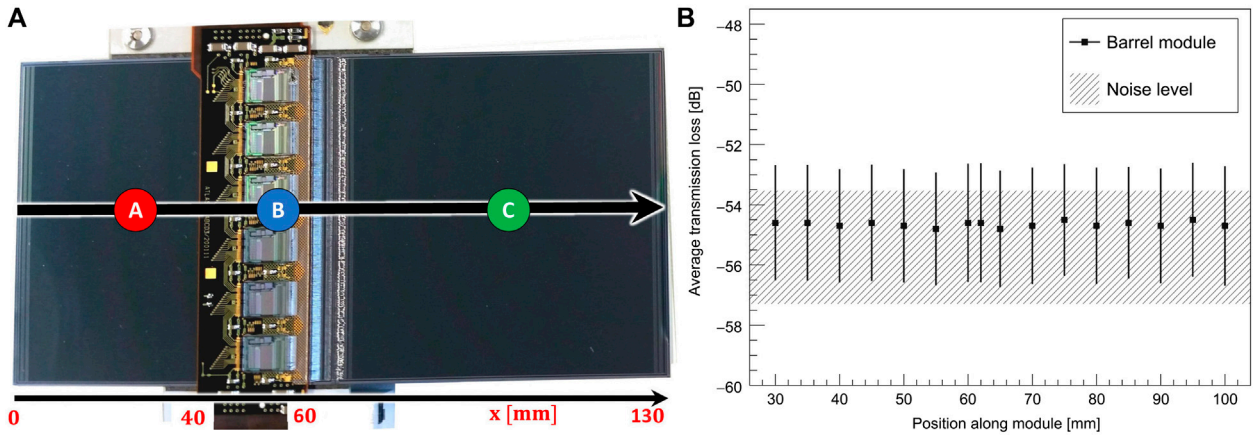


FIGURE 4 | (A) The ATLAS SCT barrel module under test. Positions for frequency scans are denoted by **(A)**, **(B)** and **(C)**. A position scan is performed along the black arrow. **(B)** Transmission loss of the barrel module averaged over the frequency band for a position scan (along the arrow in panel 4A) and the noise limited sensitivity of the spectrum analyzer. The uncertainties represent the RMS of the average measurements [26].

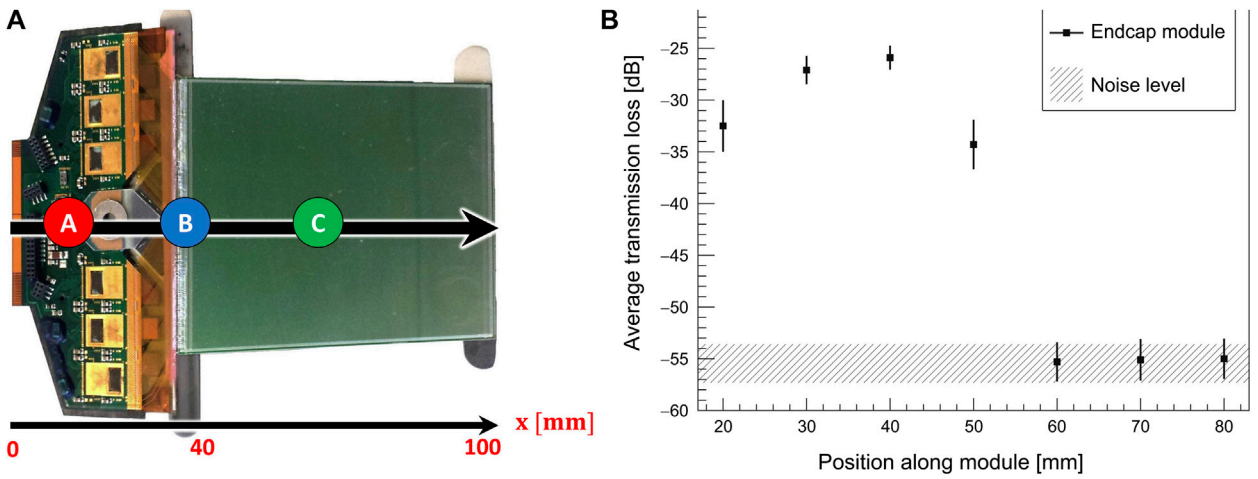


FIGURE 5 | (A) The ATLAS SCT endcap module under test. Positions for frequency scans are denoted by **(A)**, **(B)** and **(C)**. A position scan is performed along the black arrow. **(B)** Transmission loss of the endcap module averaged over the frequency band for a position scan (along the arrow in panel 5A) [26].

densely packed. The main challenge is to avoid crosstalk between parallel and subsequent chained links.

Transmission and reflection tests have been performed with a spare silicon strip detector module from the ATLAS barrel detector (**Figure 4A**). The module is mounted on a 2D movable stage and placed in line of sight of two horn antennas for transmission and reception. The module is irradiated with linearly polarized waves in the range from 57.3 to 61.3 GHz. A spectrum analyzer measures the power transmitted through the module in the radio frequency band without down conversion. The transmitted power is normalized to the power in the absence of module in-between. **Figure 4B** shows the transmission loss averaged over the chosen frequency band for a few positions along the module.

The measurement is repeated with a spare ATLAS SCT endcap module (**Figure 5A**). The frequency averaged transmission loss as function of the position on the module (**Figure 5B**) is −20 dB to −40 dB in the electronics region due to the assembly hole and the gap between hybrid and flex print. As expected the silicon strip region is opaque for the millimeter waves.

Most of the signal is absorbed by the silicon material, and small reflections may occur due to metallization of the chip and PCB. In some configurations these reflections might induce crosstalk. Absorbing materials (e.g., graphite foam) may be used to attenuate reflections [26].

Reduction of Crosstalk

All wireless links in a tracking detector are installed in stationary positions. Therefore, wireless readout systems can be designed and optimized in order to minimize the interference effects

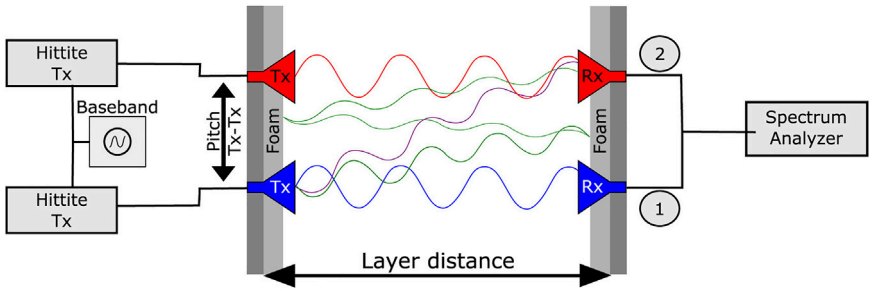


FIGURE 6 | Sketch of the setup to measure crosstalk with two links between highly reflective aluminium layers. LOS crosstalk is indicated as purple wave [26].

between different links. Nevertheless crosstalk may be an issue for extremely high link densities. However different actions can be combined to reduce crosstalk, if necessary.

One of them is acting on the antenna pattern, using directive antennas with high gain, small beam width and low secondary lobes. By increasing the gain of the antenna, the transmitted power can be decreased while keeping the signal over noise ratio (S/N) in the receiver constant.

In addition the antenna linear or circular polarization direction can be exploited. Crosstalk between adjacent links can be significantly reduced by using orthogonal polarization states.

Furthermore low mass absorbing material can be used to attenuate the spurious reflections without increasing the material budget.

Finally a tunable voltage controlled oscillator (VCO) with sharp filters or low bandwidth antennas may be used to divide the frequency band in different channels. Spreading (CDMA) is also an option with the drawback of reducing the data rate [26].

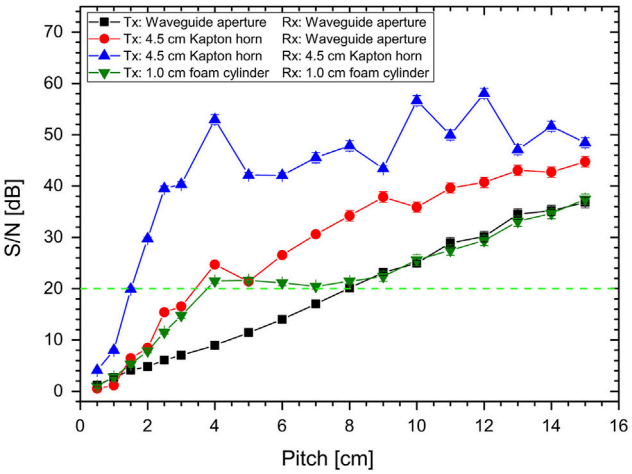


FIGURE 7 | S/N in the radio frequency spectrum with LOS induced crosstalk as function of the antenna pitch for different setups [26].

Link Density Studies

The different methods for reducing crosstalk discussed in the previous section have been experimentally studied by measuring how close two parallel 60 GHz links can be placed without interfering (**Figure 6**). The parameter S/N is defined as the ratio of the signals from transmitter #1 (useful signal) over transmitter #2 (interfering signal) measured with the spectrum analyzer at the antenna of the receiver #1. The link quality is considered to be acceptable when S/N is higher than 20 dB, which corresponds to a theoretical bit error rate (BER) of 10^{-12}, for high order modulation scheme decoding and forward error correction. For OOK the required S/N is in the range 30–35 dB function of the total jitter requirement.

The distance between transmitters and receivers is set at 10 cm (expected distance between layers) and both links have identical polarization. S/N is measured as a function of the antenna spacing for different setups (**Figure 7**). For reference non-directive waveguide apertures are used, for which S/N increases linearly as a function of the pitch.

In the absence of any mitigation a minimum pitch of about 8 cm is necessary to achieve a S/N of 20 dB. With one highly directive (17 dBi) Al-Kapton® horn antenna on the transmitting or receiving side, the minimum pitch is reduced to 4 cm, while with Al-Kapton® horn antennas installed on both sides, links can be placed as close as 2 cm next to each other without significant interference effects. Without using any directive antennas crosstalk can be reduced just by absorber shields. This has been tested by equipping all transmitters and receivers with 1 cm long, hollow graphite foam cylinders on top of the waveguide apertures to shield lateral radiation. Then shielded links can be placed as close as 4 cm next to each other.

The Bit Error Rate (BER) measures the performance of a data transmission device. With the setup described above the BER is measured for a data rate of 1.76 Gbps using Minimum Shift Keying (MSK) as modulation scheme. Both links use the same polarization state and the same carrier frequency. **Figure 8** shows the BER measured as function of the pitch in different configurations. With horn antennas bit error rates smaller than 10^{-12} are achieved for all considered pitches.

In a real tracking detector 60 GHz radiation can be partially reflected between the silicon detector layers, generating multipath and crosstalk. The silicon detector environment can be simulated by enclosing the volume with aluminium plates (**Figure 6**). The distance between the links is set to 5, 10 and 15 cm. The measurements relative to all antenna setups

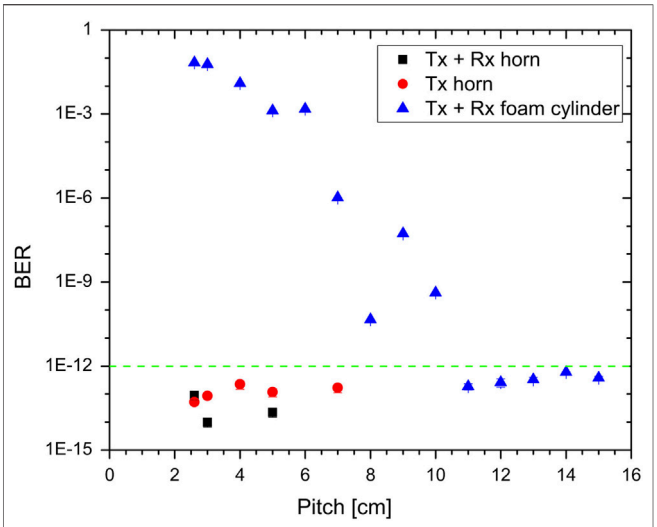


FIGURE 8 | Influence of LOS crosstalk on the BER of a wireless data transmission, shown as function of the pitch between two parallel links. Distance between transmitter and receiver is set to 10 cm. Both links are operated at the same carrier frequency [26].

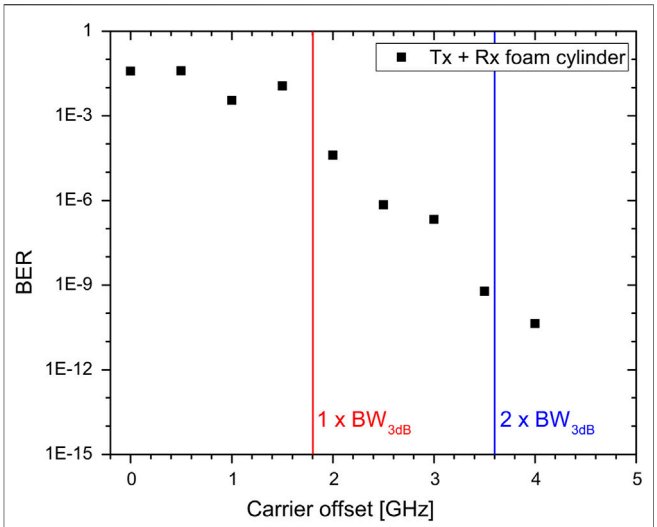


FIGURE 10 | Measured BER of two parallel-operated links at a pitch of 2.6 cm as function of the carrier frequency offset. The links are operated with MSK at up to 1.76 Gbps. No directive antennas are used. The 3 dB bandwidth of the Hititte transceivers (1.8 GHz) is indicated as red line, twice the bandwidth as a blue line [26].

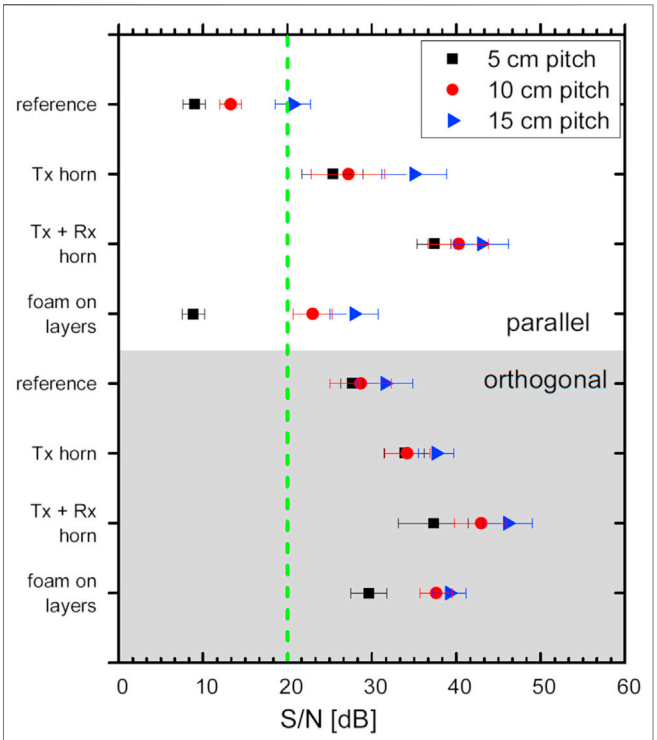


FIGURE 9 | S/N for two parallel links operated between two fully reflective aluminium layers at a distance of 10 cm. Results are shown for parallel polarization states (top, white background) and orthogonal polarization states (bottom, grey background) [26].

with both polarization states (parallel and orthogonal) between the two links are displayed in **Figure 9**. With directive antennas, stable data transmission at a pitch of 5 cm is possible for all studied configurations. For orthogonal polarization of neighbouring links, a good S/N is obtained for all tested setups even without directive antennas and a distance smaller than 5 cm seems to be possible. The combination of directive antennas and absorbing foam can increase the S/N ratio further. Thus using directive antennas, polarization, absorber materials, alone or in combination allows the operation of 60 GHz links between highly reflective materials even at a small pitch.

A transmitter chip with tunable carrier frequency allows the use of frequency channeling. The BER has been measured as a function of the carrier offset when two parallel links are positioned at a pitch of 2.6 cm using the same polarization state. Foam cylinders on top of the waveguide apertures provide a minimum shielding. The carrier frequency offset is varied in steps of 500 MHz. **Figure 10** shows that with this setup, using 3 frequency channels in parallel within the full 60 GHz band seems possible. Drawing full benefit from channeling necessitates filters with very sharp cutoff frequencies.

An estimate on the possible link density can be derived from above results. Assuming a tracking detector environment with distances of 10 cm between layers and a minimum pitch of 3.5 cm between links., OOK as modulation scheme with a spectral efficiency of $\rho_{OOK} \approx 0.5$ bps/Hz, and using the full 9 GHz bandwidth per channel yields a data rate of 4.5 Gbps per link, the resulting data rate area density is about 3.7 Tbps/m^2. This is a theoretical limit, which cannot be realistically reached experimentally. The estimated experimental data rate density is rather around 1.2 Tbps/m^2 [26].

Noise Pickup

The pickup of noise from 60 GHz communications is a potential worry for the operation of wireless links inside the detector. However,

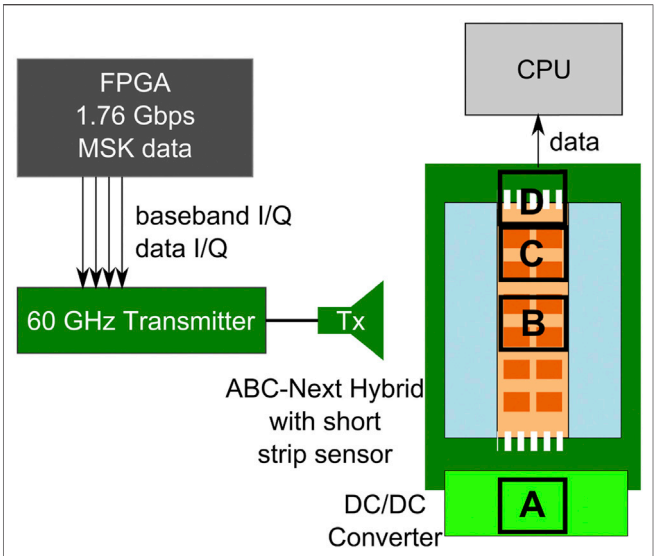


FIGURE 11 | Block diagram of the setup used for the irradiation test at Freiburg. A, B, C and D indicate different antenna positions used in the test. The 12 readout chips are illustrated as brown boxes on the orange readout hybrid [26].

as cut-off frequencies of sensors and readout chips for silicon detectors are typically below a few GHz, no interference between the 60 GHz links and tracking detector modules is expected. To demonstrate this, 60 GHz irradiation tests are performed for different silicon strip and pixel detectors.

Two ATLAS Binary Chip-Next (ABCN) endcap electronics hybrid prototypes for the Phase-2 upgrade of the silicon tracking-detector for the ATLAS experiment are tested using the setup shown on **Figure 11**. One is a bare electronics hybrid, while the other one is connected to short silicon strip sensors with a strip length of 2.4 cm. Each prototype includes 12 fully functional ABCN readout chips. Noise is measured for each channel of the readout ASICs under wireless irradiation. The hybrid module is irradiated using a 20 dBi horn antenna from an distance of 1 cm resulting in a irradiated power of about −1 dBm. Four different components are irradiated (**Figure 11**) with carrier frequencies of 57, 60 and 63 GHz: (A) the power converter, (B) and (C) the readout chips, and (D) the bonding wires. The noise level distribution of a reference measurement and a measurement with 60 GHz irradiation are compared in **Figure 12** for the electronics hybrid without silicon strips and no significant differences are found in the mean and width of the noise distributions. Similarly no significant noise increase for the second prototype connected to silicon strips is observed. As expected, no influence of the wireless signal on the noise level is found at any carrier frequency.

These results indicate that there should not be any obstacle to wireless data transmission in the 60 GHz band with respect to the operation and performance of current silicon tracking detector sensors [26].

Signal Integrity

The performance of non-coherent chips described in [2] was measured at 1–5 cm range between transmitter and receiver using low directivity antenna (**Figure 13**). The transmitter was fed by a SLVS 8b/10b encoded signal from pseudo random binary sequence at 5 Gbps. At the receiver output the binary signal showed less than 75 ps total jitter, less than 35 ps 20–80% rising/falling time and less than 1 ns latency (without the cables). A range of 10 cm can be obtained using 10 dB gain antenna.

Connectivity

A non-coherent transceiver was interfaced with Minimum Ionizing MOnolithic Active pixel Sensors (MIMOSA) from IPHC Strasbourg (**Figure 14**), which use the epitaxial layer of standard CMOS (TowerJazz High Resistivity 180 nm technology [29]) processes as detecting sensitive volume. The pixel detector board provides 5 µs refreshment rate and generates up to 3 Gbps data rate. Signal integrity tests were carried out with the detector feeding a wireless link. No error was reported at oscilloscope. Different materials were then introduced between the antennas. Signals well passed through the non-metallized area of the thin flex Kapton PCB that embeds chips, but not through low-density carbon foam that holds the detector. These results validate the proposed readout architecture, whose principle is to isolate every detector layer and export data through multiple hops between layers. At longer term, the RF transceiver may be integrated within the detector chip. The antenna could also be part of the chip with good efficiency, assuming the high resistivity substrate of the CMOS backend material. Pixel detectors still need to receive a clock or trigger reference from the network. It could then be worth modifying the 60 GHz transceiver, so that the clock is also transmitted wirelessly from chip to chip on the carrier by harmonic injection locking.

Radiation Hardness

A first irradiation experiment with protons was performed at Åbo Akademi University, Turku, Finland [30]. **Figure 15** shows a scheme and a picture of the setup. The radiation hardness of the V-band TRX chip in 65 nm CMOS, acting both in TX and RX modes, was

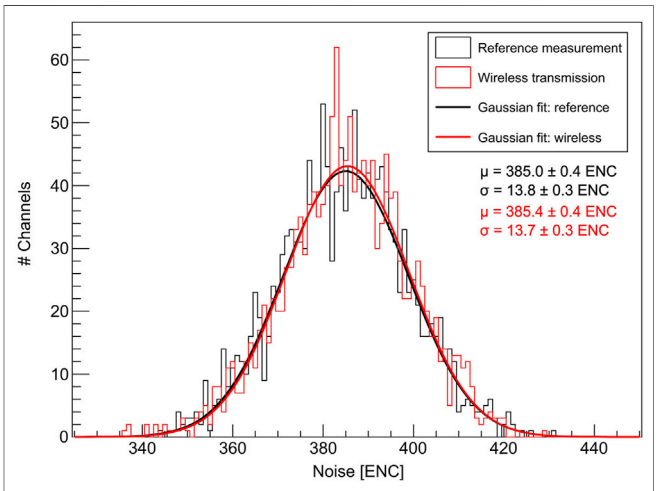


FIGURE 12 | Noise distribution of all channels of the 12 ABCN readout chips without strip sensors under irradiation. The reference measurement was performed without 60 GHz irradiation. Mean µ and width σ of the Gaussian fits are given. No significant difference is observed [26].

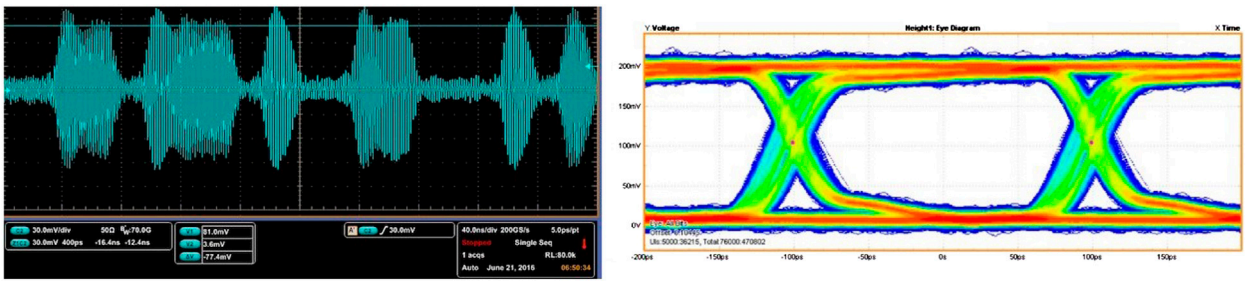


FIGURE 13 | Emitted waveform (left) and received eye diagramm (right) for 5 Gbps wireless link.

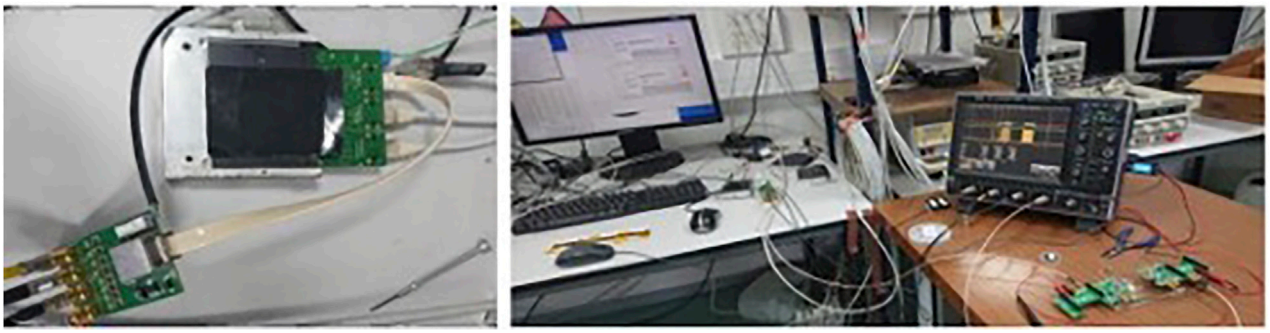


FIGURE 14 | Wireless data transmission interfacing with MIMOSA detector.

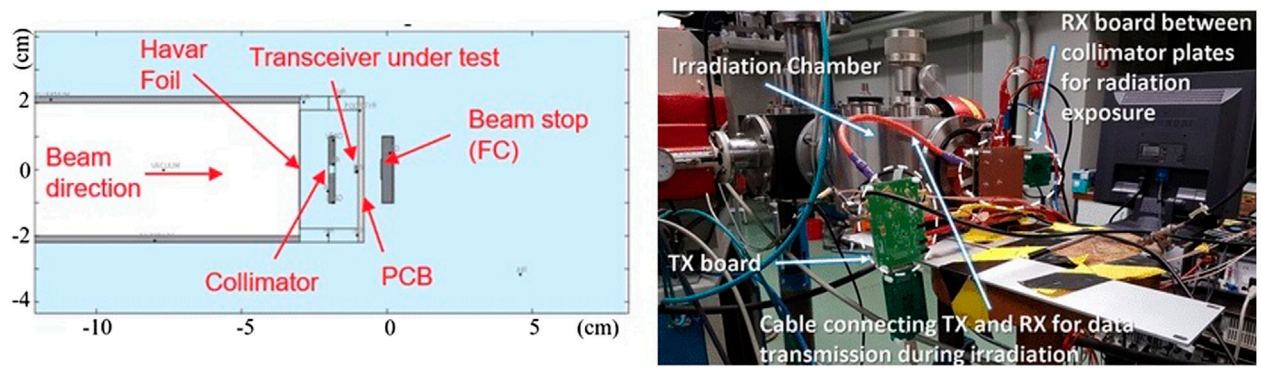


FIGURE 15 | irradiation tests at Turku Cyclotron.

investigated after 17 MeV proton irradiation. TID and fluence used during the experiment were 74 and 42 kGy and 1.4×10^{14}Neq/cm^2 and 0.8×10^{14}Neq/cm^2 for the RX and the TX, respectively. Transmission performance was measured before, during and after irradiation. The RX experienced higher loss of sensitivity than the TX power loss, accordingly to the respective TID and fluence. However, both the RX and the TX were found operational through over-the-air measurements at 5 Gbps after irradiation. If some transmission errors were observed during the irradiation, the chip recovered error free transmission after irradiation with small gain degradation.

Although the radiation levels expected in the real experiment at HL-LHC or future hadron colliders are much higher, this experiment presents a first investigation of radiation hardness that gives encouraging results despite the fact that the devices were not specifically radiation hardened by either design or process.

A second campaign of irradiation with electrons was carried out at the CLEAR (CERN Linear Electron Accelerator for Research) facility at CERN [31]. The TX and RX were exposed to a TID of 3.14 and 2.70 MGy respectively. The pre- and post-irradiation characterization of the chips revealed a 80 MHz upshift in the centre frequency. An overall link budget degradation of 10 dB was observed, which is a sum of 4 dB reduced TX output power and 6 dB degraded receiver sensitivity. Besides, circuit bandgap (voltage reference) were degraded, which affected the transceivers internal biasings and so

the transmitter and receiver gains. Despite the mentioned degradation, the chips were found operational in the post-irradiation characterization. Moreover, margins have been taken in chip design in order to guarantee performance through P,V,T variations and production yield, so that calibration means can restore part of the chip performance. Indeed TX power can be increased by 3dB by changing the settings, while LNA gain can be boosted also to restore the sensitivity. Radiation hardening shall however be considered on C65SPACE (which includes libraries with hardened cells, tested up to 0.3 Mrad dose but could handle higher radiations) or more advanced SOI technology for the highest radiation levels expected at future colliders such as HL-LHC or FCC-hh.

Future Developments and Test

At long term, the integration of wireless technologies in tracking detector would certainly require the design of a dedicated radiation hardened mmw transceiver IP, to be integrated either as a companion chip or within silicon trackers. For this purpose the ST C65SPACE radiation hardened ASIC technology or 45 nm Partially Depleted SOI technologies seem good candidates. The latter would allow the integration of antenna directly on the High Resistivity substrate of its back-end. For a better and versatile interfacing, the transceiver IP could also contain serializer/deserializer. External antennas or focusing lenses should be envisaged with the objectives of extending the range and ease crosstalk management. The crosstalk can be further improved by alternating antennas radiating in vertical and horizontal polarization for better isolation. The antenna should radiate a vertical beam (normal to the PCB plane) in order to fit easily in the detector.

At shorter term is envisaged the integration of the available non coherent mmw transceivers BGA with external antennas, as System In Package or System On Board. The antenna shall provide a narrow beam broadside for dense wireless links. As a proof of concept, a three (or four) layer silicon detector will be build[3], as shown on **Figure 3**. This mock-up of a central tracker will be equipped with the transceiver chip in BGA package. This will allow to study the added noise and data transmission quality (impact on eye diagram, Jitter, Bit Error Rate) over the different layers and get the antenna specifications: gain, polarization, bandwidth and radiation robustness. Then we will integrate the detectors to the setup to check the readout capabilities while using multi-hop wireless link. Eventually a multilink scenario will be considered to check if the isolation between channels is sufficient for the considered application. This would bring us closer to the full scale implementation and would help specify and integrate the future wireless systems in detectors at future colliders.

CONCLUSION

The recent advances in mmw technologies allows low latencies, high data rates and low power wireless data transmission at short range. The association of non-coherent RF architecture and directive antennas are particularly relevant for the considered innermost part of the readout network, based on multi-hop transmission. Early feasibility tests with prototypes and commercial products have demonstrated that crosstalk is not an issue for high density of links and that no pickup noise is to be expected. The projected theoretical limit with current technology is 3.7 Tbps/m2 and <10 pJ/bit. Although the commercial products were not designed nor manufactured as radiation hard components, they showed good behaviour under irradiation with protons and electrons. They could be easily hardened to be able to sustain radiation levels expected at future hadron colliders. There is no show-stopper to the early deployment and proof of concept using hardened RFICs and associated electronics in HEP experiments. The mmw technological roadmap shows soaring data rates using higher frequency bands above 100 GHz, challenging optical links for data aggregation and routing at short to medium range.

Well-chosen detector technology and geometry, combined with flexible wireless techniques might help in minimizing dead material by reducing the amount of cables and connectors. The absence of constrain to mechanical structures allows radial readout following the event topology that would be optimal for a fast trigger decision and would enhance on-detector intelligence, which could also open up new possibilities of communication between chips.

Technologies should be mature enough to be envisaged as a part of the readout network of future detectors at future colliders.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

The present document is based on the work of the whole WADAPT consortium[4], which has been partly presented at the 40th International Conference on High Energy physics—ICHEP2020 [2]. The authors are grateful to their colleagues of the Heidelberg University, Sebastian Dittmeier, André Schöning, Hans Kristian Soltveit, Dirk Wiedner, for giving them the permission to use in extenso large parts of their publication [26] for the sake of the completeness of this review. Special thanks are due to Sebastian Dittmeier for the tests of the Hittite transceiver, to Imran Aziz for the irradiation tests, to Frederic Lagrange from STMicroelectronics for providing equipment and support, to our colleagues[5] of the Institut Pluridisciplinaire Hubert Curien, Strasbourg, France for the tests of connectivity to MIMOSA sensors as well as to Åbo Akademi University, Finland for use of their irradiation facility. Warm thanks go to Hans Kristian Soltveit for the careful reading of this document.

[3]This project is supported by AIDAinnova.

[4]R. Brenner, D. Dancila, C. Dehos, P. De Lurgio, Z. Djurcic, G. Drake, J.L. Gonzalez Jiménez, L. Gustafsson, D.W. Kim, E. Locci, U. Pfeiffer, D. Röhrich, D. Rydberg, A. Schöning, A. Siligaris, H.K. Soltveit, K. Ullaland, P. Vincent, P.R. Vasquez, D. Wiedner, S. Yang.

[5]J. Baudot, G. Claus, M. Goffe.

REFERENCES

1. Brenner R, Ceuterickx S, Dehos C, De Lurgio P, Djurcic Z, Drake G, et al. *Development of Wireless Techniques in Data and Power Transmission Application for Particle-Physics Detectors*. (2017) CERN-LHCC-2017-002; LHCC-I-028 Available from: https://cds.cern.ch/record/2243480?ln=fr.
2. Dehos C, Locci E, Brenner R, Dancila D, de Lurgio P, Djurcic Z, et al. Wireless Allowing Data and Power Transfer. In: Proceesing of the 40th International Conference on High Energy physics - ICHEP (2020) ; Jul 2020; Prague, Czech Republic. p. 832.
3. The ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider. *J Instrum* (2008) 3:S08003. doi:10.1088/1748-0221/3/08/S08003
4. The CMS Collaboration. The CMS experiment at the CERN Large Hadron Collider. *J Instrum* (2008) 3:S08004. doi:10.1088/1748-0221/3/08/S08004
5. Béjar Alonso I, Brüning O, Fessia P, Lamont M, Rossi L, Tavian L, et al. High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report. In: *CERN Yellow Reports: Monographs* (2020). p. 378. CERN-2020-010. doi:10.23731/CYRM-2020-0010
6. ATLAS PUB Note. *Expected Size and Read-Out Rate of the Data of the ATLAS ITk Pixel Detector* (2022). Available from: https://cds.cern.ch/record/2800852/files/ATL-ITK-PUB-2022-001.pdf.
7. Sguazzoni G. The CMS Pixel Detector for the High Luminosity LHC. *Nucl.Instrum.Meth.A* (2019) 958:162089. doi:10.1016/j.nima.2019.04.043
8. Affolder A, Allongue B, Blanchot G, Faccio F, Fuentes C, Greenall A, et al. DC-DC Converters with Reduced Mass for Trackers at the HL-LHC. *J Inst* (2011) 6:C11035. doi:10.1088/1748-0221/6/11/c11035
9. Rappaport TS, Murdock JN, Gutierrez F. State of the Art in 60-GHz Integrated Circuits and Systems for Wireless Communications. *Proc IEEE* (2011) 99(No. 8):1390–436. doi:10.1109/jproc.2011.2143650
10. Siligaris A, Hamada Y, Mounet C, Raynaud C, Martineau B, Deparis N, et al. A 60 GHz Power Amplifier with 14.5 dBm Saturation Power and 25% Peak PAE in CMOS 65 Nm SOI. *IEEE J Solid-State Circuits (Jssc)* (2010) 45(7):1286–94. doi:10.1109/jssc.2010.2049456
11. Siligaris A, Richard O, Martineau B, Mounet C, Chaix F, Ferragut R, et al. A 65 Nm CMOS Fully Integrated Transceiver Module for 60 GHz Wireless HD Applications. In: Proceeding of the International Solid-State Circuits Conference (ISSCC); San Francisco (2011). p. 20–4.
12. Siligaris A, Richard O, Martineau B, Mounet C, Chaix F, Ferragut R, et al. A 65-nm CMOS Fully Integrated Transceiver Module for 60-GHz Wireless HD Applications. *IEEE J Solid-State Circuits*, 46 (2011). p. 3005–17. doi:10.1109/jssc.2011.2166662
13. Kaouach H, Dussopt L, Lanteri J, Koleck T, Sauleau R. Wideband Low-Loss Linear and Circular Polarization Transmit-Arrays in V-Band. *IEEE Trans Antennae Propagation* (2011) 59:2513–23. doi:10.1109/tap.2011.2152331
14. Pfeiffer UR, Goren D. A 20 dBm Fully-Integrated 60 GHz SiGe Power Amplifier with Automatic Level Control. *IEEE J Solid-state Circuits* (2007) 42(7):1455–63. doi:10.1109/jssc.2007.899116
15. Reynolds SK, Floyd BA, Pfeiffer UR, Beukema T, Grzyb J, Haymes C, et al. A Silicon 60-GHz Receiver and Transmitter Chipset for Broadband Communications. *IEEE J Solid-state Circuits* (2006) 41:2820–31. doi:10.1109/jssc.2006.884820
16. Diet A, Villegas M, Baudoin G, Robert F. Radio-communications Architectures. In: A Bazzi, editor. *Radio Communications*. Sciyo: Intech (2010). p. 1–35. ffhal-00553229f.
17. Siligaris A, Deparis N, Pilard R, Gloria D, Loyez C, Rolland N, et al. A 60 GHz UWB Impulse Radio Transmitter with Integrated Antenna in CMOS65nm SOI Technology. In: Proceeding of the 2011 IEEE 11th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems; Jan. 2011; Phoenix, AZ. IEEE (2011). p. 153–6. doi:10.1109/SIRF.2011.5719303
18. Dussopt L, Lamy Y, Joblot S, Lanteri J, Salti H, Bar P, et al. Silicon Interposer with Integrated Antenna Array for Millimeter-Wave Short-Range Communications. In: Proceeding of the 2012 IEEE/MTT-S International Microwave Symposium Digest; June 2012; Montreal QC. IEEE (2012). p. 1–3. doi:10.1109/MWSYM.2012.6259424
19. Soltveit HK. Multi-Gigabit Wireless Data Transfer for High EnergyPhysics Applications. In: Proceedings of The European Physical Society Conference on High Energy Physics — PoS(EPS-HEP2017); October 2017 (2017). doi:10.22323/1.314.0520
20. Gonzalez-Jimenez JL, Dehos C, Cassiau N, Siligaris A, Clemente A, D'Errico R, et al. Channel-bonding CMOS Transceiver for 100 Gbps Wireless point-to-point Links. *EURASIP J Wireless Commun Networking*, 2020 (2020). p. 117. doi:10.1186/s13638-020-01741-1
21. Manzillo FF, Luis Gonzalez-Jimenez J, Clemente A, Siligaris A, Blampey B, Dehos C. Low-cost, High-Gain Antenna Module Integrating a CMOS Frequency Multiplier Driver for Communications at D-Band. In: Proceedings of The 2019 IEEE Radio Frequency Integrated Circuits Symposium (RFIC); June 2019; Boston, MA, USA. IEEE (2019). p. 19–22. doi:10.1109/RFIC.2019.8701772
22. Voineau F, Dehos C, Martineau B, Sie M, Perchicot M, Nguyen NH, et al. A 12 Gb/s 64QAM and OFDM Compatible Millimeter-Wave Communication Link Using a Novel Plastic Waveguide Design. In: Proceedings of The 2018 IEEE Radio and Wireless Symposium (RWS); Jan 2018; Anaheim, CA. IEEE (2018). p. 250–2. doi:10.1109/RWS.2018.8305001
23. Muñoz Sánchez F. Module and Electronics Developments for the ATLAS ITk Pixel System. In: Proceedings of The 19th iWoRiD; Krakow, Poland (2017).
24. Brenner R, Cheng S. Multigigabit Wireless Transfer of Trigger Data through Millimetre Wave Technology. *J Instrumentation* (2010) 5:c07002. doi:10.1088/1748-0221/5/07/c07002
25. Dehos C, González JL, De Domenico A, Kténas D, Dussopt L. Millimeter-wave Access and Backhauling: the Solution to the Exponential Data Traffic Increase in 5G mobile Communications Systems? *IEEE Commun Mag* (2014) 52(9):88–95. doi:10.1109/MCOM.2014.6894457
26. Dittmeier S, Schöning A, Soltveit HK, Wiedner D. Feasibility Studies for a Wireless 60 GHz Tracking Detector Readout. *Nucl.Instrum.Meth.* (2016) A830:417–26. doi:10.1016/j.nima.2016.06.016
27. Analog Devices. *Millimeterwave Transmitter 57 - 64 GHz*. Available from: https://1224 www.analog.com/media/en/technical-documentation/data-sheets/hmc6000.pdf (Accessed 2012).
28. Wireless-Connectivity. *Wireless-Connectivity*. Available from: https://www.st.com/en/wireless-connectivity/st60a2g0.html (Accessed 2022).
29. Pernegger H. Monolithic Pixel Development in TowerJazz 180 Nm CMOS for the Outer Pixel Layers in the ATLAS experiment. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2018) 924:92–8. doi:10.1016/j.nima.2018.07.043
30. Aziz I, Dancila D, Dittmeier S, Siligaris A, Dehos C, De Lurgio PM, et al. Effects of Proton Irradiation on 60 GHz CMOS Transceiver Chip for Multi-Gbps Communication in High-Energy Physics Experiments. *J Eng* (2019) 8:5391. doi:10.1049/joe.2018.5402
31. Aziz I. Ultra-wideband Millimeter-waveAntenna Arrays and Front-End SystemsFor High Data Rate 5G and High Energy Physics. *PhD Thesis Digital Compr Summ Uppsala Dissertations Fac Sci Tech* (2021) 2005:68.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Brenner, Dehos and Locci. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

REVIEW
published: 22 June 2022
doi: 10.3389/fphy.2022.897719

Jets and Jet Substructure at Future Colliders

Johan Bonilla[1], Grigorios Chachamis[2], Barry M. Dillon[3], Sergei V. Chekanov[4], Robin Erbacher[1], Loukas Gouskos[5], Andreas Hinzmann[6], Stefan Höche[7], B. Todd Huffman[8], Ashutosh. V. Kotwal[9], Deepak Kar[10], Roman Kogler[11], Clemens Lange[12], Matt LeBlanc[5], Roy Lemmon[13], Christine McLean[14], Benjamin Nachman[15], Mark S. Neubauer[16], Tilman Plehn[3], Salvatore Rappoccio[14]*, Debarati Roy[17], Jennifer Roloff[18], Giordon Stark[19], Nhan Tran[7]*, Marcel Vos[20], Chih-Hsiang Yeh[21] and Shin-Shan Yu[21]*

[1]Department of Physics and Astronomy, University of California, Davis, Davis, CA, United States, [2]Laboratório de Instrumentação e Física Experimental de Partículas (LIP), Lisboa, Portugal, [3], Universität Heidelberg, Heidelberg, Germany, [4]HEP Division, Argonne National Laboratory, Lemont, IL, United States, [5]Experimental Physics Department, Organisation Européenne pour la Recherche Nucléaire (CERN), Geneva, Switzerland, [6]Department of Physics, University of Hamburg, Hamburg, Germany, [7]Fermi National Accelerator Laboratory, Batavia, IL, United States, [8]Department of Physics, Oxford University, Oxford, United Kingdom, [9]Department of Physics, Duke University, Durham, NC, United States, [10]School of Physics, University of Witwatersrand, Johannesburg, South Africa, [11]Deutsches Elektronen-Synchrotron, DESY, Hamburg, Germany, [12]Paul Scherrer Institute, Villigen, Switzerland, [13]Daresbury Laboratory, Warrington, United Kingdom, [14]University at Buffalo, State University of New York, Amherst, NY, United States, [15]Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States, [16]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, United States, [17]Amity Institute of Applied Sciences, Amity University Uttar Pradesh, Noida, India, [18]Brookhaven National Laboratory, Upton, NY, United States, [19]Santa Cruz Institute for Particle Physics, UC Santa Cruz, Santa Cruz, CA, United States, [20]IFIC (UV/CSIC) Valencia, Paterna, Spain, [21]Department of Physics and Center for High Energy and High Field Physics, National Central University, Taoyuan, Taiwan

OPEN ACCESS

Edited by:
Michael Schmitt,
Northwestern University,
United States

Reviewed by:
Annapaola De Cosa,
ETH Zürich, Switzerland
Janusz Gluza,
University of Silesia in Katowice,
Poland

***Correspondence:**
Benjamin Nachman
bpnachman@lbl.gov
Salvatore Rappoccio
srrappoc@buffalo.edu
Nhan Tran
ntran@fnal.gov

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: 16 March 2022
Accepted: 20 May 2022
Published: 22 June 2022

Citation:
Bonilla J, Chachamis G, Dillon BM, Chekanov SV, Erbacher R, Gouskos L, Hinzmann A, Höche S, Huffman BT, Kotwal AV, Kar D, Kogler R, Lange C, LeBlanc M, Lemmon R, McLean C, Nachman B, Neubauer MS, Plehn T, Rappoccio S, Roy D, Roloff J, Stark G, Tran N, Vos M, Yeh C-H and Yu S-S (2022) Jets and Jet Substructure at Future Colliders. Front. Phys. 10:897719. doi: 10.3389/fphy.2022.897719

Even though jet substructure was not an original design consideration for the Large Hadron Collider (LHC) experiments, it has emerged as an essential tool for the current physics program. We examine the role of jet substructure on the motivation for and design of future energy Frontier colliders. In particular, we discuss the need for a vibrant theory and experimental research and development program to extend jet substructure physics into the new regimes probed by future colliders. Jet substructure has organically evolved with a close connection between theorists and experimentalists and has catalyzed exciting innovations in both communities. We expect such developments will play an important role in the future energy Frontier physics program.

Keywords: jets, jet substructure, collider, artificial intelligence, machine learning, snowmass, top quark, Higgs boson

1 INTRODUCTION

Jets produced from high energy quarks and gluons through quantum chromodynamics (QCD) have a complex composition. This jet substructure (JSS) has emerged as a powerful framework for studying the Standard Model (SM) at particle colliders, and provides a key set of tools for probing nature at the highest energy scales accessible by terrestrial experiments [1–8]. While not an experimental or theoretical consideration of the Large Hadron Collider (LHC) experiments' original designs, JSS is now being widely used to extend the sensitivity of searches for new particles, to enhance the precision of measurements of highly-Lorentz-boosted SM particles, as well as to probe the fundamental and emergent properties of the strong force in new ways. Along the way, the JSS community has been a catalyst for new detector concepts, new analysis tools (e.g., using deep learning), new theory techniques, and more. Jet substructure has transformed the physics

program of the LHC and it can play a central role in the physics case for and the design considerations of future colliders.

As the particle physics community decides what the direction of the field should be in the middle part of the 21st century, it is useful to assess the state of JSS techniques that have developed over the last decades and to highlight the utility in various future collider scenarios. Efforts to investigate these scenarios are currently under way by the broader community, with pros and cons for many different strategies, for instance in the European Committee for Future Accelerators [9], and as part of the Snowmass 2021 process in the US (for which this paper is a contribution) [10]. While it is not yet clear what the future energy Frontier machine(s) will be, it is clear that jets and JSS will play an important role in the physics program of the future.

In this forward-looking perspective paper[1], we will investigate the opportunities and challenges associated with the various types of future colliders in the context of JSS. We will discuss both lepton and hadron colliders, including Higgs factories and ultra high energy machines. This paper is organized as follows. In **Section 2**, we give a brief introduction to various signatures of interest for JSS physics. We then outline the multiple avenues of research that the will be important in the context of Snowmass 2021. We believe that numerous topics of relevance for the Snowmass process should be discussed and evaluated with explicit considerations of the impacts for and benefits from JSS theory, phenomenology, and experimental tools (both hardware and software). These topics will be covered in a section on Theoretical Innovation (**Section 3**), Experimental Innovation (**Section 4**), and Enhancing Sensitivity (**Section 5**). We forgo a conclusion section in favor of the executive summary preceding this introduction.

2 SIGNATURES OF INTEREST

There are a large number of signatures that can benefit from JSS at future colliders. In general, JSS techniques are applied to tag Lorentz-boosted massive particles ($H/W/Z$ bosons, top quarks, and BSM particles) and to explore the structure of the strong force in final state radiation on small angular scales. This section briefly introduces various categories in the context of both SM measurements and BSM searches.

2.1 Light Quark and Gluon Jets

2.1.1 Measurements

High energy quark and gluon jets provide important probes of a variety of quantum chromodynamic (QCD) phenomena. These jets can be used to study perturbative aspects of QCD as well as features of QCD that cannot currently be described with perturbation theory. For the latter case, there are cases where scaling relations can be predicated and tested across a wide range of energies. These final states can be used to measure the strong coupling constant, to extract various universal objects within factorized QCD, to tune Parton Shower Monte Carlo generators, as well as other tasks. Quark and gluon jets were also studied at previous colliders, but higher energy machines allow for a suppression of non-perturbative effects as well as a larger lever arm for testing scaling behaviors.

2.1.2 Searches

Quark and gluon jets are statistically distinguishable due to their different fragmentation processes. Quark vs. gluon jet tagging has been a standard benchmark for the development of new classical and machine learning-based jet taggers. Many SM and BSM final states of interest are dominated either by quark or gluon jets, in contrast to the dominant background processes. Quark versus gluon jet tagging [11–15] can help enhance such signals, although these jets are not as seperable as other objects described below.

See also Refs. [16–18] for further details.

2.2 Bottom Quarks

Bottom quarks are prevalent in BSM decays as well as in the decays of H/Z bosons, and top quarks. Bottom quark jets are highly separable from other jets due to the long lifetime of the bottom quark and the heavy mass of bottom-flavored hadrons. In addition to lifetime information, jet substructure can be used to further separate these jets from other jets [19, 20].

A similar story is true to a lesser extent for charm quark jets [21, 22] and to an even lesser extent for strange quark jets [23–25].

2.3 *H* Boson

A main goal of the HL-LHC, as well as future Higgs factories, is to study the H boson [26, 27] in as much detail as possible. This includes detailed measurements of the branching fractions (BF). In the $H \to b\bar{b}$ and $H \to c\bar{c}$ final states, current analyses at the LHC [28–34] utilize kinematic criteria that identify hadronically decaying H bosons that have Lorentz factors larger than 1. These are moderately to fully boosted topologies. identify these final states. In addition, BSM physics that decay to H bosons (or other Higgs-like scalars) can also utilize these reconstruction techniques as is done in the current LHC experiments (a review can be found in Ref. [35]). In particular, specifically for bottom and charm quark final states, flavor and lifetime information can be used in addition to the jet substructure to improve categorization. Many all future collider scenarios result in copious Higgs bosons produced with large Lorentz boosts, so the techniques developed at the LHC will be broadly applicable for these cases as well.

2.4 *W/Z* Bosons

The cross sections and branching ratios of W and Z bosons are extremely well known *via* leptonic channels and previous LEP measurements [36]. However, W and Z bosons often participate in BSM scenarios, so can be present in many final states of these models (see a review in Ref. [37]). For example, in SM extensions with an additional real [38] or complex [39] scalar field S, resonant hS and SS production [40–42] can lead to an enhanced rate of highly-boosted W/Z bosons.

[1]This paper is not a review and is not comprehensive. See the reviews cited earlier for an in depth view of the state of JSS.

The identification of W and Z bosons is similar to the H boson, however the masses are slightly lower and they often do not decay to bottom or charm quarks, so there are fewer handles to use to identify them. This often leads to lower efficiency and purity in selections [43–47].

The phenomena of W and Z bosons radiating off of very high energy jets ("Weak-strahlung") is a new area that is not very likely at the LHC (although there are measurements of W bosons nearby jets [48]). Even standard QCD jets that originate from quarks or gluons can have additional information from jet substructure. New colliders provide an opportunity to study the phenomenon, and this may contribute to mis-identification rates of other algorithms.

In addition, W/Z bosons in vector-boson fusion initial states can also be more highly boosted than in $s-$ or $t-$ channel creation. Boosted techniques can also be used to appropriately identify these collisions at future colliders.

Finally, study of the vector boson scattering (VBS) process informs the degree to which the Higgs mechanism is the source of EWSB and thus provides an important test of the SM. Additionally, new physics that alters the quartic gauge couplings (QGC) [49, 50], or involves new resonances [51, 52], predict enhancements for VBS at high p_{T} of the vector bosons and invariant mass of the diboson system. At large m (VV) of most interest, 2 W/Z bosons are produced with large momentum and a hadronically-decaying boson can be reconstructed using boosted-boson tagging techniques that exploit jet substructure [43–47].

2.5 Top Quark

The top quark is a special quark with a Yukawa coupling close to unity. This makes it a likely participant in many BSM models to explain the hierarchy problem. The top quark nearly always decays to a W boson and a bottom quark [36]. At the LHC, even SM production of top quark pairs can result often in boosted final states [53–58]. In addition, many BSM scenarios have boosted top quarks participating in the event (a review can be found in Ref. [37]).

The jet substructure of top quarks is, in some sense, an ideal case, since there are two heavy SM particle masses to utilize (the top quark and W boson), as well as lifetime and flavor information in the final state particles. This provides a strong handle to identify top quarks.

Especially at higher-energy future colliders, the analysis of collisions containing top quarks will be ever more reliant on jet substructure and boosted topologies. Similarly to the W and Z bosons, there may also be top quark production within a jet that originates from light quarks or gluons *via* gluon splitting to $t\bar{t}$, similarly to the case at the Tevatron and LHC for bottom quarks. These types of events will need to be handled separately from events without these gluon splittings. Jet substructure and boosted techniques will play an increasingly important role here also.

2.5.1 Multi-Class Tagging

While most tagging studies focus on binary classification (one signal vs. one combined background), it is also possible to simultaneously tag multiple signals at the same time (see e.g., Ref. [59]). Multiclass classification methods output a score for each signal and background type that often corresponds to the probability that the jet belongs to the class given the inputs (with prior probabilities as in the dataset). While such approaches may not necessarily improve classification accuracy (with sufficient training examples), they can provide flexibility for downstream analyses.

2.6 Background Processes

For all of the signatures described above, there are a variety of physics backgrounds that obfuscate the target signatures. At hadron colliders, this is the result of multiple, nearly simultaneous collisions (pileup) as well as underlying event, and multiparton interactions. A variety of *jet grooming* techniques have been developed to mitigate these effects (see e.g., Refs. [1–4]). While similar backgrounds in e^+e^- are often much smaller, beam-induced backgrounds in muon colliders [60] could potentially benefit from similar techniques developed for hadron colliders.

3 THEORETICAL INNOVATION

In the last several decades, major advances in theoretical techniques have drastically improved our understanding of the nature of QCD radiation (a review is found in Ref. [5]). A combination of fixed-order, resummation, non-perturbative, and machine-learning techniques have opened new avenues of study, guided by extensive measurements of these processes at the LHC and elsewhere. Some studies of these topics with respect to collider scenarios is highlighted in Ref. [18]. In this Section, we focus on the developments of Monte-Carlo (MC) event generators, particularly as applied to new collider scenarios and to improve JSS modeling. Monte-Carlo (MC) event generators provide the link between the theoretical calculations and experimental measurements through a fully differential simulation of final states [61]. These are a combination of fixed-order, resummed, and non-perturbative effects [62–67]. At present, the majority of the uncertainty that lies in JSS is in the so-called "physics model" [68, 69], which includes the parton shower and hadronization, the former of which performs the QCD evolution, and the latter of which is performed with either the Lund string model [70, 71] or the cluster model [72–74].

The MC event generators most commonly used to compare to experimental measurements at the LHC are Herwig [75], Pythia [76] and Sherpa [77]. They contain various parton-shower models for the simulation of jet evolution, and cover a broad spectrum of matching and merging techniques. Several recent studies compared the physics performance of these generators for a large number of processes of relevance to the LHC [16, 78–85] and observed good agreement in their predictions for identical input parameters. For any given generator, the prediction may however strongly depend on those parameters, i.e., on the generator tune. Improvements in these tools will give a better event-by-event simulation of collisions with JSS, and will allow

better modeling of background processes as well as better inputs to advanced ML-based techniques.

One typical parametric uncertainty is the value of the strong coupling. Another common systematic uncertainty is the recoil scheme in the parton shower, which impacts a Monte-Carlo prediction in a different way than an analytical resummation, due to momentum and probability conserving effects in the event generator. These effects must however not influence the Monte-Carlo result in those regions where momentum conservation becomes irrelevant, and where analytic results can be obtained for certain observables. Much effort has been devoted recently to understanding these constraints in the context of parton-shower algorithms [86–88], and in providing parton showers that satisfy the theoretical boundary conditions [89–91]. In addition, some observables require the understanding of sub-leading color and spin effects, which are typically absent in parton-showers used for LHC physics. There has been renewed interest in implementing algorithms to include these spin correlations [92–96], and in including sub-leading color corrections for non-global observables [97, 98]. Some efforts have also been made to devise a generic approach for implementing higher-order corrections to the parton-shower splitting kernels in a fully differential form [99–103]. All these improvements will help to link analytic predictions for resummed jet observables to event generator predictions.

Systematic uncertainties also arise in the combination of fixed-order computations with parton showers. Matching algorithms for next-to-leading order (NLO) QCD calculations [104, 105] mainly differ in their treatment of real-radiative corrections. When observables become sensitive to either radiation (e.g., jet-p_T) or inhibited radiation (e.g., jet veto), this difference can create the dominant uncertainty. Similarly, merging algorithms, both at leading order [106–108] and at next-to-leading order [109–111] have associated uncertainties, which are mostly related to the matching algorithm for the underlying NLO calculations, and to the treatment of unitarity [111, 112]. Uncertainties in current NNLO matching algorithms arise from the precise technique being used to devise the resummed result at small transverse momentum in the case of resummation based approaches [113–115], and again from the treatment of unitarity in all approaches [111, 112, 116–118].

Finally, systematic undertainties may arise from the implementation of semi-hard physics effects, such as multiple scattering [119, 120] and hadronization [74, 121–123]. It is to be kept in mind that often the study of hadronization uncertainties is performed by replacing not only the hadronization model itself, but also the parton shower. This procedure is ill-advised, as the true hadronization uncertainty is almost always overestimated (however, see Ref. [124]). Studies using different hadronization models with identical perturbative input have demonstrated that in many cases the hadronization uncertainties are subdominant [84, 125, 126].

4 EXPERIMENTAL INNOVATION

Future colliders often have substantively different characteristics compared to the LHC. Higher-energy *pp* colliders will have more radiation and pileup, with the SM particles being produced with enormous Lorentz boosts and often in the forward region of the detector. Muon colliders will have beam-induced backgrounds. Electron-positron colliders have simpler environments due to lack of pileup and a precise measure of the z position of interactions. These all come with challenges and opportunities that can be exploited. This can come in the form of detector optimization for JSS, improved reconstruction algorithms, and in improved calibration and systematic uncertainties. These are covered in the following section.

4.1 Detector Optimization

There are several detector technologies that will improve JSS and related techniques. These include finer calorimeter granularity [127, 128], more hermetic coverage of tracking detectors, and precise measurements of timing information. The experience of the LHC has shown that such information can be used to more accurately reconstruct the interaction of hadrons with various detector elements, much of which is used in the "particle flow" (PF) [129] concept already deployed by the LHC experiments. At future muon colliders, "beam background" detectors could also in principle be deployed to reduce the impact on JSS.

4.1.1 Electron-Positron Colliders

The main detector concepts developed for electron-positron collider experiments are based on PF. With transparent, hermetic trackers and highly granular calorimeters, the ILD [130] and SiD [131] experiments at the ILC, as well as the CLIC detector [132] and the CLD design [133] for the FCC-*ee*, are designed to efficiently associate tracks and calorimeter energy deposits. A global detector R&D program has proven the feasibility of highly granular calorimeters [134] and large-scale systems are under construction for the ALICE [135], ATLAS [136], and CMS [137] upgrades. The optimization of the overall design was primarily driven by the jet energy resolution, but as a collateral benefit, these concepts offer excellent substructure performance. Jet substructure studies based on full simulation have been performed in Ref. [138].

In addition to the intrinsic particle identification capabilities, the fine transverse granularity allows close showers to be separated and provides good matching to tracks in the inner preshower signals, and also to muon tracks, making this calorimeter a good candidate for efficient particle-flow reconstruction. The need for disentangling signals produced by overlapping electromagnetic and hadron showers is likely to require longitudinal segmentation as well. Several ways to implement this segmentation were envisioned and are being studied, e.g., the classical division of the calorimeter in several compartments, an arrangement with fibres starting at different depths, the extended use of the timing information, etc. The specific advantages and drawbacks of each approach need to be studied through both simulations and beam tests. High-granularity calorimetry associated with a silicon tracker will be

a promising option to reach jet energy resolutions around 5%–20% with PF reconstruction.

4.1.2 Muon Collider

Proposed muon colliders offer a physics reach for discoveries similar to that of proposed high-energy hadron colliders, while maintaining appealing experimental aspects of lepton collider environments such as a lack of pileup and underlying event, as well as precise determination of the z position of the interaction. A critical difference between muon and electron accelerators is the presence of large beam-induced background (BIB) processes for muon machines, which arise due to muons in the beam decaying *via* $\mu \rightarrow e\nu\bar{\nu}$ before colliding. The resultant electrons interact with experimental elements along the beamline, creating electromagnetic showers of soft photons and neutral particles that can interact with detectors.

Detectors at future muon colliders will need to incorporate specifically-designed shielding and subsystems to mitigate BIB processes. The exact characteristics of the BIB depend strongly on the machine centre-of-mass energy and accelerator lattice, and must be studied in-detail for different scenarios. For studies during the Snowmass 2021 community planning exercise, the performance of a modified version of the CLIC detector has been benchmarked at a $\sqrt{s}$ =1.5 TeV (3 TeV) collider. This detector includes a modified vertex detector barrel that does not overlap with regions of large BIB activity, and shielding nozzles made of Tungsten and borated polyethlene to absorb contributions from beam-induced particles. Sets of adjacent sensors in the inner detector can also be used to mitigate contributions from BIB processes, by exploiting angular correlations as done in the CMS track trigger. The experimental conditions at a muon collider will also necessitate an increased material budget for the inner tracking systems, up to 10 times larger per-layer than that foreseen for ILC detectors due to additional cooling, power and support structures.

Early studies of this detector indicate that BIB contributions will be approximately evenly distributed in the calorimeter ($\eta - \phi$), suggesting that the advanced pileup mitigation techniques studied at the LHC could also provide a versatile handle with which to remove BIB contamination during reconstruction (**Section 4.2**). While the jet reconstruction efficiency for early jet reconstruction approaches at future muon colliders is above 90% for high-p_{T} jets, the decreased efficiency at lower jet p_{T} could also imply decreased performance when reconstruction jet substructure observables which rely on subjet identification (*e.g., N*-subjettiness [139, 140]) or soft radiation patterns (*e.g.,* D_2 [141]).

4.1.3 High-Energy Hadron Collider

There are currently two main hadron-hadron collider proposals, the FCC-hh at CERN and the SPPC in China, both targeting pp collisions at a center of mass energy of about 100 TeV. Driven by the physics requirements, the 100 TeV machine will deliver an integrated luminosity of around 25 ab^{-1} per experiment, reaching an instantaneous luminosity of $3 \times 10^{35}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, almost an order of magnitude larger than expected from the HL-LHC. These are extremely ambitious projects requiring breakthroughs in accelerator technology, detector design, and physics object reconstruction, and a coherent effort in all aspects is required.

To meet the physics requirements, the detectors for a 100 TeV machine should be able to reconstruct multi-TeV physics objects, while in parallel provide the necessary precision to measure the SM processes which typically results in high-energy final states at very high rapidity. The detector coverage should be extended with respect to the LHC detectors, since due to the almost a factor of five increase in the center of mass energy, many processes are expected to be extremely forward. For instance, SM ZZ production would produce two Z bosons with multi-TeV energies, with transverse momenta less than 100 GeV. These would have relativistic boosts of γ = 20, with opening angles between the Z boson decay products of about 0.1 radian. Detector capabilities to reconstruct these objects are fairly challenging (for instance, the average Z boson from ZZ production would shower mostly within one of the current LHC calorimeter cells). Concrete detector proposals are not yet in place, however different studies have been carried out to motivate the main aspects of the design.

An additional challenge is that the detector design should take in to account the harsh conditions expected at a 100 TeV machine. The foreseen upgrades of the LHC experiments for HL-LHC give a useful insight of the the challenges and the technology requirements expected in a future machine.

In addition to the extremely high energies that occur at very high rapidities necessitating finely granular detector elements, one of the big challenges at 100 TeV colliders is the large pileup. At the LHC, the average pileup is around 25, and it is expected to reach values of around 150–200 during the HL-LHC operation. This will result in significant degradation in the physics object reconstruction performance and hence on the physics outcome without dedicated detector systems and reconstruction algorithms. To this end, new developments are required in both the detector and reconstruction fronts. On the detector front, ATLAS and CMS experiments are developing fast timing detectors to improve the track-to-vertex association [142, 143]. These technologies achieve a timing resolution $\mathcal{O}(30)$ ps and studies using simulated samples show that are able to restore the physics object reconstruction performance obtained with much smaller pileup. At a 100 TeV machine, a factor of five larger pileup is expected posing even stringent criteria on the detector design. Likely, the developments on the precision timing detectors towards the HL-LHC will provide a solid ground to build upon. To cope with the pileup expected at 100 TeV, the timing resolution of the detectors should be improved by around a factor of 5-6, reaching a timing resolution better than 10 ps.

The calorimetry systems must provide excellent energy resolution over a wide range of energies in the central and forward regions, and increased hermetic coverage with respect to the LHC ones (reaching $|\eta| < 6$). Studies have shown [144] that another parameter of particular importance for JSS measurements in the ultra-relativistic regime, is the granularity of the detector. These studies showed that calorimeters must have 10 times finer granularity than the ones used at the LHC to achieve similar levels of performance in the main JSS observables in the this high-p_T regime. The extreme levels of radiation present

in a 100 TeV collider pose another challenge for the calorimeter design.

Technologies developed and successfully used at the LHC can serve as a promising starting point. One option for the electromagnetic and hadronic calorimeters, ECAL and HCAL, respectively can be based on the concepts used for the ATLAS calorimetry system. Their ECAL system uses Liquid Argon to generate the signal from the traversing particles. This technology provides both powerful performance together with the necessary radiation tolerance. In the case of the barrel region of HCAL, a more cost-efficient solution can be explored. For instance, the ATLAS HCAL uses organic scintillating tiles as active material. For the absorber, a combination of lead and steel provides promising results. However, due to the larger levels of radiation in the endcap and forward regions, this technology is not viable. Technologies based on liquid argon can be employed also in this case. Another option in this region could be a silicon-based or hybrid silicon/photomultiplier calorimeter similar to that being deployed by CMS in the HL-LHC upgrades, the High-Granularity Calorimeter (HGCAL) [145]. This also provides a large amount of resolution for substructure determination. This detector design also provides timing information (with $\mathcal{O}(30)$ps) allowing even for a 4D particle shower reconstruction. This approach can be powerful in suppressing the effect from pileup in the calorimeter system, and also aid the reconstruction of exotic signatures. The energy resolution in electromagnetic showers is characterized by a stochastic term ~16%/$\sqrt{E}$. Another idea for the ECAL system is based on monolithic active pixel sensors (MAPS). Studies in simulation using 50, ×, 50 μm pixels and a sensitive layer thickness of 18 μm yield a stochastic term of ~13%$\sqrt{E}$ [146].

The FCC-hh collaboration developed a baseline detector based on these principles [146]. The detector concept relies heavily on the ATLAS technology for both ECAL and HCAL, however changes in the design of the detector and its granularity have been considered. For instance, to complement the tracking system in JSS, the $\eta - \phi$ granularity of ECAL (HCAL) is around $\Delta\eta \times \Delta\phi = 0.01 \times 0.009$ ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$) in the barrel region, around four times finer compared to the LHC detector. This transverse granularity for the 100 TeV collision environment was determined using fast Monte Carlo simulations for boosted jets at tens-of-TeV scale [147]. Detailed studies [146, 148] using Full Simulation demonstrated that this technology could attain a stochastic and constant term of 8 (48)%/$\sqrt{E}$ and 0.2 (2)%, respectively for electromagnetic (hadronic) showers, with small dependence on $|\eta|$ and neglecting pileup interactions. This can attain jet energy resolutions of $< 5\%$ for jets with $p_T > 1$ TeV. Jet substructure variables for hadronic jets from highly Lorentz-boosted weak bosons from resonances between 5–40 TeV were studied in Ref. [127], using several spatial sizes of calorimeter cells. The current scale of LHC cell sizes around $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ were insufficient to ascertain the jet substructure. The study confirmed the HCAL design of the baseline FCC-hh with ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$). It is interesting to note that, for very boosted jets with transverse momenta close to 20 TeV, further decrease of cell size to ($\Delta\eta \times \Delta\phi = 0.0043 \times 0.0043$) did not show a further improvement in performance.

4.2 Reconstruction Algorithms

Reconstruction algorithms for jets and jet substructure have been widely developed in the last decade. Different collider scenarios can utilize different aspects of these advancements to address their unique challenges and opportunities as compared to the LHC. However, overall there are well-established techniques to achieve the desired performance level in all scenarios, as will be described in this section.

4.2.1 Jet Reconstruction

Precise and well-understood jet finding, clustering, and calibration is a key initial step to deploying powerful JSS techniques. The conceptual task is similar at the different future colliders under consideration. However, the varying energy range, whether or not the center-of-mass is known or not, and the level of beam backgrounds play a role in the optimal approach. Furthermore, good jet performance is reliant on well-understood and calibrated inputs for each of the subdetector elements and at the single particle level.

In the case of the ILD and FCC-ee detectors, software compensation has been shown to reduce the jet energy resolution significantly [149]. The lack of pileup results in smaller stochastic terms, and an optimal assignment of tracks and clusters in the particle flow algorithm can lead to superior energy resolutions. However, differences are present in the simulation of shower shapes, in particular the energy and radius of the interaction region, which need further studies and improved simulations [150]. Detailed measurements of the spatial and temporal development of showers in test-beam setups with fully integrated detector prototypes will help to improve the systematic uncertainties in the detector simulation, which is a crucial ingredient for precision measurements at future colliders.

For higher energies and more granular detector technologies, some initial studies have been performed. The energy calibration of calorimeter cells, composite clusters, single particles and jets is a challenging task at a 100 TeV pp collider. First studies exist on the energy calibration of the single-particle response of a FCC-hh detector, with electronic noise added to single cells and a simulation of in-time pileup [151]. In this study, energy deposits inside the calorimeter are summed into clusters using the sliding windows algorithm. For an optimal single-particle response, dead material corrections and a layer correction, accounting for the different sampling fractions depending on the depth of the shower, are necessary. The achieved jet resolutions are within the design goals with stochastic terms below 50%, but rely on extrapolations from detector simulations. Hadronic and electromagnetic shower components up to several TeV need to be simulated, where extrapolations to these high energies come with large uncertainties. Differences in the hadronic shower simulation models in Geant4 [152] have been reported for pions in the energy range between 2 and 10 GeV [153]. Detailed studies at higher energies will be needed to achieve the best possible precision at future colliders.

4.2.1.1 Electron-Positron Colliders

Compared to the previous generation of high-energy electron-positron colliders, the complexity of final states increases considerably. However, this complexity is comparable to that already observed in the LHC experiments. For instance, the hadronic Higgs-strahlung analysis at a Higgs factory requires excellent jet clustering performance in four-jet final states [154, 155]. At higher energy, di-Higgs, top quark pair and $t\bar{t}H$ production lead to six-jet and even eight-jet final states and jet clustering becomes the dominant experimental limitation [156]. Improved algorithms can have a profound impact on the potential to measure e.g., the Higgs self-coupling.

Machine-induced backgrounds at e^+e^- colliders are generally benign compared to the pile-up levels encountered at the LHC, but can have a non-negligible impact on jet reconstruction, especially at higher energy. The VLC algorithm [157] modifies the beam distance criterion of the generalized e^+e^- k_t algorithm and has been shown to provide a much more robust performance in comparison to the classical sequential clustering algorithms for e^+e^- collisions [156] in the presence of $\gamma\gamma \rightarrow$ background. A thrust-based algorithm is found to yield better performance than e^+e^- k_t in two-jet events at the CEPC [155]. The XCone algorithm [158] can naturally accommodate the boosted and resolved regimes and provides a close connection to calculations in Soft Collinear Effective Theory.

4.2.1.2 Muon Colliders

Since the major advantage of a muon collider is the ability to reach higher $\sqrt{s}$ than electron-positron colliders in a smaller area, the muon collider will produce final states that are generally more complicated than electron-positron colliders. Like other lepton colliders, the z position of the interaction is also known precisely, and there is no pileup as in hadron colliders. As such, it is expected that jet algorithms developed for electron-positron colliders should also apply well to muon colliders. However, due to typically higher energies, boosted topologies tend to be more prevalent.

4.2.1.3 Hadron Colliders

The challenges of a jet reconstruction at a hadron collider are well-known and extremely well-studied. While jet substructure reconstruction and tagging techniques were not directly considered in the design of the initial detectors at LHC and their reconstruction algorithms, they provide excellent performance after several years of evolution in algorithms. For future hadron colliders, jet substructure reconstruction is already considered in their design and it is expected that similar techniques as currently deployed at the LHC will find broad applicability.

The experiments at the LHC rely mostly on jets with a fixed distance parameter, where mostly the anti-k_t algorithm with R = 0.4, 0.8, 1.0 [159, 160] are used. The rigidity of the jet boundaries helps in pileup mitigation with an area-based approach [161–163] and the calibration of isolated jets [164, 165]. The experiences from the LHC have allowed extremely precise determination of jet energy and mass scales and resolutions (a review can be found in Ref. [6]), and have overall excellent precision.

While a larger value of R reduces hadronization corrections in jet p_T which scale as $1/R$, the influence of pileup and the underlying event increases with R^2 [166]. At the LHC experiments, it was possible to generally balance these effects with a few fixed-R algorithms, variable-R (VR) algorithms [167, 168] provide a promising alternative at future hadron colliders, which will have a larger dynamic range of jet energies, and have already been investigated by the LHC experiments [169–171].

Particle-flow algorithms [129, 172], or more generally algorithms combining tracking and calorimeter information [173] are the state-of-the art to reconstruct jet substructure with fine granularity and good energy resolution. Reconstruction challenges faced at the LHC, such as events containing up to 50 pileup interactions [172, 174], jet substructure of highly boosted W/Z/top quarks with multi-TeV transverse momenta, have been overcome successfully [171, 175].

4.2.2 Jet Substructure

4.2.2.1 Electron-Positron Colliders

Boosted object reconstruction at electron-positron colliders has been studied in full simulation by the CLIC group [138, 176], with a focus on boosted top quark tagging. This study confirms the excellent response of the CLIC detector concept for a large number of substructure observables.

In the electron-positron collider program at $\sqrt{s} \sim 250$ GeV, jet substructure finds applications in many measurements. A good example is the measurement of the Higgs coupling to gluons, where the differences between quark and gluon jets can be used together with flavour tagging information to distinguish the $H \rightarrow gg$ decay from $H \rightarrow b\bar{b}$ and $H \rightarrow c\bar{c}$. Jet substructure observables and grooming techniques are likely of value in determinations of the strong coupling α_S. This area has been identified as one of the open questions [177], but so far detailed phenomenology and experimental studies are lacking. A lepton collider also offers excellent opportunities for jet substructure measurements that can be used to develop a better understanding of fragmentation and hadronization. Carefully validated first-principle calculations and models for Monte Carlo generators in the clean e^+e^- without QCD radiation in the initial state can be very valuable in the preparation of a high-energy hadron collider.

4.2.2.2 Muon and Hadron Colliders in the Multi-TeV Regime

Boosted object reconstruction and tagging is crucial at a muon collider [178] or advanced linear collider [179, 180] operated at a center-of-mass energy in the multi-TeV regime, in addition the clear needs at future high energy hadron colliders.

The JSS tools developed in the last 2 decades provide a very solid baseline for the developments for a future 100 TeV hadron collider. However, in such high energies there are additional challenges to be met in both the detector design (discussed in **Section 4.1.3**) and in the development of the algorithms. First, the physics program at 100 TeV requires both standard model measurements with high precision using boosted objects in a

p_T-regime similar to the one at the (HL-)LHC (~0.5–2 TeV), while in parallel explore the energy Frontier with ultra relativistic particles with momenta up to ~10–15 TeV. Particularly in this p_T-regime, the decay products from the heavy objects result in hadronic jets that overlap significantly and are extremely difficult to reconstruct and explore the internal jet structure. It is therefore critical to have sufficient detector granularity in future colliders to sufficiently reconstruct JSS.

In addition to detector considerations, one algorithmic approach followed to overcome this challenge is to use only track-based variables for the design of the JSS algorithms, exploiting the much finer granularity of the tracking system compared to the calorimeters. These will be discussed in detail in **Section 5.2.1**.

4.2.3 Mitigating Beam Backgrounds

As noted above, beam backgrounds from electron-positron colliders is relatively benign compared to muon and hadron colliders though techniques have been developed to account for them.

4.2.3.1 Muon Colliders

At a high-energy muon collider, interactions of the decay products of the muon beams with accelerator and detector elements create an intense flux of particles through the experiments. While the background screening and mitigation strategy is still under development, it seems likely that the residual background level [60, 181] requires a combination of active background mitigation in the low-level reconstruction algorithm and the use of robust high-level reconstruction algorithms. Examples of possibly effective low-level reconstruction techniques could include vertex association, tracklet pointing, and timing information.

4.2.3.2 Hadron Colliders

Pileup mitigation is an important aspect of the low-level calibration of calorimeter cells, as pileup adds a diffuse noise term with large fluctuations. High granularity of the calorimeter is a pre-requisite for the mitigation of these effects, as it ensures an unambiguous combination with information from the tracking detectors. Machine learning techniques can help to improve the jet resolution by identifying electromagnetic deposits within jets, which are then calibrated to the electromagnetic scale. This can lead to an improvement in the single particle response by about 50% [151], but the existing studies need to be extended to more realistic conditions including pileup and electronics noise.

Jet reconstruction at the LHC is complicated by pileup. Pileup impacts jet reconstruction in a variety of ways, creating additional jets, changing the jet energy scale, and smearing out the jet energy resolution. It is particularly detrimental to jet substructure reconstruction, which can be affected by the presence of low-p_T pileup particles. LHC experiments use a combination of several strategies to reduce the impact of pileup, which have enabled high-quality jet substructure taggers and measurements, even under high pileup conditions. More study is needed to understand the impact of pileup on future hadron colliders, such as the FCC-hh and the SPPC, but the prospects for these colliders can be informed by the performance at the LHC.

Experiments at the LHC rely on a variety of different techniques to reduce the effect of pileup on jet reconstruction, including the topocluster reconstruction [182], particle flow using the primary vertex association for tracks [129, 172], Constituent Subtraction [183], SoftKiller [184], and the Pileup Per Particle Identification (PUPPI) algorithm [174, 185, 186]. For jet substructure reconstruction, grooming algorithms also provide some amount of pileup suppression, in addition to the other benefits they provide.

At the HL-LHC, pileup conditions will become even more challenging, with an average number of interactions per bunch crossing of around 200. Nevertheless, both ATLAS and CMS expect to maintain good performance, making use of detector upgrades and advanced reconstruction algorithms, based on studies of small-R and large-R jet reconstruction [187, 188]. Detector upgrades will also enable studies on the use of timing detectors [143, 189] and, in the case of CMS, a high granularity calorimeter [190]. Existing pileup mitigation algorithms will become even more important for jet reconstruction, and novel methods for pileup mitigation are also being explored, such as machine learning to improve pileup identification and subtraction [191–194].

Pileup conditions at the FCC-hh are expected to reach around five times those of the HL-LHC, with up to 1,000 simultaneous proton-proton collisions. With this density of interactions, high quality spatial and timing resolution will be critical in order to resolve the different pileup vertices and associate tracks to them. ATLAS and CMS both rely on vertex association of tracks to reduce pileup for particle flow algorithms, and in the case of CMS, for the PUPPI algorithm. This means that the use of 4D tracking will be critical for jet substructure reconstruction at the FCC-hh/SPPC. While charged particles are able to provide useful inputs to jet substructure reconstruction, neutral particles provide additional information that can be used to improve the performance of jet taggers. To use this information effectively will require advances in particle flow reconstruction in dense environments as well as dedicated pileup mitigation algorithms. The HL-LHC will enable critical studies of new tools which can be used to reduce pileup effects at future colliders like the FCC-hh and SPPC, such as the use of timing detectors for object reconstruction, as well as the development of pileup mitigation algorithms for reconstructed inputs.

5 ENHANCING SENSITIVITY

In this section, we highlight applications and techniques for using JSS information to enhance the sensitivity of both measurements and searches at colliders. First, we discuss novel and more exotic signatures of JSS which illustrates the broader application of the techniques we have discussed to search for potential new physics. Then, we will describe a number of important and emerging techniques for analyzing JSS information. In both cases, we cannot cover all approaches as JSS techniques are continually evolving in novel applications. Instead, we present here a broad

set of examples to give the reader a sense of the possibilities. In addition to the direct physics possibilities, JSS serves as a test bed for new and creative ideas in theory and analysis. The following section titles do not uniquely categorize the exam ples, which could be classified in a variety of ways.

5.1 Uncovered Scenarios

Traditional event reconstruction is mostly based on the principle that physics objects of interest can be individually reconstructed and well-isolated from other objects. However, SM and BSM signatures can give rise to highly collimated objects, manifesting in unusual topologies which are relatively rare at the (HL-)LHC, but will be much more prevelent at future colliders. Unconventional signatures can include cases where jets are composed of leptons and hadrons, only leptons, only photons, hadrons and missing transverse energy etc. In addition to the jet kinematics and JSS, the jet timing [195–197] information and other information can be used for classification. Examples include jets containing one or more hard leptons [198–202], displaced vertices [200], hard photons [203, 204], or significant missing transverse momentum [205–208]. Some of these anomalous signatures are already started being explored at the LHC [209–213]. Timing information can be useful to gain sensitivity in the searches with delayed jets [214]. It will also enhance the accuracy of prompt jet and MET reconstruction, that can boost the sensitivity of several new physics searches. Moreover other detector upgrades for high radiation tolerance, unprecedented granularity particularly in the forward region [145, 215], extension of the detector acceptance [216–218], a significantly sophisticated design upgrade of the trigger system [219] etc., will effectively lead us to broaden the search corners.

5.1.1 Photon Jets

Axion-like particles (ALPs) are predicted by several extensions of the SM (e.g., spontaneous breaking of a global symmetry, hierarchy problem, an interesting connection to the puzzle of dark matter). The discovery potential of ALPs in the future LHC era can well be estimated in the mass range of ALPs, which is inaccessible to previous experiments [203]. The jet kinematics and a few JSS variables (e.g., hadronic energy fraction of a jet, number of charged tracks in a jet, N-subjettiness, fraction of the jet p_T carried by the leading subjet, energy correlation function of the three hardest subjets) or the jet image study based on CNN technique [220, 221] are found to be extremely useful to disentangle photon-jet events from the single photon or QCD events. The so-called photon jet can be produced from the decay of boosted ALPs in the HL-LHC period. A detailed study of the reconstruction of a photon-jet, its calibration and performance in the future LHC environment or beyond needs to be carefully undertaken.

5.1.2 Delayed Jets

Several BSM predictions (e.g., supersymmetry (SUSY) with gauge-mediated SUSY breaking [222–225], hidden valley models [226], a Higgs boson decaying to glueballs where the Higgs boson is the portal to a dark QCD sector whose lightest states are the long-lived glueballs) [227, 228] lead to the unusual signature of non prompt or delayed jets which are sensitive to the proper measurement of jet timing [197]. These non prompt or delayed jets are usually modeled to be produced by the displaced decays of the heavy long-lived particles in BSM. The sensitivity of these long-lived particle searches using non prompt or delayed jets is found to be significantly enhanced by the precision timing information of the jet. The time profile of a jet can be used as an independent probe of jet properties. Similar to a choice of a jet clustering algorithm, the choice of a jet timing definition determines its properties and performance. The evaluation of various jet timing definitions is carried out depending on the closest representation to the parton level information as well as on the basis of minimizing the spread in the arrival times of the particles. Among the various jet timing definitions studied, the definition based on the p_T weighted sum of the arrival times of the jet constituents exhibits the most promising performance both for prompt and delayed jets. However, the jet timing performance of a prompt jet is estimated to depend on its η whereas the jet timing performance of a delayed jet is sensitive to the full kinematics of the event.

5.1.3 Dark QCD

Searches for dark matter (DM) particles in colliders have remained unsuccessful so far. Consequently in recent years, some focus has shifted to unusual final states, which are not covered by typical searches at the LHC. Semi-visible jets [205, 206] arise in strongly interacting dark sectors, where parton evolution includes dark sector emissions, resulting in jets overlapping with undetected particles that often result in missing transverse momentum aligned with one of the jets. This signature is usually discarded in the experiments, as it is usually from mismeasured jets. The implementation of semi-visible jets is done using the Pythia Hidden valley module [229, 230] to duplicate the QCD sector parton shower. In studies [207, 208], several jet substructure observables have been examined to compare semi-visible jets (signal) and light quark/gluon jets (background). The focus was on the more challenging scenario of t-channel production mode of semi-visible jets, where the absence of a resonance mass peak makes identifying the substructure difference more critical. The key parameter in the mode is the ratio of the rate of stable dark hadrons over the total rate of hadron, denoted by R_{inv}. In general, it was found that D_2 [141], C_2 [231] and ECF2 [231] observables were highly sensitive. The overall interpretation is that the semi-visible jets result in more multi-pronged substructure. This was verified by clustering stable dark hadrons in the jets, which resulted in the differences in the substructure observables disappearing. This indicate that the substructure becomes less two-pronged with visible and dark hadrons in them, and the absence of the dark hadrons create the two-pronged structure. Detailed studies of this phenomenon can be found in Ref. [232].

5.2 New Observables

As more information is obtained in the realm of JSS, new observables can be constructed that have interesting properties, either experimentally or theoretically. We discuss a few examples in this section.

5.2.1 Track-Based Observables

One of the challenges of an extremely high energy collider, such as a 100 TeV proton-proton collider, is so-called "hyper-boosted" jets, whose decay products will be collimated into areas the size of single calorimeter cells [233]. This fact, coupled with additional contamination from excess radiation - pileup, ISR, FSR, and UE, means that current jet substructure approaches will not be sufficient at a future high-energy hadron collider. One of the proposed mitigation strategies is to use track-based observables to augment calorimetric information [233–236].

Studies have shown [233, 237, 238] that this improves the identification performance in the ultra-relativistic limit. On the other hand, this result in imperfect measurement of the mass of the jet. Simple mass re-scaling techniques, e.g., as in Refs. [233, 237], or more sophisticated ML-based and other approaches utilized at the LHC [239, 240], provide promising solutions to improve the mass reconstruction. However, a calorimetry system with sufficient granularity can be very important in JSS, even at the ultra-relativistic regime, as detailed in [144]. Based on these results, calorimeters at 100 TeV should have $\mathcal{O}(10)$ finner granularity than the LHC calorimeters.

One study aimed to apply these strategies to the identification of hyper-boosted top quark jets [233]. First, the jet radius was scaled inversely with p_T in order to remove excess radiation. While calorimetric information was then still sufficient to measure jet energy, tracking information was added in order to resolve substructure information, including the following track-based observables:

- **Jet mass:** $m = \frac{p_T}{p_T^{tracks}} m_{tracks}$. The track-based mass m_{tracks} is scaled in order to recover the neutral particle information that is not measured by the tracker.
- **Prongy-ness variables:** N-subjettiness [241] and n-point energy correlation functions [242]. These variables measure the likelihood of a jet to have a given number of subjets, and in this case only include track information.

Clear improvement in the identification of top quarks vs. both light quark jets and gluon jets is seen when using these track-based observables, as compared to calorimeter-based observables.

In [236], it is shown that the HEPTopTagger [243, 244] can be successfully modified with a track-based approach (called HPTTopTagger) in order to identify tops at a future hadron collider. This technique is also applied to extremely boosted hadronic W- and Z-tagging, with the so-called HPTWTagger and HPTZTagger, respectively. Additionally in [245], it is pointed out that these track-based observables could be further enhanced by so-called "tracking calorimeters" in which detailed information about individual particle decays could be reconstructed [148].

5.2.2 $Flow_{n,5}$

Taking into account the extremely collimated nature of heavy object jets, the quantities $Flow_{n,5}$ are introduced in [246]:

$$Flow_{n,5} = \sum_p \frac{|p_T^p|}{|p_T^{jet}|}, \tag{1}$$

where n goes from 1 to 5 and p, p_T^p, and p_T^{jet} are the jet constituents, jet constituent transverse momentum and jet transverse momentum, respectively. The sum runs over the jet constituents so that the following holds:

$$\frac{n-1}{5} R \leq \Delta R(p, jet) < \frac{n}{5} R, \tag{2}$$

where $\Delta R(p, jet) = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is the angular separation between the jet axis and a particular jet constituent, and R is the jet size. These variables are applied to distinguishing boosted hadronically-decaying Z bosons from Randall Sundrum graviton decays ($G_{RS} \to ZZ$) to light quarks originating from $G_{RS} \to q\bar{q}$, specifically in the case where the G_{RS} mass is equal to 32 TeV. It is found that the combination of jet mass and the $Flow_{n,5}$ variables outperforms the combination of jet mass and the τ_2/τ_1 N-subjettiness ratio.

5.3 Novel Physics Effects

New showering effects begin to emerge at multi-TeV energies, including gluon splitting to top quark pairs, weak bosons radiating from jets, and radiation off of top quarks. These can affect boosted object identification overall, but these particularly affect boosted top quark identification, because they correspond to real on-shell weak bosons or top quarks, or enhanced radiation off of quarks. These effects are explored in the case of boosted top quark identification [245].

At very high energies, a gluon can split directly into a top quark pair [245]. This phenomenon will therefore increase gluon mistag rates. To mitigate this affect, it is important to recognize the fact that a gluon jet will have more constituents than a prompt top quark, and so the gluon-induced top will carry a smaller percentage of the jet's total energy. A useful discriminating variable would therefore be the transverse momentum ratio of a top-tagged subjet to its host large-radius jet: $p_{T,\text{top-subjet}}/p_{T,\text{fatjet}}$.

At extremely high energies, particles will radiate W, Z, and h bosons [245]. This can lead to light quarks jets that look like heavy particle jets. In the case of semi-leptonic top-tagging, W-strahlung can be particularly problematic. To mitigate this problem, one can take advantage of the fact that W-strahlung emissions peak at an angle of about $5m_W p_T$, whereas a top decay happens within a cone of approximately m_t/p_T [247]. Therefore, upper bound on the angle between the b-jet and the muon can be used to discriminate between tops and light quarks that radiate W bosons. Further study is required to mitigate the effect of weakstrahlung on other heavy particle tagging scenarios. Additional kinematic handles and AI/ML-based techniques may be deployed to provide further discrimination.

Similarly, identifying $WV \to \ell\nu qq$ from heavy particle decays is an important but challenging problem due to overlapping lepton and jet signatures [248]. ML-based taggers, such as convolutional (CNN) and/or fully connected (DNN) trained to distinguish signal (boosted WV) and background (QCD multijets) based on calorimeter and tracking features in jet

constituents, can be used to enhance sensitivity to new physics in future hadron colliders.

Even after applying a shrinking jet radius, a reconstructed top jet will still include some semi-hard final state radiation. This leads to around 10% of these jets having obfuscated substructure and masses well above the top mass. In this case, one can improve top/gluon discrimination by treating top quarks similarly to light quarks, and take advantage of the fact that a top will have less wide-angle radiation. As a simple example, it was shown that adding a track counting variable to a top tagger could improve discrimination, reducing gluon mistag rates by up to 20% [245].

5.4 Novel Analysis Techniques

The theoretical and experimental innovation discussed above will also require novel analysis techniques. These cannot be entirely of a computational nature, but will also require re-imagining the inputs to JSS and their processing. Selected examples on iterative generator tuning, anomaly detection, and machine-learning assisted techniques are outlined below.

5.4.1 Iterative Monte Carlo Generator Tuning

JSS techniques are sensitive to simulation effects such as underlying event and parton shower modeling, see also **Section 3**. While this directly affects the sensitivity of physics analyses that are making use of JSS techniques, it provides also the opportunity to constrain and improve physics modeling by performing dedicated measurements. In the past, these measurements have been performed in a one-off manner, i.e. the measurements are published [249, 250], then used for a future tuning campaign [251, 252] in a systematic way [253]. This approach, however, integrates over a huge phase space and often yields suboptimal values for JSS [251]. With the help of declarative and therefore consistently repeatable workflows [254] and machine-learning techniques, this approach can be significantly improved. The generator settings can be adjusted iteratively by repeating dedicated JSS measurements, consequently yielding optimal settings for the given suite of measurements. This can similarly be achieved by using machine-learning techniques to determine optimal generator settings (see e.g., Ref. [255]), for example to minimize related uncertainties. By adding further measurements, also not those directly related to JSS, significantly better simulation can be achieved. Enhanced tuning (and also a variety of measurements) may be enabled by unbinned and high-dimensional differential cross section measurements that are not possible by ML (see e.g., Ref. [256]).

5.4.2 Anomaly Detection

One of the most promising applications of machine-learning in ATLAS and CMS could be model-agnostic anomaly searches. There are a vast number of interesting new physics scenarios that we would like to search for at the LHC, however using traditional hypothesis testing techniques it is not possible to search for all them. Anomaly detection techniques aim to circumvent this problem by automatically identifying potential BSM contributions. These anomalies could be outliers (low probability density) or over/under-densities in phase space with respect to the SM. In this approach, a specific signal hypothesis is not required, although there is a tradeoff between performance on a given scenario and model dependence. Anomaly detection can be applied to individual objects/jets or to entire events. Modern deep-learning techniques dramatically increase the sensitivity of anomaly detection methods through their ability to use low-level, high-dimensional inputs. The technical concept behind these new anomaly searches is unsupervised, weakly supervised, and/or semi-supervised training of deep classification networks (see Refs. [257–260] for recent reviews).

A notable application is in the use of autoencoder neural networks optimised to compress and reconstruct event data. The accuracy of the reconstruction can then be used as the observable with which to identify the anomalies for instance in jets [208, 261–263]. Anomalous events may be expected to occur much less often in the data and thus result in less accurate reconstruction by the autoencoder. A promising path to improve this method is to extend the discriminative power from the physics phase space to include the latent space of the neural networks. This can be achieved, for example, using rapidity-mass matrices for standard autoencoders [264] (Dirichlet) variational autoencoders [265, 266] or invertible normalizing flow network [267], benchmarked for dark-matter-inspired jet signatures. For any kind of neural network application to jet physics, self-supervised learning of symmetries, fundamental invariances, and detector effects is an exciting new direction which is expected to significantly improve the understanding and the experimental stability of neural networks applied to subjet physics [268].

Related applications of anomaly detection, such as the classification without labels (CWoLa) method [269], are promising tools to enhance bump hunt analyses [270, 271]; this approach is also the first ML-based anomaly detection method to be applied to collider data [272]. In Ref. [258] the results of the LHC Olympics showcase many different methods on a resonant anomaly detection challenge. Recent developments have brought in a better understanding of these deep-learning techniques and new ideas for background estimation [273] and linearized explanations of decision classifiers [274, 275]. Ongoing and future work will certainly lead to more progress in all of these areas.

5.4.3 Hit-Based Inputs for High-Energy Flavor Tagging

Studies are on-going at the ATLAS and CMS experiments to incorporate some of the ideas first explored in [276]. When central jet energies exceed 500 GeV several effects make tracking difficult and the ability to discriminate jets containing B hadrons decreases. However, because a primary B hadron will often absorb most of the jet's energy, it has a high probability of crossing the innermost layer or layers of trackers in colliding beam machines prior to decay. Using the fact that hit patterns might "jump" from one layer to the next, or that charged tracks would cluster more tightly around the jet axis could be used as contributing input to sophisticated ML algorithms to improve their performance in the high energy regime. Initial studies are indicating that some additional efficiency can be gained up to

1,500 GeV with hit-based inputs added to neural network-based taggers [277]. If found effective, this technique might influence tracker design at future colliders where high energy jets will be even more common than currently at the LHC.

6 EXECUTIVE SUMMARY

In lieu of conclusions, we offer a summary of jet substructure at future colliders. Jet substructure (JSS) has emerged as a powerful framework for studying the Standard Model (SM) and provides a key set of tools for probing nature at the highest energy scales accessible by terrestrial experiments. While not an experimental or theoretical consideration of the design of the original LHC experiments, JSS is now being widely used to extend the sensitivity of searches for new particles, to enhance the precision of measurements of highly-Lorentz-boosted SM particles, as well as to probe the fundamental and emergent properties of the strong force in new ways. Along the way, the JSS community has been a catalyst for new detector concepts, new analysis tools (e.g., deep learning), new theory techniques, and more. Jet substructure has been transformative for the physics program of the LHC and it can play a central role in the physics case for future colliders.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

The funding was for support of the authors while writing this grant. This work was also submitted to the Snowmass 2021 Community Planning Exercise in the US. BN was supported by the Department of Energy, Office of Science under contract number DE-AC02-05CH11231. GC acknowledges support by the Fundação para a Ciência e a Tecnologia (Portugal) under project CERN/FIS-PAR/0024/2019 and contract "Investigador auxiliar FCT - Individual Call/03216/2017" and from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 824093. SR and CM were supported by the National Science Foundation under grant 2111229. AH gratefully acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) in the Emmy-Noether program (HI 1952/1-1 and HI 1952/1-2) and under Germany's Excellence Strategy—EXC 2121 "Quantum Universe"—390833306. DK is funded by National Research Foundation (NRF), South Africa through Competitive Programme for Rated Researchers (CPRR), Grant No: 118515. BMD was supported by a Postdoctoral Research Fellowship from Alexander von Humboldt Foundation. BH and RL acknowledge the support of STFC, United Kingdom. NT is supported by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the DOE, Office of Science, Office of High Energy Physics and the DOE Early Career Research program under Award No. DE-0000247070.

REFERENCES

1. Abdesselam A, Bergeaas Kuutmann E, Bitenc U, Brooijmans G, Butterworth J, Bruckman de Renstrom P, et al. Boosted Objects: A Probe of beyond the Standard Model Physics. *Eur Phys J C* (2011) 71:1661.
2. Altheimer A, Arora S, Asquith L, Brooijmans G, Butterworth J, Campanelli M, et al. Jet Substructure at the Tevatron and LHC: New Results, New Tools, New Benchmarks. *J Phys G: Nucl Part Phys* (2012) 39:063001. doi:10.1088/0954-3899/39/6/063001
3. Altheimer A, Arce A, Asquith L, Backus Mayes J, Kuutmann EB, Berger J, et al. Boosted Objects and Jet Substructure at the LHC. Report of BOOST2012, Held at IFIC Valencia, 23rd-27th of July 2012. *Eur Phys J C* (2014) 74:2792. doi:10.1140/epjc/s10052-014-2792-8
4. Adams D, Arce A, Asquith L, Backovic M, Barillari T, Berta P, et al. Towards an Understanding of the Correlations in Jet Substructure. *Eur Phys J C* (2015) 75:409.
5. Larkoski AJ, Moult I, Nachman B. Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning. *Phys Rep* (2020) 841:1–63. doi:10.1016/j.physrep.2019.11.001
6. Kogler R, Nachman B, Schmidt A, Asquith L, Campanelli M, Delitzsch C, et al. Jet Substructure at the Large Hadron Collider: Experimental Review. *Rev Mod Phys* (2019) 91:045003. doi:10.1103/revmodphys.91.045003
7. Marzani S, Soyez G, Spannowsky M. Looking inside Jets. In: *An Introduction to Jet Substructure and Boosted-Object Phenomenology, Vol. 958*. Berlin, Germany: Springer (2019). doi:10.1007/978-3-030-15709-8
8. Kogler R. Advances in Jet Substructure at the LHC. In: *Algorithms, Measurements and Searches for New Physical Phenomena, Vol. 284 of Springer Tracts Mod. Phys*. Berlin, Germany: Springer (2021). doi:10.1007/978-3-030-72858-8
9. Group EDRRP. The 2021 ECFA Detector Research and Development Roadmap. In: *Tech. Rep*. Geneva: CERN (2020). doi:10.17181/CERN.XDPL.W2EX
10. The Particle Physics Community Planning Exercise. *The Particle Physics Community Planning Exercise* (2021). Available from: https://snowmass21.org. (April 15, 2022).
11. Gallicchio J, Schwartz MD. Quark and Gluon Tagging at the LHC. *Phys Rev Lett* (2011) 107:172001. doi:10.1103/physrevlett.107.172001
12. Gallicchio J, Schwartz MD. Quark and Gluon Jet Substructure. *JHEP* (2013) 04:090. doi:10.1007/jhep04(2013)090
13. ATLAS collaboration. Light-quark and Gluon Jet Discrimination in Pp Collisions at $\sqrt{s}$ = 7 TeV with the ATLAS Detector. *Eur Phys J C* (2014) 74:3023.
14. CMS Collaboration collaboration. Performance of Quark/gluon Discrimination in 8 TeV Pp Data. In: *Tech. Rep*. Geneva: CERN (2013).
15. CMS Collaboration collaboration. Jet Algorithms Performance in 13 TeV Data. In: *Tech. Rep*. Geneva: CERN (2017).
16. Gras P, Höche S, Kar D, Larkoski A, Lönnblad L, Plätzer S, et al. Systematics of Quark/gluon Tagging. *JHEP* (2017) 07:091. doi:10.1007/jhep07(2017)091
17. Andersen JR, Bellm J, Bendavid J, Berger N, Bhatia D, Biedermann B, et al. Les Houches 2017: Physics at TeV Colliders Standard Model Working Group Report. *arXiv:1803.07977* (2018) 3.
18. Amoroso S, Azzurri P, Bendavid J, Bothmann E, Britzger D, Brooks H, et al. Les Houches 2019: Physics at TeV Colliders: Standard Model Working Group Report. In: *11th Les Houches Workshop on Physics at TeV Colliders: PhysTeV Les Houches* (2020). p. 3. 2003.01700.
19. CMS collaboration. Identification of Heavy-Flavour Jets with the CMS Detector in Pp Collisions at 13 TeV. *JINST* (2018) 13:P05011.
20. ATLAS collaboration. ATLAS B-Jet Identification Performance and Efficiency Measurement with $t\bar{t}$ Events in Pp Collisions at $\sqrt{s}$ = 13 TeV. *Eur Phys J C* (2019) 79:970.
21. CMS collaboration. A New Calibration Method for Charm Jet Identification Validated with Proton-Proton Collision Events at $\sqrt{s}$ =13 TeV. *arXiv: 2111.03027* (2021).

22. ATLAS collaboration. Measurement of the C-Jet Mistagging Efficiency in $t\bar{t}$ Events Using Pp Collision Data at $\sqrt{s}$ = 13 TeV Collected with the ATLAS Detector. *Eur Phys J C* (2022) 82:95.
23. Erdmann J, Nackenhorst O, Zeißner SV. Maximum Performance of Strange-Jet Tagging at Hadron Colliders. *J Inst* (2021) 16:P08039. doi:10.1088/1748-0221/16/08/p08039
24. Nakai Y, Shih D, Thomas S. Strange Jet Tagging. *arXiv:2003.09517* (2003).
25. Erdmann J. A Tagger for Strange Jets Based on Tracking Information Using Long Short-Term Memory. *J Inst* (2020) 15:P01021. doi:10.1088/1748-0221/15/01/p01021
26. ATLAS collaboration. Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC. *Phys Lett B* (2012) 716:1.
27. CMS collaboration. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys Lett B* (2012) 716:30.
28. ATLAS collaboration. Observation of $H \to b\bar{b}$ Decays and VH Production with the ATLAS Detector. *Phys Lett B* (2018) 786:59.
29. CMS collaboration. Observation of Higgs Boson Decay to Bottom Quarks. *Phys Rev Lett* (2018) 121:121801.
30. ATLAS collaboration. Measurements of WH and ZH Production in the $H \to b\bar{b}$ Decay Channel in Pp Collisions at 13 TeV with the ATLAS Detector. *Eur Phys J C* (2021) 81:178.
31. ATLAS collaboration. Measurement of the Associated Production of a Higgs Boson Decaying into B-Quarks with a Vector Boson at High Transverse Momentum in Pp Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector. *Phys Lett B* (2021) 816:136204.
32. ATLAS collaboration. Search for the Standard Model Higgs Boson Produced by Vector-Boson Fusion and Decaying to Bottom Quarks in $\sqrt{s}$ = 8 TeV Pp Collisions with the ATLAS Detector. *JHEP* (2016) 11:112.
33. CMS collaboration. Search for the Standard Model Higgs Boson Produced through Vector Boson Fusion and Decaying to $b\bar{b}$. *Phys Rev D* (2015) 92:032008.
34. ATLAS collaboration. Measurements of Higgs Bosons Decaying to Bottom Quarks from Vector Boson Fusion Production with the ATLAS experiment at $\sqrt{s}$ = 13 TeV. *Eur Phys J C* (2021) 81:537.
35. Gouzevitch M, Carvalho A. A Review of Higgs Boson Pair Production. *Rev Phys* (2020) 5:100039. doi:10.1016/j.revip.2020.100039
36. Particle Data Group collaboration. Review of Particle Physics. *PTEP* (2020) 2020:083C01.
37. Rappoccio S. The Experimental Status of Direct Searches for Exotic Physics beyond the Standard Model at the Large Hadron Collider. *Rev Phys* (2019) 4:100027. doi:10.1016/j.revip.2018.100027
38. von Buddenbrock S, Chakrabarty N, Cornell AS, Kar D, Kumar M, Mandal T, et al. Phenomenological Signatures of Additional Scalar Bosons at the LHC. *Eur Phys J C* (2016) 76:580. doi:10.1140/epjc/s10052-016-4435-8
39. Dawson S, Sullivan M. Enhanced Di-higgs Boson Production in the Complex Higgs Singlet Model. *Phys Rev D* (2018) 97:015022. doi:10.1103/physrevd.97.015022
40. Robens T, Stefaniak T, Wittbrodt J. Two-real-scalar-singlet Extension of the SM: LHC Phenomenology and Benchmark Scenarios. *Eur Phys J C* (2020) 80:151. doi:10.1140/epjc/s10052-020-7655-x
41. Dawson S, Sullivan M. Enhanced Di-higgs Boson Production in the Complex Higgs Singlet Model. *Phys Rev D* (2018) 97:015022. doi:10.1103/physrevd.97.015022
42. Muehlleitner M, Sampaio MO, Santos R, Wittbrodt J. Phenomenological Comparison of Models with Extended Higgs Sectors. *JHEP* (2017) 08:132.
43. Cacciari M, Salam GP, Soyez G. The Anti-ktjet Clustering Algorithm. *J High Energ Phys.* (2008) 2008:063. doi:10.1088/1126-6708/2008/04/063
44. Krohn D, Thaler J, Wang LT. Jet Trimming. *JHEP* (2010) 02:084. doi:10.1007/jhep02(2010)084
45. Ellis SD, Vermilion CK, Walsh JR. Techniques for Improved Heavy Particle Searches with Jet Substructure. *Phys Rev* (2009) D80:051501. doi:10.1103/physrevd.80.051501
46. Butterworth JM, Davison AR, Rubin M, Salam GP. Jet Substructure as a New Higgs-Search Channel at the Large Hadron Collider. *Phys Rev Lett* (2008) 100:242001. doi:10.1103/physrevlett.100.242001
47. ATLAS collaboration. Performance of Jet Substructure Techniques for Large-R Jets in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV Using the ATLAS Detector. *JHEP* (2013) 09:076.
48. ATLAS collaboration. Measurement of W Boson Angular Distributions in Events with High Transverse Momentum Jets at $\sqrt{s}$ = 8 TeV Using the ATLAS Detector. *Phys Lett B* (2017) 765:132.
49. Eboli OJP, Gonzalez-Garcia MC, Lietti SM. Bosonic Quartic Couplings at CERN LHC. *Phys Rev* (2004) D69:095005. doi:10.1103/physrevd.69.095005
50. Eboli OJP, Gonzalez-Garcia MC, Mizukoshi JK. p p — > j j e+- mu+- nu nu and j j e+- mu-+ nu nu at O(alpha(em)**6) and O(alpha(em)**4 alpha(s)**2) for the study of the quartic electroweak gauge boson vertex at CERN LHC. *Phys Rev* (2006) D74:073005.
51. Chang J, Cheung K, Lu CT, Yuan TC. WW Scattering in the Era of post-Higgs-boson Discovery. *Phys Rev* (2013) D87:093005.
52. Espriu D, Yencho B. Longitudinal WW Scattering in Light of the "Higgs Boson" Discovery. *Phys Rev* (2013) D87:055017.
53. CMS collaboration. Measurement of the Integrated and Differential $t\bar{t}$ Production Cross Sections for High-P_t Top Quarks in Pp Collisions at $\sqrt{s}$ = 8 TeV. *Phys Rev D* (2016) 94:072002.
54. CMS collaboration. Measurement of the Jet Mass in Highly Boosted $t\bar{t}$ Events from Pp Collisions at $\sqrt{s}$ = 8 TeV. *Eur Phys J C* (2017) 77:467.
55. CMS collaboration. Measurement of the Jet Mass Distribution and Top Quark Mass in Hadronic Decays of Boosted Top Quarks in Pp Collisions at $\sqrt{s}$ = 13 TeV. *Phys Rev Lett* (2020) 124:202001.
56. CMS collaboration. Measurement of Differential $t\bar{t}$ Production Cross Sections Using Top Quarks at Large Transverse Momenta in Pp Collisions at $\sqrt{s}$ = 13 TeV. *Phys Rev D* (2021) 103:052008.
57. CMS collaboration. Measurement of Differential $t\bar{t}$ Production Cross Sections in the Full Kinematic Range Using Lepton+jets Events from Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV. *Phys Rev D* (2021) 104:092013.
58. ATLAS collaboration. Measurements of Differential Cross-Sections in Top-Quark Pair Events with a High Transverse Momentum Top Quark and Limits on beyond the Standard Model Contributions to Top-Quark Pair Production with the ATLAS Detector at $\sqrt{s}$ = 13 TeV. *arXiv:2202.12134* (2022).
59. Collaboration C. Identification of Heavy, Energetic, Hadronically Decaying Particles Using Machine-Learning Techniques. *J Instrumentation* (2020) 15:P06005.
60. Collamati F, Curatolo C, Lucchesi D, Mereghetti A, Mokhov N, Palmer M, et al. Advanced Assessment of Beam-Induced Background at a Muon Collider. *J Inst* (2021) 16:P11009. doi:10.1088/1748-0221/16/11/p11009
61. Buckley A, Butterworth J, Gieseke S, Grellscheid D, Höche S, Hoeth H, et al. General-purpose Event Generators for LHC Physics. *Phys Rep* (2011) 504:145–233. doi:10.1016/j.physrep.2011.03.005
62. Berger CF, Bern Z, Dixon LJ, Febres Cordero F, Forde D, Ita H, et al. An Automated Implementation of On-Shell Methods for One-Loop Amplitudes. *Phys Rev D* (2008) 78:036003. doi:10.1103/physrevd.78.036003
63. Bevilacqua G, Czakon M, Garzelli MV, van Hameren A, Kardos A, Papadopoulos CG, et al. HELAC-NLO. *Comput Phys Commun* (2013) 184:986–97. doi:10.1016/j.cpc.2012.10.033
64. Cascioli F, Maierhöfer P, Pozzorini S. Scattering Amplitudes with Open Loops. *Phys Rev Lett* (2012) 108:111601. doi:10.1103/physrevlett.108.111601
65. Hirschi V, Frederix R, Frixione S, Garzelli MV, Maltoni F, Pittau R. Automation of One-Loop QCD Corrections. *JHEP* (2011) 05:044.
66. Cullen G, van Deurzen H, Greiner N, Heinrich G, Luisoni G, Mastrolia P, et al. GOSAM-2.0: a Tool for Automated One-Loop Calculations within the Standard Model and beyond. *Eur Phys J C* (2014) 74:3001. doi:10.1140/epjc/s10052-014-3001-5
67. Actis S, Denner A, Hofer L, Lang JN, Scharf A, Uccirati S. R E C O L A-REcursive Computation of One-Loop Amplitudes. *Comput Phys Commun* (2017) 214:140–73. doi:10.1016/j.cpc.2017.01.004
68. ATLAS collaboration. Measurement of the Soft-Drop Jet Mass in Pp Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector. *Phys Rev Lett* (2018) 121:092001.
69. CMS collaboration. Measurements of the Differential Jet Cross Section as a Function of the Jet Mass in Dijet Events from Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV. *JHEP* (2018) 11:113.
70. Andersson B, Gustafson G, Söderberg B. A General Model for Jet Fragmentation. *Z Phys C - Particles Fields* (1983) 20:317–29. doi:10.1007/bf01407824

71. Andersson B, Gustafson G, Ingelman G, Sjöstrand T. Parton Fragmentation and String Dynamics. *Phys Rep* (1983) 97:31–145. doi:10.1016/0370-1573(83)90080-7
72. Gottschalk TD. A realistic model for e+e− annihilation including parton bremsstrahlung effects. *Nucl Phys B* (1983) 214:201–22. doi:10.1016/0550-3213(83)90658-2
73. Gottschalk TD. An improved description of hadronization in the QCD cluster model for e+e− annihilation. *Nucl Phys B* (1984) 239:349–81. doi:10.1016/0550-3213(84)90253-0
74. Webber BR. A QCD Model for Jet Fragmentation Including Soft Gluon Interference. *Nucl Phys B* (1984) 238:492–528. doi:10.1016/0550-3213(84)90333-x
75. Bellm J, Gieseke S, Grellscheid D, Plätzer S, Rauch M, Reuschle C, et al. Herwig 7.0/Herwig++ 3.0 Release Note. *Eur Phys J C* (2016) 76:196. doi:10.1140/epjc/s10052-016-4018-8
76. Sjöstrand T, Ask S, Christiansen JR, Corke R, Desai N, Ilten P, et al. An Introduction to PYTHIA 8.2. *Comput Phys Commun* (2015) 191:159–77. doi:10.1016/j.cpc.2015.01.024
77. Sherpa Collaboration. Event Generation with Sherpa 2.2. *Scipost Phys* (2019) 7:034.
78. Kanaki A, Papadopoulos CG. HELAC: A Package to Compute Electroweak Helicity Amplitudes. *Comput Phys Commun* (2000) 132:306–15. doi:10.1016/s0010-4655(00)00151-x
79. Krauss F, Kuhn R, Soff G. AMEGIC++ 1.0, A Matrix Element Generator in C++. *J High Energ Phys.* (2002) 2002:044. doi:10.1088/1126-6708/2002/02/044
80. Mangano ML, Piccinini F, Polosa AD, Pittau MR, Polosa AD. ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions. *J High Energ Phys.* (2003) 2003:001. doi:10.1088/1126-6708/2003/07/001
81. Gleisberg T, Höche S. Comix, a New Matrix Element Generator. *J High Energ Phys.* (2008) 2008:039. doi:10.1088/1126-6708/2008/12/039
82. Alwall J, Frederix R, Frixione S, Hirschi V, Maltoni F, Mattelaer O, et al. The Automated Computation of Tree-Level and Next-To-Leading Order Differential Cross Sections, and Their Matching to Parton Shower Simulations. *JHEP* (2014) 07:079. doi:10.1007/jhep07(2014)079
83. Andersen JR, Bendavid J, Ciulli V, Denner A, Frederix R, Grazzini M, et al. Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report. In: *9th Les Houches Workshop on Physics at TeV Colliders* (2016). p. 5. arXiv:1605.04692.
84. Bellm J, Buckley A, Chen X, Ridder AGD, Gehrmann T, Glover N, et al. Jet Cross Sections at the LHC and the Quest for Higher Precision. *Eur Phys J C* (2020) 80:93. doi:10.1140/epjc/s10052-019-7574-x
85. Buckley A, Chen X, Cruz-Martinez J, Ferrario Ravasio S, Gehrmann T, Glover EWN, et al. A Comparative Study of Higgs Boson Production from Vector-Boson Fusion. *J High Energ Phys* (2021) 2021:108. doi:10.1007/jhep11(2021)108
86. Höche S, Reichelt D, Siegert F. Momentum Conservation and Unitarity in Parton Showers and NLL Resummation. *J High Energ Phys* (2018) 2018:118. doi:10.1007/jhep01(2018)118
87. Dasgupta M, Dreyer FA, Hamilton K, Monni PF, Salam GP. Logarithmic Accuracy of Parton Showers: a Fixed-Order Study. *JHEP* (2018) 09:033. doi:10.1007/jhep09(2018)033
88. Nagy Z, Soper DE. Summations of Large Logarithms by Parton Showers. *Phys Rev D* (2021) 104:054049. doi:10.1103/physrevd.104.054049
89. Nagy Z, Soper DE. Parton Showers with Quantum Interference. *J High Energ Phys.* (2007) 2007:114. doi:10.1088/1126-6708/2007/09/114
90. Bewick G, Ferrario Ravasio S, Richardson P, Seymour MH. Logarithmic Accuracy of Angular-Ordered Parton Showers. *JHEP* (2020) 04:019. doi:10.1007/jhep04(2020)019
91. Dasgupta M, Dreyer FA, Hamilton K, Monni PF, Salam GP, Soyez G. Parton Showers beyond Leading Logarithmic Accuracy. *Phys Rev Lett* (2020) 125:052002. doi:10.1103/PhysRevLett.125.052002
92. Knowles IG. A Linear Algorithm for Calculating Spin Correlations in Hadronic Collisions. *Comput Phys Commun* (1990) 58:271–84. doi:10.1016/0010-4655(90)90063-7
93. Nagy Z, Soper DE. Parton Showers with Quantum Interference: Leading Color, with Spin. *J High Energ Phys.* (2008) 2008:025. doi:10.1088/1126-6708/2008/07/025
94. Richardson P, Webster S. Spin Correlations in Parton Shower Simulations. *Eur Phys J C* (2020) 80:83. doi:10.1140/epjc/s10052-019-7429-5
95. Karlberg A, Salam GP, Scyboz L, Verheyen R. Spin Correlations in Final-State Parton Showers and Jet Observables. *Eur Phys J C* (2021) 81:681. doi:10.1140/epjc/s10052-021-09378-0
96. Hamilton K, Karlberg A, Salam GP, Scyboz L, Verheyen R. Soft Spin Correlations in Final-State Parton Showers. *arXiv:2111.01161* (2021).
97. Nagy Z, Soper DE. Parton Showers with More Exact Color Evolution. *Phys Rev D* (2019) 99:054009. doi:10.1103/physrevd.99.054009
98. Hamilton K, Medves R, Salam GP, Scyboz L, Soyez G. Colour and Logarithmic Accuracy in Final-State Parton Showers. *arXiv:2011.10054* (2020).
99. Hartgring L, Laenen E, Skands P. Antenna Showers with One-Loop Matrix Elements. *J High Energ Phys* (2013) 2013:127. doi:10.1007/jhep10(2013)127
100. Li HT, Skands P. A Framework for Second-Order Parton Showers. *Phys Lett B* (2017) 771:59–66. doi:10.1016/j.physletb.2017.05.011
101. Höche S, Prestel S. Triple Collinear Emissions in Parton Showers. *Phys Rev D* (2017) 96:074017. doi:10.1103/physrevd.96.074017
102. Dulat F, Höche S, Prestel S. Leading-Color Fully Differential Two-Loop Soft Corrections to QCD Dipole Showers. *Phys Rev D* (2018) 98:074013. doi:10.1103/physrevd.98.074013
103. Gellersen L, Höche S, Prestel S. Disentangling Soft and Collinear Effects in QCD Parton Showers. *arXiv:2110.05964* (2021).
104. Frixione S, Webber BR. Matching NLO QCD Computations and Parton Shower Simulations. *J High Energ Phys.* (2002) 2002:029. doi:10.1088/1126-6708/2002/06/029
105. Nason P. A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms. *J High Energ Phys.* (2004) 2004:040. doi:10.1088/1126-6708/2004/11/040
106. Catani S, Krauss F, Kuhn BRR, Webber BR. QCD Matrix Elements + Parton Showers. *J High Energ Phys.* (2001) 2001:063. doi:10.1088/1126-6708/2001/11/063
107. Mangano ML, Moretti M, Pittau R. Multijet matrix elements and shower evolution in hadronic collisions: -jets as a case study $Wb\bar{b}$ + n jets as a case study. *Nucl Phys B* (2002) 632:343–62. doi:10.1016/s0550-3213(02)00249-3
108. Lönnblad L. Correcting the Colour-Dipole Cascade Model with Fixed Order Matrix Elements. *J High Energ Phys.* (2002) 2002:046. doi:10.1088/1126-6708/2002/05/046
109. Hoeche S, Krauss F, Schonherr M, Siegert F. QCD Matrix Elements + Parton Showers: The NLO Case. *JHEP* (2013) 04:027.
110. Frederix R, Frixione S. Merging Meets Matching in MC@NLO. *JHEP* (2012) 12:061. doi:10.1007/jhep12(2012)061
111. Lönnblad L, Prestel S. Merging Multi-Leg NLO Matrix Elements with Parton Showers. *J High Energ Phys* (2013) 2013:166. doi:10.1007/jhep03(2013)166
112. Bellm J, Gieseke S, Plätzer S. Merging NLO Multi-Jet Calculations with Improved Unitarization. *Eur Phys J C* (2018) 78:244. doi:10.1140/epjc/s10052-018-5723-2
113. Hamilton K, Nason P, Oleari C, Zanderighi G. Merging H/W/Z + 0 and 1 Jet at NLO with No Merging Scale: a Path to Parton Shower + NNLO Matching. *JHEP* (2013) 05:082.
114. Alioli S, Bauer CW, Berggren C, Tackmann FJ, Walsh JR, Zuberi S. Matching Fully Differential NNLO Calculations and Parton Showers. *JHEP* (2014) 06:089. doi:10.1007/jhep06(2014)089
115. Monni PF, Nason P, Re E, Wiesemann M, Zanderighi G. MiNNLOPS: a New Method to Match NNLO QCD to Parton Showers. *J High Energ Phys* (2020) 2020:143. doi:10.1007/jhep05(2020)143
116. Höche S, Li Y, Prestel S. Drell-Yan Lepton Pair Production at NNLO QCD with Parton Showers. *Phys Rev D* (2015) 91:074015. doi:10.1103/physrevd.91.074015
117. Prestel S. Matching N3LO QCD Calculations to Parton Showers. *JHEP* (2021) 11:041. doi:10.1007/jhep11(2021)041
118. Bertone V, Prestel S. Combining N3LO QCD Calculations and Parton Showers for Hadronic Collision Events. *arXiv:2202.01082* (2022).
119. Sjöstrand T, van Zijl M. A Multiple-Interaction Model for the Event Structure in Hadron Collisions. *Phys Rev D* (1987) 36:2019–41. doi:10.1103/physrevd.36.2019

120. Butterworth JM, Forshaw JR, Seymour MH. Multiparton Interactions in Photoproduction at HERA. *Z Phys C - Particles Fields* (1996) 72:637–46. doi:10.1007/s002880050286
121. Bengtsson HU, Sjöstrand T. The Lund Monte Carlo for Hadronic Processes - PYTHIA Version 4.8. *Comput Phys Commun* (1987) 46:43–82. doi:10.1016/0010-4655(87)90036-1
122. Andersson B. *The Lund Model, Vol. 7*. Cambridge: Cambridge University Press (2005). p. 7. doi:10.1017/CBO9780511524363
123. Winter JC, Krauss F, Soff G. A Modified Cluster-Hadronisation Model. *Eur Phys J C* (2004) 36:381–95. doi:10.1140/epjc/s2004-01960-8
124. Ghosh A, Nachman B. A Cautionary Tale of Decorrelating Theory Uncertainties. *Eur Phys J C* (2022) 82:46. doi:10.1140/epjc/s10052-022-10012-w
125. Hoeche S, Krauss F, Schonherr M, Siegert F. A Critical Appraisal of NLO+PS Matching Methods. *JHEP* (2012) 09:049.
126. SMNLO MULTILEG Working GroupSM MC Working Group collaboration. The SM and NLO Multileg and SM MC Working Groups: Summary Report. In: *7th Les Houches Workshop on Physics at TeV Colliders*, 3 (2012). p. 1–220. arXiv:1203.6803.
127. Yeh CH, Chekanov SV, Kotwal AV, Proudfoot J, Sen S, Tran NV, et al. Studies of Granularity of a Hadronic Calorimeter for Tens-Of-TeV Jets at a 100 TeV Pp Collider. *J Inst* (2019) 14:P05008. doi:10.1088/1748-0221/14/05/p05008
128. Coleman E, Freytsis M, Hinzmann A, Narain M, Thaler J, Tran N, et al. The Importance of Calorimetry for Highly-Boosted Jet Substructure. *J Inst* (2018) 13:T01003. doi:10.1088/1748-0221/13/01/t01003
129. CMS collaboration. Particle-flow Reconstruction and Global Event Description with the CMS Detector. *JINST* (2017) 12:P10003.
130. ILD collaboration. The ILD Detector at the ILC. *arXiv:1912.04601* (2019).
131. Breidenbach M, Brau JE, Burrows P, Markiewicz T, Stanitzki M, Strube J, et al. Updating the SiD Detector Concept. *arXiv:2110.09965* (2021).
132. CLICdp collaboration. A Detector for CLIC: Main Parameters and Performance. *arXiv:1812.07337* (2021).
133. Bacchetta N, Blaising JJ, Brondolin E, Dam M, Dannheim D, Elsener K, et al. CLD – A Detector Concept for the FCC-Ee. *arXiv:1911.12230* (2019).
134. CALICE collaboration. Tests of a Particle Flow Algorithm with CALICE Test Beam Data. *JINST* (2011) 6:P07005.
135. ALICE Collaboration. Letter of Intent: A Forward Calorimeter (FoCal) in the ALICE experiment. In: *CERN-LHCC-2020-009, LHCC-I-036* (2020).
136. ATLAS Collaboration Collaboration. ATLAS Phase-II Upgrade Scoping Document. In: *Tech. Rep*. Geneva: CERN (2015).
137. Contardo D, Klute M, Mans J, Silvestris L, Butler J. Technical Proposal for the Phase-II Upgrade of the CMS Detector. In: *Tech. Rep*. Geneva: CERN (2015).
138. Ström R, Roloff P. Physics Potential for Boosted Topologies in Top-Quark Pair Production at a Multi-TeV Compact Linear Collider. *arXiv:2008.05526* (2020).
139. Thaler J, Van Tilburg K. Identifying Boosted Objects with N-Subjettiness. *JHEP* (2011) 03:015. doi:10.1007/jhep03(2011)015
140. Thaler J, Van Tilburg K. Maximizing Boosted Top Identification by Minimizing N-Subjettiness. *JHEP* (2012) 02:093. doi:10.1007/jhep02(2012)093
141. Larkoski AJ, Moult I, Neill D. Power Counting to Better Jet Observables. *JHEP* (2014) 12:009. doi:10.1007/jhep12(2014)009
142. ATLAS Collaboration collaboration. Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade. In: *Tech. Tep*. Geneva: CERN (2020).
143. CMS collaboration. *A MIP Timing Detector for the CMS Phase-2 Upgrade* (2019).
144. Coleman E, Freytsis M, Hinzmann A, Narain M, Thaler J, Tran N, et al. The Importance of Calorimetry for Highly-Boosted Jet Substructure. *J Inst* (2018) 13:T01003. doi:10.1088/1748-0221/13/01/t01003
145. CMS collaboration. The Phase-2 Upgrade of the CMS Endcap Calorimeter. In: *Tech. Rep*. Geneva: CERN (2017).
146. FCC collaboration. FCC-hh: The Hadron Collider: Future Circular Collider Conceptual Design Report Volume 3. *Eur Phys J ST* (2019) 228:755.
147. Chekanov SV. Performance Requirements for Hadron Calorimeters. In: First Annual Meeting of the Future Circular Collider study; 23-29 March 2015; Washington, DC (2015).
148. Chekanov SV, Beydler M, Kotwal AV, Gray L, Sen S, Tran NV, et al. Initial Performance Studies of a General-Purpose Detector for Multi-TeV Physics at a 100 TeVppcollider. *J Inst* (2017) 12:P06009. doi:10.1088/1748-0221/12/06/p06009
149. CALICE collaboration. Hadronic Energy Resolution of a Highly Granular Scintillator-Steel Hadron Calorimeter Using Software Compensation Techniques. *JINST* (2012) 7:P09017.
150. Abusleme Hoffman AC, Parès G, Fritzsch T, Rothermund M, Jansen H, Krüger K, et al. Detector Technologies for CLIC. *arXiv:1905.02520* (2019).
151. Aleksa M, Allport P, Bosley R, Faltova J, Gentil J, Goncalo R, et al. Calorimeters for the FCC-Hh. *arXiv:1912.09962* (2019).
152. GEANT4 collaboration. GEANT4–a Simulation Toolkit. *Nucl Instrum Meth A* (2003) 506:250.
153. CALICE collaboration. Characterisation of Different Stages of Hadronic Showers Using the CALICE Si-W ECAL Physics Prototype. *Nucl Instrum Meth A* (2019) 937:41.
154. Thomson MA. Model-independent measurement of the e $${+}$$ + e $${-}$$ - $$\rightarrow $$ HZ cross section at a future e $${+}$$ + e $${-}$$ - linear collider using hadronic Z decays. *Eur Phys J C* (2016) 76:72. doi:10.1140/epjc/s10052-016-3911-5
155. Lai PZ, Ruan M, Kuo CM. Jet Performance at the Circular Electron-Positron Collider. *J Inst* (2021) 16:P07037. doi:10.1088/1748-0221/16/07/p07037
156. Boronat M, Fuster J, Garcia I, Roloff P, Simoniello R, Vos M. Jet Reconstruction at High-Energy Electron-Positron Colliders. *Eur Phys J C* (2018) 78:144. doi:10.1140/epjc/s10052-018-5594-6
157. Boronat M, Fuster J, García I, Ros E, Vos M. A Robust Jet Reconstruction Algorithm for High-Energy Lepton Colliders. *Phys Lett B* (2015) 750:95–9. doi:10.1016/j.physletb.2015.08.055
158. Stewart IW, Tackmann FJ, Thaler J, Vermilion CK, Wilkason TF. XCone: N-Jettiness as an Exclusive Cone Jet Algorithm. *JHEP* (2015) 11:072. doi:10.1007/jhep11(2015)072
159. Cacciari M, Salam GP. Dispelling the N^3 Myth for the K_t Jet-Finder. *Phys Lett B* (2006) 641:57.
160. Cacciari M, Salam GP, Soyez G. FastJet User Manual. *Eur Phys J C* (2012) 72:1896. doi:10.1140/epjc/s10052-012-1896-2
161. Cacciari M, Salam GP, Soyez G. The Catchment Area of Jets. *J High Energ Phys.* (2008) 2008:005. doi:10.1088/1126-6708/2008/04/005
162. CMS collaboration. Pileup Removal Algorithms. In: *CMS Physics Analysis Summary CMS-PAS-JME-14-001* (2014).
163. ATLAS Collaboration. Performance of Pile-Up Mitigation Techniques for Jets in Pp Collisions at $\sqrt{s}$ = 8 Tev Using the Atlas Detector. *Eur Phys J C* (2016) 76:581.
164. ATLAS Collaboration. Jet Energy Scale Measurements and Their Systematic Uncertainties in Proton-Proton Collisions at $\sqrt{s}$ = 13 Tev with the Atlas Detector. *Phys Rev D* (2017) 96:072002.
165. CMS Collaboration. Jet Energy Scale and Resolution in the Cms experiment in Pp Collisions at 8 Tev. *JINST* (2017) 12:P02014.
166. Dasgupta M, Magnea L, Salam GP. Non-perturbative QCD Effects in Jets at Hadron Colliders. *J High Energ Phys.* (2008) 2008:055. doi:10.1088/1126-6708/2008/02/055
167. Krohn D, Thaler J, Wang LT. Jets with Variable R. *J High Energ Phys.* (2009) 2009:059. doi:10.1088/1126-6708/2009/06/059
168. Lapsien T, Kogler R, Haller J. A New Tagger for Hadronically Decaying Heavy Particles at the LHC. *Eur Phys J C* (2016) 76:600. doi:10.1140/epjc/s10052-016-4443-8
169. ATLAS collaboration. Boosted Object Tagging with Variable-R Jets in the ATLAS Detector. In: *ATLAS Physics Note ATL-PHYS-PUB-2016-013* (2016).
170. ATLAS collaboration. Variable Radius, Exclusive-k_T, and center-of-mass Subjet Reconstruction for Higgs ($\rightarrow b\bar{b}$) Tagging in ATLAS. In: *ATLAS Physics Note ATL-PHYS-PUB-2017-010* (2017).
171. CMS collaboration. Identification of Heavy, Energetic, Hadronically Decaying Particles Using Machine-Learning Techniques. *JINST* (2020) 15:P06005.
172. ATLAS collaboration. Jet Reconstruction and Performance Using Particle Flow with the ATLAS Detector. *Eur Phys J C* (2017) 77:466.

173. ATLAS collaboration. *Improving Jet Substructure Performance in ATLAS Using Track-CaloClusters* (2017).
174. CMS collaboration. Pileup Mitigation at CMS in 13 TeV Data. *JINST* (2020) 15:P09018.
175. CMS collaboration. V Tagging Observables and Correlations. In: *Physics Analysis Summary CMS-PAS-JME-14-002* (2014).
176. CLICdp collaboration. Top-Quark Physics at the CLIC Electron-Positron Linear Collider. *JHEP* (2019) 11:003.
177. Fujii K, Grojean C, Peskin ME, Barklow T, Gao Y, Kanemura S, et al. ILC Study Questions for Snowmass 2021. *arXiv:2007.03650* (2021).
178. Delahaye JP, Diemoz M, Long K, Mansoulié B, Pastrone N, Rivkin L, et al. Muon Colliders. *arXiv:1901.06150* (2019).
179. ALEGRO collaboration. Towards an Advanced Linear International Collider. *arXiv:1901.10370* (2019).
180. Bai M, Barklow T, Bartoldus R, Breidenbach M, Grenier P, Huang Z, et al. C^3: A "Cool" Route to the Higgs Boson and beyond. *arXiv:2110.15800* (2021) 10.
181. Bartosik N, Andreetto P, Buonincontri L, Casarsa M, Gianelle A, Griso SP, et al. Full Detector Simulation with Unprecedented Background Occupancy at a Muon Collider. *Comput Softw Big Sci* (2021) 5:21. doi:10.1007/s41781-021-00067-x
182. ATLAS collaboration. Topological Cell Clustering in the ATLAS Calorimeters and its Performance in LHC Run 1. *Eur Phys J C* (2017) 77:490.
183. Berta P, Spousta M, Miller DW, Leitner R. Particle-level Pileup Subtraction for Jets and Jet Shapes. *JHEP* (2014) 06:092. doi:10.1007/jhep06(2014)092
184. ATLAS collaboration. Optimisation of Large-Radius Jet Reconstruction for the ATLAS Detector in 13 TeV Proton–Proton Collisions. *Eur Phys J C* (2021) 81:334.
185. Bertolini D, Harris P, Low M, Tran N. Pileup Per Particle Identification. *JHEP* (2014) 10:059. doi:10.1007/jhep10(2014)059
186. Cacciari M, Salam GP, Soyez G. SoftKiller, a Particle-Level Pileup Removal Method. *Eur Phys J C* (2015) 75:59. doi:10.1140/epjc/s10052-015-3267-2
187. ATLAS Collaboration. Constituent-level Pile-Up Mitigation Techniques in ATLAS. In: *ATLAS-CONF-2017-065* (2017).
188. CMS. Pileup-per-particle Identification: Optimisation for Run 2 Legacy and beyond. In: *Detector Performance Note CMS-DP-2021-001* (2021).
189. ATLAS Collaboration. A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade: Technical Design Report. In: *atlas-tdr-031; Cern-Lhcc-2020-007* (2020).
190. CMS collaboration. *The Phase-2 Upgrade of the CMS Endcap Calorimeter* (2017).
191. Komiske PT, Metodiev EM, Nachman B, Schwartz MD. Pileup Mitigation with Machine Learning (PUMML). *JHEP* (2017) 12:051. doi:10.1007/jhep12(2017)051
192. Arjona Martínez J, Cerri O, Spiropulu M, Pierini JRM, Vlimant J-R. Pileup Mitigation at the Large Hadron Collider with Graph Neural Networks. *Eur Phys J Plus* (2019) 134:333. doi:10.1140/epjp/i2019-12710-3
193. Mikuni V, Canelli F. ABCNet: An Attention-Based Method for Particle Tagging. *Eur Phys J Plus* (2020) 135:463. doi:10.1140/epjp/s13360-020-00497-3
194. Maier B, Narayanan SM, de Castro G, Goncharov M, Paus C, Schott M. Pile-Up Mitigation Using Attention. *arXiv:2107.02779* (2021).
195. Schwaller P, Stolarski D, Weiler A. *Emerging Jets, JHEP* (2015) 05:059.
196. CMS collaboration. Search for New Particles Decaying to a Jet and an Emerging Jet. *JHEP* (2019) 02:179.
197. Chiu WH, Liu Z, Low M, Wang LT. Jet Timing. *JHEP* (2022) 01:014. doi:10.1007/jhep01(2022)014
198. Chatterjee S, Godbole R, Roy TS. Jets with Electrons from Boosted Top Quarks. *J High Energ Phys* (2020) 2020:170. doi:10.1007/jhep01(2020)170
199. Mitra M, Ruiz R, Scott DJ, Spannowsky M. Neutrino Jets from High-Mass W_R Gauge Bosons in TeV-Scale Left-Right Symmetric Models. *Phys Rev D* (2016) 94:095016. doi:10.1103/physrevd.94.095016
200. Nemevšek M, Nesti F, Popara G. Keung-Senjanović Process at the LHC: From Lepton Number Violation to Displaced Vertices to Invisible Decays. *Phys Rev D* (2018) 97:115018.
201. du Plessis K, Flores MM, Kar D, Sinha S, van der Schyf H. Hitting Two BSM Particles with One Lepton-Jet: Search for a Top Partner Decaying to a Dark Photon, Resulting in a Lepton-Jet. *arXiv:2112.08425* (2021).
202. Dube S, Gadkari D, Thalapillil AM. Lepton-Jets and Low-Mass Sterile Neutrinos at Hadron Colliders. *Phys Rev D* (2017) 96:055031. doi:10.1103/physrevd.96.055031
203. Wang D, Wu L, Yang JM, Zhang M. Photon-jet Events as a Probe of Axionlike Particles at the LHC. *Phys Rev D* (2021) 104:095016. doi:10.1103/physrevd.104.095016
204. Sheff B, Steinberg N, Wells JD. Higgs Boson Decays into Narrow Diphoton Jets and Their Search Strategies at the Large Hadron Collider. *Phys Rev D* (2021) 104:036009. doi:10.1103/physrevd.104.036009
205. Cohen T, Lisanti M, Lou HK. Semivisible Jets: Dark Matter Undercover at the LHC. *Phys Rev Lett* (2015) 115:171804. doi:10.1103/physrevlett.115.171804
206. Cohen T, Lisanti M, Lou HK, Mishra-Sharma S. LHC Searches for Dark Sector Showers. *J High Energ Phys* (2017) 2017:196. doi:10.1007/jhep11(2017)196
207. Kar D, Sinha S. Exploring Jet Substructure in Semi-visible Jets. *Scipost Phys* (2021) 10:084. doi:10.21468/scipostphys.10.4.084
208. Canelli F, de Cosa A, Pottier LL, Niedziela J, Pedro K, Pierini M. Autoencoders for Semivisible Jet Detection. *arXiv:2112.02864* (2021).
209. CMS collaboration. Search for a Right-Handed W Boson and a Heavy Neutrino in Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV. *arXiv:1803.11116* (2018).
210. ATLAS collaboration. Search for a Right-Handed Gauge Boson Decaying into a High-Momentum Heavy Neutrino and a Charged Lepton in Pp Collisions with the ATLAS Detector at $\sqrt{s}$ = 13 TeV. *Phys Lett B* (2019) 798:134942.
211. CMS collaboration. Search for Resonant Production of Strongly Coupled Dark Matter in Proton-Proton Collisions at 13 TeV. *arXiv:2112.11125* (2021).
212. ATLAS collaboration. Search for Light Long-Lived Neutral Particles Produced in Pp Collisions at $\sqrt{s}$ = 13 TeV and Decaying into Collimated Leptons or Light Hadrons with the ATLAS Detector. *Eur Phys J C* (2020) 80:450.
213. CMS collaboration. Search for Long-Lived Particles Using Nonprompt Jets and Missing Transverse Momentum with Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV. *Phys Lett B* (2019) 797:134876.
214. Liu J, Liu Z, Wang LT. Enhancing Long-Lived Particles Searches at the LHC with Precision Timing Information. *Phys Rev Lett* (2019) 122:131801. doi:10.1103/physrevlett.122.131801
215. Allaire C, Benitez J, Bomben M, Calderini G, Carulla M, Cavallaro E, et al. Beam Test Measurements of Low Gain Avalanche Detector Single Pads and Arrays for the ATLAS High Granularity Timing Detector. *J Inst* (2018) 13:P06017. doi:10.1088/1748-0221/13/06/p06017
216. ATLAS Collaboration collaboration. Technical Design Report for the ATLAS Inner Tracker Pixel Detector. In: *Tech. Rep.* Geneva: CERN (2017).
217. ATLAS Collaboration collaboration. Technical Design Report for the ATLAS Inner Tracker Strip Detector. In: *Tech. Rep.* Geneva: CERN (2017).
218. CMS Collaboration collaboration. The Phase-2 Upgrade of the CMS Tracker. In: *Tech. Rep.* Geneva: CERN (2017).
219. CMS Collaboration collaboration. The Phase-2 Upgrade of the CMS Level-1 Trigger. In: *Tech. Rep.* Geneva: CERN (2020).
220. de Oliveira L, Kagan M, Mackey L, Nachman B, Schwartzman A. Jet-images — Deep Learning Edition. *JHEP* (2016) 07:069.
221. Ren J, Wang D, Wu L, Yang JM, Zhang M. Detecting an Axion-like Particle with Machine Learning at the LHC. *J High Energ Phys* (2021) 2021:138. doi:10.1007/jhep11(2021)138
222. Giudice GF, Rattazzi R. Theories with Gauge-Mediated Supersymmetry Breaking. *Phys Rep* (1999) 322:419–99. doi:10.1016/s0370-1573(99)00042-3
223. Arkani-Hamed N, Dimopoulos S. Supersymmetric Unification without Low Energy Supersymmetry and Signatures for fine-tuning at the LHC. *J High Energ Phys.* (2005) 2005:073. doi:10.1088/1126-6708/2005/06/073
224. Giudice GF, Romanino A. Split Supersymmetry. *Nucl Phys B* (2004) 699:65–89. doi:10.1016/j.nuclphysb.2004.08.001
225. Fan J, Reece M, Ruderman JT. Stealth Supersymmetry. *JHEP* (2011) 11:012. doi:10.1007/jhep11(2011)012
226. Strassler MJ, Zurek KM. Echoes of a Hidden valley at Hadron Colliders. *Phys Lett B* (2007) 651:374–9. doi:10.1016/j.physletb.2007.06.055
227. Meade P, Reece M, Shih D. Long-Lived Neutralino NLSPs. *JHEP* (2010) 10:067. doi:10.1007/jhep10(2010)067

228. Craig N, Katz A, Strassler M, Sundrum R. Naturalness in the Dark at the LHC. *J High Energ Phys* (2015) 2015:105. doi:10.1007/jhep07(2015)105
229. Carloni L, Sjöstrand T. Visible Effects of Invisible Hidden Valley Radiation. *J High Energ Phys* (2010) 2010:105. doi:10.1007/jhep09(2010)105
230. Carloni L, Rathsman J, Sjostrand T. Discerning Secluded Sector Gauge Structures. *JHEP* (2011) 04:091. doi:10.1007/jhep04(2011)091
231. Larkoski AJ, Salam GP, Thaler J. Energy Correlation Functions for Jet Substructure. *J High Energ Phys* (2013) 2013:108. doi:10.1007/jhep06(2013)108
232. Cohen T, Doss J, Freytsis M. Jet Substructure from Dark Sector Showers. *J High Energ Phys* (2020) 2020:118. doi:10.1007/jhep09(2020)118
233. Larkoski AJ, Maltoni F, Selvaggi M. Tracking Down Hyper-Boosted Top Quarks. *arXiv:1503.03347* (2015). doi:10.1007/jhep06(2015)032
234. Chang HM, Procura M, Thaler J, Waalewijn WJ. Calculating Track-Based Observables for the Lhc. *Phys Rev Lett* (2013) 111:102002. doi:10.1103/PhysRevLett.111.102002
235. Elder BT, Thaler J. Aspects of Track-Assisted Mass. *J High Energ Phys* (2019) 2019. doi:10.1007/jhep03(2019)104
236. Spannowsky M, Stoll M. Tracking New Physics at the Lhc and beyond. *Phys Rev D* (2015) 92. doi:10.1103/physrevd.92.054033
237. ATLAS Collaboration collaboration. Jet Mass Reconstruction with the ATLAS Detector in Early Run 2 Data. In: *Tech. Rep.* Geneva: CERN (2016).
238. Gouskos L, Sung A, Incandela J. Search for Stop Scalar Quarks at FCC-Hh. In: *Tech. Rep.* Geneva: CERN (2018).
239. ATLAS Collaboration collaboration. Improving Jet Substructure Performance in ATLAS Using Track-CaloClusters. In: *Tech. Rep.* Geneva: CERN (2017).
240. CMS Collaboration collaboration. *Mass Regression of Highly-Boosted Jets Using Graph Neural Networks* (2021).
241. Thaler J, Van Tilburg K. Identifying Boosted Objects with N-Subjettiness. *J High Energ Phys* (2011) 2011. doi:10.1007/jhep03(2011)015
242. Larkoski AJ, Salam GP, Thaler J. Energy Correlation Functions for Jet Substructure. *J High Energ Phys* (2013) 2013. doi:10.1007/jhep06(2013)108
243. Plehn T, Salam GP, Spannowsky M. Fat Jets for a Light Higgs Boson. *Phys Rev Lett* (2010) 104:111801. doi:10.1103/PhysRevLett.104.111801
244. Plehn T, Spannowsky M, Takeuchi M, Zerwas D. Stop Reconstruction with Tagged Tops. *J High Energ Phys* (2010) 2010. doi:10.1007/jhep10(2010)078
245. Han Z, Son M, Tweedie B. Top-tagging at the Energy Frontier. *Phys Rev D* (2018) 97. doi:10.1103/physrevd.97.036023
246. Mangano ML, Zanderighi G, Saavedra JAA, Alekhin S, Badger S, Bauer CW, et al. Physics at a 100 Tev Pp Collider: Standard Model Processes. *arXiv:1607.01831* (2016).
247. Rehermann K, Tweedie B. Efficient Identification of Boosted Semileptonic Top Quarks at the Lhc. *J High Energ Phys* (2011) 2011. doi:10.1007/jhep03(2011)059
248. Agashe K, Collins JH, Du P, Hong S, Kim D, Mishra RK. Detecting a Boosted Diboson Resonance. *JHEP* (2018) 11:027. doi:10.1007/jhep11(2018)027
249. CMS collaboration. Measurement of Jet Substructure Observables in $\mathrm{t\bar{t}}$ Events from Proton-Proton Collisions at $\sqrt{s}$ = 13TeV. *Phys Rev D* (2018) 98:092014.
250. ATLAS Collaboration. Measurement of Jet-Substructure Observables in Top Quark, W Boson and Light Jet Production in Proton-Proton Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector. *JHEP* (2019) 08:033.
251. CMS collaboration. Development and Validation of HERWIG 7 Tunes from CMS Underlying-Event Measurements. *Eur Phys J C* (2021) 81:312.
252. CMS collaboration. *A New Set of CMS Tunes for Novel Colour Reconnection Models in PYTHIA8 Based on Underlying-Event Data* (2021).
253. Buckley A, Hoeth H, Lacker H, Schulz H, von Seggern JE. Systematic Event Generator Tuning for the LHC. *Eur Phys J C* (2010) 65:331–57. doi:10.1140/epjc/s10052-009-1196-7
254. Šimko T, Lange C, Heinrich LA, Lintuluoto AE, MacDonell DM, Mečionis A, et al. Scalable Declarative Hep Analysis Workflows for Containerised Compute Clouds. *Front Big Data* (2021) 4.
255. Andreassen A, Nachman B. Neural Networks for Full Phase-Space Reweighting and Parameter Tuning. *Phys Rev D* (2020) 101:091901(R). doi:10.1103/physrevd.101.091901
256. Arratia M, Butter A, Campanelli M, Croft V, Gillberg D, Ghosh A, et al. Publishing Unbinned Differential Cross Section Results. *J Inst* (2022) 17: P01024. doi:10.1088/1748-0221/17/01/p01024
257. Nachman B. Anomaly Detection for Physics Analysis and Less Than Supervised Learning. *arXiv:2010.14554* (2020).
258. Kasieczka G, Nachman B, Shih D, Amram O, Andreassen A, Benkendorfer K, et al. The LHC Olympics 2020 a Community challenge for Anomaly Detection in High Energy Physics. *Rep Prog Phys* (2021) 84:124201. doi:10.1088/1361-6633/ac36b9
259. Aarrestad T, van Beekveld M, Bona M, Boveia A, Caron S, Davies J, et al. The Dark Machines Anomaly Score Challenge: Benchmark Data and Model Independent Event Classification for the Large Hadron Collider. *Scipost Phys* (2022) 12:043. doi:10.21468/scipostphys.12.1.043
260. Karagiorgi G, Kasieczka G, Kravitz S, Nachman B, Shih D. Machine Learning in the Search for New Fundamental Physics. *arXiv:2112.03769* (2021).
261. Farina M, Nakai Y, Shih D. Searching for New Physics with Deep Autoencoders. *Phys Rev D* (2020) 101:075021. doi:10.1103/physrevd.101.075021
262. Heimel T, Kasieczka G, Plehn T, Thompson JM. QCD or what? *Scipost Phys* (2019) 6:030. doi:10.21468/scipostphys.6.3.030
263. Macaluso S, Shih D. Pulling Out All the Tops with Computer Vision and Deep Learning. *J High Energ Phys* (2018) 2018:121. doi:10.1007/jhep10(2018)121
264. Chekanov SV, Hopkins W. Event-based Anomaly Detection for New Physics Searches at the LHC Using Machine Learning. *arXiv:2111.12119* (2021).
265. Dillon BM, Faroughy DA, Kamenik JF, Szewc M. Learning Latent Jet Structure. *Symmetry* (2021) 13:1167. doi:10.3390/sym13071167
266. Dillon BM, Plehn T, Sauer C, Sorrenson P. Better Latent Spaces for Better Autoencoders. *Scipost Phys* (2021) 11:061. doi:10.21468/scipostphys.11.3.061
267. Buss T, Dillon BM, Finke T, Krämer M, Morandini A, Mück A, et al. What's Anomalous in LHC Jets? *arXiv:2202.00686* (2022).
268. Dillon BM, Kasieczka G, Olischlager H, Plehn T, Sorrenson P, Vogel L. Symmetries, Safety, and Self-Supervision. *arXiv:2108.04253* (2021).
269. Metodiev EM, Nachman B, Thaler J. Classification without Labels: Learning from Mixed Samples in High Energy Physics. *J High Energ Phys* (2017) 2017: 174. doi:10.1007/jhep10(2017)174
270. Collins J, Howe K, Nachman B. Anomaly Detection for Resonant New Physics with Machine Learning. *Phys Rev Lett* (2018) 121:241803. doi:10.1103/physrevlett.121.241803
271. Collins JH, Howe K, Nachman B. Extending the Search for New Resonances with Machine Learning. *Phys Rev* (2019) D99:014038. doi:10.1103/physrevd.99.014038
272. ATLAS collaboration. Dijet Resonance Search with Weak Supervision Using $\sqrt{s}$ = 13 TeV Pp Collisions in the ATLAS Detector. *Phys Rev Lett* (2020) 125:131801.
273. Mikuni V, Nachman B, Shih D. Online-compatible Unsupervised Non-resonant Anomaly Detection. *arXiv:2111.06417* (2021).
274. Agarwal G, Hay L, Iashvili I, Mannix B, McLean C, Morris M, et al. Explainable AI for ML Jet Taggers Using Expert Variables and Layerwise Relevance Propagation. *J High Energ Phys* (2021) 2021:208. doi:10.1007/jhep05(2021)208
275. Mokhtar F, Kansal R, Diaz D, Duarte J, Pata J, Pierini M, et al. Explaining Machine-Learned Particle-Flow Reconstruction. *35th Conf Neural Inf Process Syst* (2021) 11:211112840.
276. Huffman BT, Jackson C, Tseng J. Taggingbquarks at Extreme Energies without Tracks. *J Phys G: Nucl Part Phys* (2016) 43:085001. doi:10.1088/0954-3899/43/8/085001
277. Todd Huffman B, Russell T, Tseng J. Tagging B Quarks without Tracks Using an Artificial Neural Network Algorithm. *arXiv:1701.06832* (2017).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Bonilla, Chachamis, Dillon, Chekanov, Erbacher, Gouskos, Hinzmann, Höche, Huffman, Kotwal, Kar, Kogler, Lange, LeBlanc, Lemmon, McLean, Nachman, Neubauer, Plehn, Rappoccio, Roy, Roloff, Stark, Tran, Vos, Yeh and Yu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

ORIGINAL RESEARCH
published: 24 June 2022
doi: 10.3389/fphy.2022.887738

Quantum Systems for Enhanced High Energy Particle Physics Detectors

M. Doser[1], E. Auffray[1], F.M. Brunbauer[1], I. Frank[1,2], H. Hillemanns[1], G. Orlandini[1,3] and G. Kornakov[4]*

[1]CERN, Geneva, Switzerland, [2]Faculty of Physics, Ludwig Maximilian University of Munich, Munich, Germany, [3]Dept. of Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, [4]Faculty of Physics, Warsaw University of Technology, Warsaw, Poland

Developments in quantum technologies in the last decades have led to a wide range of applications, but have also resulted in numerous novel approaches to explore the low energy particle physics parameter space. The potential for applications of quantum technologies to high energy particle physics endeavors has however not yet been investigated to the same extent. In this paper, we propose a number of areas where specific approaches built on quantum systems such as low-dimensional systems (quantum dots, 2D atomic layers) or manipulations of ensembles of quantum systems (single atom or polyatomic systems in detectors or on detector surfaces) might lead to improved high energy particle physics detectors, specifically in the areas of calorimetry, tracking or timing.

Keywords: quantum sensors, detectors, particle physics, calorimetry, tracking, particle identification, HEP

OPEN ACCESS

Edited by:
Petra Merkel,
Fermi National Accelerator Laboratory (DOE), United States

Reviewed by:
Pavel Murat,
Fermi Research Alliance, United States
Thomas Cecil,
Argonne National Laboratory (DOE), United States

****Correspondence:***
M. Doser
michael.doser@cern.ch

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *01 March 2022*
Accepted: *06 May 2022*
Published: *24 June 2022*

Citation:
Doser M, Auffray E, Brunbauer FM, Frank I, Hillemanns H, Orlandini G and Kornakov G (2022) Quantum Systems for Enhanced High Energy Particle Physics Detectors. Front. Phys. 10:887738. doi: 10.3389/fphy.2022.887738

1 INTRODUCTION

The development of a wide range of highly sensitive technologies based on the manipulations of small numbers of atoms or on quantum effects that arise at ultra-low temperatures has led to the rapid proliferation of a very wide range of quantum devices, many of which are now beginning to see commercial applications. At the same time, the extraordinary sensitivity of these devices, which rely on discrete state changes from one quantum state to another, makes them ideal detectors for probing very weak interactions between putative ultra-light particles or fields and the quantum devices themselves. This has led to their wide uptake in the field of low energy particle physics and the rapid exploration in recent years of the low energy phase space associated with e.g., axions, ALP's, and many other dark matter candidates (numerous reviews, among them [1–4], have covered these applications).

This same sensitivity would appear to make these devices unsuitable for high energy physics applications, whose detection mechanisms mostly rely on detecting and reconstructing individual particles' properties through the quasi-continuous effects of their interactions with matter, integrating the charge deposited by the continuous process of multiple ionizations of the atoms of the bulk of a detector by the interacting particle. Forming a usable signal that can be differentiated from thermal and statistical fluctuations requires large numbers of such ionization processes to have taken place. Furthermore, existing detector families are already very well suited to high resolution tracking, calorimetry or particle identification.

In this paper, we discuss a small number of quantum devices or systems in which quantum effects play a major role in view of applying them to the areas of particle tracking, particle identification or calorimetry. We particularly focus on applications that potentially could result in information that is currently difficult to obtain, or where some of the existing technologies' boundary conditions or

limitations might be alleviated or circumvented. These are tied to the need for ever better time resolution to deal with increasing pile-up in high energy collider experiments; to the desire to minimize the material budget of detectors so as to minimize multiple scattering, in particular for vertexing; to the interest in determining electromagnetic shower profiles for better particle identification and energy measurement; to the desire to improve on the existing techniques to avoid detector backgrounds or widen the range of employable materials; or to the potential of providing additional information to particle trajectories that would aid in identifying e.g., particle helicities.

The structure of this paper is built around a handful of selected quantum systems, exploring the potential impact of each of these in different areas of high energy particle detection and reconstruction; in most cases, the proposed detection systems will be hybrids of established technologies and of elements whose quantum nature potentially enhances the functionality of the former. The paper does not claim to be exhaustive, but instead compiles a few of what we consider to be some of the more promising near-term applications, and which are highlighted for each of the investigated families of quantum devices in dedicated sections: low-dimensional materials, nano-engineered semiconducting devices, implementation of polarizable support structures, or manipulation of individual atoms in large volume detectors. Given the rather speculative and often very preliminary nature of the detection schemes described below, it has to be emphasized that significant research and development efforts will be required in validating the concepts proposed here for a range of high energy particle detection approaches and in establishing their viability and usefulness.

2 LOW-DIMENSIONAL MATERIALS

Low dimensional materials (nanodots, atomically thin monolayers) offer a wide degree of tunability of their parameters, be it through their elemental composition and crystal stucture in the case of 2-D layers, or their composition and geometric size in the case of nanodots. Their properties range from photon emission to modification of surface properties (when used as coatings) to mechanical barriers (differential transmission of electrons and ions).

2.1 Low Dimensional Materials for Scintillating Detectors

Scintillator-based detection systems are in wide use since many years and in many applications ranging from nuclear and particle physics experiments to medical imaging and security. Their physical properties like density, light yield, linearity of the detector response and operational speed, but also their resistance to harsh radiation load, their insensitivity to small changes in operational parameters and the widely available production capabilities make them one of the most popular devices for the detection and the energy measurement of charged and neutral particles interacting with material structures. The need for the above-mentioned ever better timing resolution in particle physics experiments places, however, increasingly stringent requirements on the time measurement performance of scintillator-based detection systems. The timing performance of future particle detector experiments is key to cope with the need to disentangle bunches of colliding particles with ever smaller temporal separation.

2.1.1 Nanomaterial Scintillators

Conventional commonly used scintillators produce an amount of light proportional to the energy deposited by charged or neutral particles. The energy transfer from initial ionisation in the bulk material to the luminescence centers is complex and leads to an intrinsic time-resolution limit in photoproduction due to the stochastic relaxation processes of the hot electron-hole pairs produced by the impact of radiation on the crystal material. This generates irreducible time jitter and limits the time resolution. To go below this intrinsic limitation, which is a characteristic property of conventional scintillation materials, various ways of exploiting faster photon production mechanisms have been investigated [5–12] among which the development of semiconductor nanomaterials represents a promising route towards fast timing; these have thus been extensively studied over the last years see [13–17].

While conventional semiconductor bulk material is characterized by continuous conduction and valence bands, the reduction of the size of a crystal down to a so-called nanocrystal of typically 1–10 nm size, close or below the Bohr radius, results in the energy levels of both conduction and valence band becoming discrete and quantized due to quantum confinement. The variation of the energy gap as a function of the size of nanomaterials and of the density of states as a function of the degree of confinement both offer the possibility of tuning their opto-electronic properties, such as for instance the emission wavelength, which can be varied from red to blue by decreasing the size of a nanocrystal (see **Figure 1** left) [18, 19]. Depending on the number of dimensions of the confinement, nanocrystals are classified as quantum wells or ultrathin films (one dimension of confinement resulting in a two-dimensional (2D) object), quantum wires (two dimensional confinement resulting in a remaining one dimensional (1D) object) and quantum dots (confinement in all three directions resulting in a zero dimensional (0D) object [20]. The available energy levels in such objects are discretized as a function of the object dimensionality (0D, 1D, 2D) and their size and shape [20–22].

In direct-band-gap-engineered semiconductor nanostructures, one effect of quantum confinement consists of a significant enhancement of Coulomb interactions between charge carriers of electron-hole pairs, coherent and multi-exciton states [19, 23, 24]. This plays a significant role in enhancing the transition dipole moment of absorption and emission and can thus increase the rate of fast radiative transitions resulting in scintillation decay times below 1 ns.

The timing resolution of scintillators is to first order proportional to the square root of the photon density (number of produced photons per time interval), which can itself be

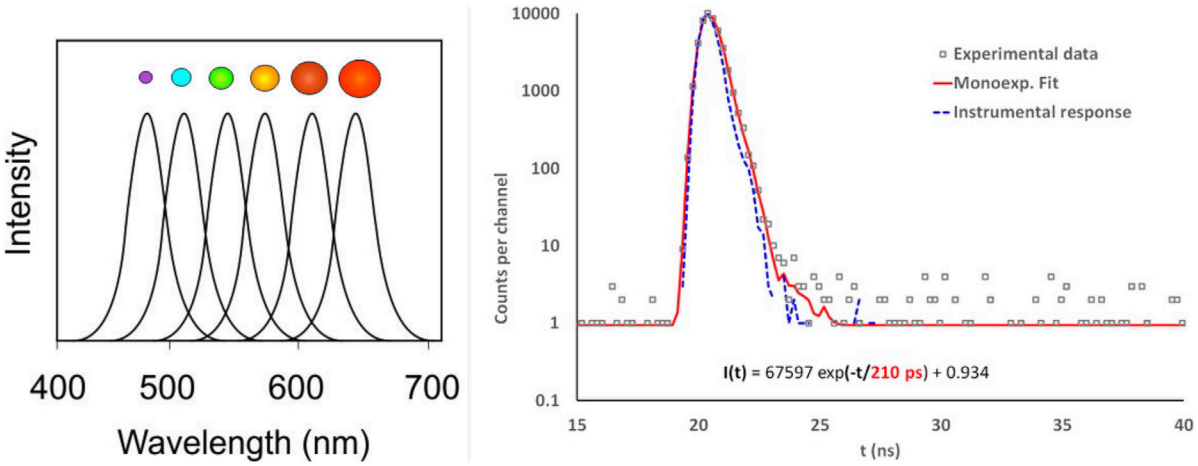


FIGURE 1 | Left: Emission spectrum as a function of nanocrystal size (from [22]). Right: decay time spectra of ZnO(Ga) under irradiation by X-rays (from [15]).

expressed as the ratio of the emission decay time and the light yield of the scintillator. Therefore, to minimise timing resolution, scintillator development aims at achieving a maximum light yield with the shortest possible decay times. Future scintillator based timing layers or time-of-flight detectors aim at achieving a timing resolution below 30 ps in order to be suitable for pileup rejection at high luminosity colliders and for particle identification, and be competitive with e.g., SPAD's [25] or LGAD's [26]. Achieving such timing resolutions with scintillating nanomaterials with a sub-nanosecond decay time would provide flexibility in matching experiment specific performance requirements as well as constraints in terms of costs, radiation hardness and infrastructure needs.

Several types of scintillating nanomaterials with different levels of confinement (nanoplatelet, quantum wire, quantum dots) have been studied over many years. Among them, CdSe and CdSe/CdS [13, 14, 27–29], CdZn/ZnS [30], ZnO quantum dot or nanoplatelets [15, 31–33], InGaN/GaN multi quantum wells [16, 34–36], the Cesium lead halide perovskyte CsPbX3 (X = Cl, Br,I) [17, 37, 38] for instance reach a fast photon emission with characteristic radiative decay times in the range of nanosecond or subnanosecond. An example of the decay time obtained for ZnO(Ga) nanomaterials is given in **Figure 1** right.

The very short decay times of such nanocrystals together with the possibility to tune their emission spectra open new prospects for timing detectors for particle physics experiments, such as precision timing layers for time tagging of collision tracks or scintillators for the energy measurement of particles in combination with high time resolution. Furthermore, if the nanocrystal emission spectrum is tuned into the infrared wavelength band between 1 and 5 μm, for which silicon is transparent (or can be made so via surface treatments [39]), any photons emitted away from the scintillating layer can be detected remotely, even through further semiconductor-based tracking layers, thus opening new possibilities in detector design and functionality.

2.1.2 Time Tagging and Calorimetry

In order to exploit the physical and optical properties of nanocrystals for radiation detectors in various particle physics experiments, R & D efforts need to focus on maximizing the energy deposit in the nanomaterial to have a sufficient number of photons with a very fast decay, increasing the Stokes shift to avoid self-absorption and improving the light transport and light collection of the fast emission. While the production of large volumes of pure nano-crystal based detection devices represents a major technical challenge, layers of nano-crystals can be combined in multiple ways with conventional scintillator materials in so-called hetero-structured scintillators or MetaCrystals [40–43], allowing to simultaneously exploit the properties of bulk scintillators, e.g., in terms of absorption power for the measurement of the energy, and the fast light emission of nano-crystals for timing measurement purposes. One possibility could e.g., be to deposit nano-materials as thin layers of several μm thickness on conventional bulk scintillators [37, 43, 44], together building a sample of alternating scintillator/nano-crystal layers. In this approach the standard scintillator and the nano-crystals are optically separated, thus preserving the high-Z scintillator performance and light collection characteristics, while at the same time adding prompt photons to the signal. The performance of such approaches essentially depends on having a minimum thickness and density of nano-crystal layers to allow for the emission of sufficient amount of prompt photons and on having a sufficient transparency to allow prompt photons to reach the photon detection device. The tunable absorption and emission characteristics of nano-crystals may furthermore allow to convert e.g. ultraviolet scintillation light (e.g., cross luminescence material such as BaF2) or Cherenkov light (eg PbF2, Lead glasses) into visible light, which is more efficiently detectable by photo-detectors, thus being functionally similar to a fast wavelength shifter.

Despite the abovementioned difficulties in producing large volumes of nano-crystals, attempts have been made to produce a

stack consisting of multiple waveguides of thin 10–50 μm thick epitaxially grown layers of InAs/GaAs quantum dot scintillators [45, 46]. Having a segmented photodetector array for the readout of each waveguide integrated in such stack, one can achieve an impressive detector performance in terms of light yield and timing characteristics. Various assembly technologies are currently under study to overcome the technical challenges related to the separation of the epitaxial layers from its substrate and its combination into a stack.

Another possibility consists of depositing one or more layers of nano-crystals directly on photo-sensitive devices in order to increase the sensitivity of the photo-detector towards X-ray and γ-ray energies or charged particles for time tagging purposes and thus to significantly enhance the range of applications of such devices. Among the issues to be addressed in these approaches figure a detailed understanding of the surface chemistry of nanocrystals in view of their deposition in form of thin layers of nanomaterials on conventional scintillators or photosensitive semiconductor devices and the transport of light.Whereas nano-crystals can currently be added to scintillating crystals only on their surfaces or as layers alternating with conventional crystal materials, they can on the other side be dispersed in liquids or embedded in host materials such as organic materials or glasses in order to enhance or replace the conventional scintillation mechanism of liquid or plastic scintillators or doped glasses by a scintillation with faster emission characteristic, which allows their use as a fast component of the above-mentioned hetero-structure, in the form of shashlik detectors or as integrated (quasi-continuous) wavelength shifters.

However, various aspects and issues have to be addressed and improved in order to bring the performance of nanocrystal composites to that of conventional detectors, such as the optimization of the energy transfer between the host and the nanomaterial. Also the concentration of nanomaterials must be optimized in terms of density and homogeneity, while at the same time having a good transparency and avoiding the scattering of the scintillation light. Several developments in these directions are performed [17, 30–33, 38, 47–49] and some projects have already been proposed. One example consists of the production of plastic scintillators exploiting CsPbBr3 perovskites as high Z sensitizers, resulting in a large Stokes shift, a high emission yield and a fast emission lifetime of few ns [38]. The resulting scintillation performance is comparable to conventional inorganic and plastic scintillators, making such scintillators a usable tool for waveguiding over long optical distances and for the detection of high energy photons and charged particles without absorption losses. The Esquire project [50] proposed to use scintillating quantum dot containing isotope componants such as CdSe/ZnS embedded in a host matrix for the study of rare events such as the neutrinoless Double Beta Decay (0νDBD). More recently AIDAinnova project approved a blue sky project "NanoCal" on the proof of concept of a fine-sampling calorimeter with nanocomposite materials [51, 52].

2.1.3 Chromatic Calorimetry

Recent developments in the tuneability and narrow emission bandwidth (~ 20 nm) of quantum dots, quantum wells, carbonized polymer dots, monolayer assemblies or perovskite nanocrystals [53–55] opens the door to a novel approach to measuring the development of an electromagnetic or hadronic shower within a scintillator, with the potential of obtaining a longitudinal tomography of the shower profile with a single monolithic device, via the means of *chromatic calorimetry*. Specifically, a calorimeter module would need to be built from a single high density transparent material that is differentially doped (at high concentrations) along its length with nanodots with different emission wavelengths, those with the longest wavelengths towards the beginning of the module, and those with the shortest wavelengths towards the end. With the currently demonstrated emission bandwidths of 20 nm, and even constraining emissions to take place only in the visible spectrum, overall around twenty different differentiable emission regions can be envisaged, thus providing for fine grained shower development measurements. The radiation tolerance of these specific nanocrystals remains however to be established. Such a device that would function like a polychromatic embedded wave length shifter thus maps the position and local intensity of the stimulating radiation within the overall module onto the wavelength and intensity of the produced fluorescence light; multiple emission regions can be uniquely identified in a single measurement. One major challenge in implementing such a scheme resides in incorporating nanodots in existing dense crystals during their growth; as mentioned earlier, possible alternatives could be to either intercalate thick dense transparent absorber regions with thin, lower-density, radiation resistant, nanodot doped transparent layers; or to embed both nanodots and microsopic high density crystals within the same low density organic or glass bulk material. Another challenge is the radiation resistance of the bulk material, but also of the nanodots: carbonized polymer dots may not reach the expected radiation resistance of e.g., $CsPbCl_3$ [55], although irradiation tests on triangular carbon nanodots have yet to be carried out.

Naturally, re-absorption (and thus frequency shifting) of light emitted at different points along the cascade has to be avoided: both the bulk material and the subsequent nanodots must be transparent to the photons originating from successively earlier points in the cascade. Positioning of the nanodots emitting the lowest energy photons at the upstream end of the module, and those emitting the most energetic ones at the downstream end (**Figure 2**, bottom right), together with a nanodot absorption spectrum that only down-converts higher energy photons into lower energy ones, ensures this spectral transparency.

Figure 2 shows the absorption (left column) and emission (right column) spectra of six different types of nanodots; the emission spectrum of e.g., carbonized polymer dots of around 680 nm lies above the absorption line of the subsequent triangular carbon quantum dot that only absorbs up to 582 nm, while emitting at 593 nm, which itself cannot be absorbed by the yellow-light emitting nanodot, and so forth. The final Perovskite nanocrystal in this example can not absorb any of the wavelengths of the earlier nanodots, since its absorption spectrum lies below 420 nm. In this manner, UV or higher energy light produced by the shower stimulates emission by

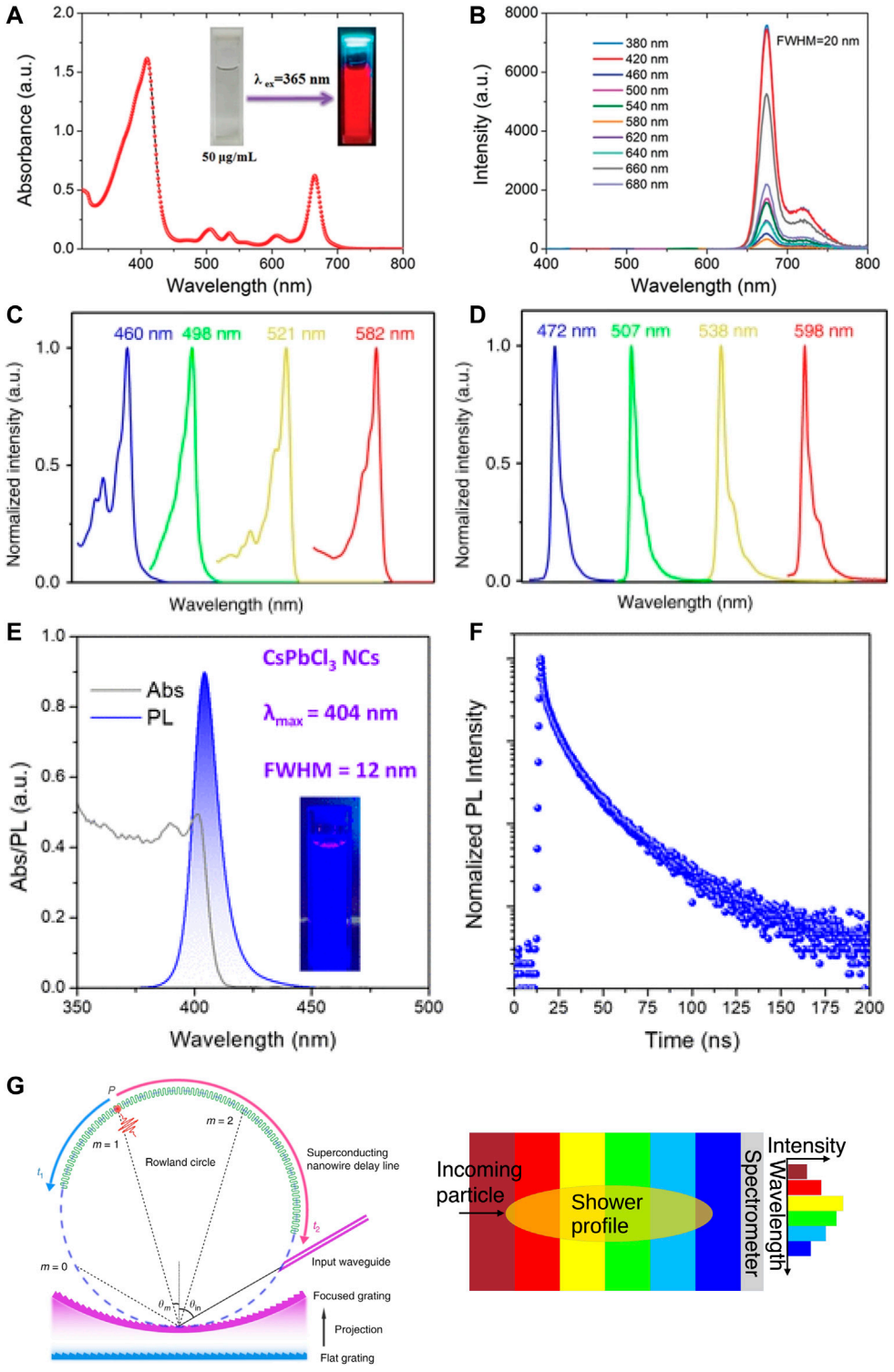


FIGURE 2 | Absorption and photoluminescence (PL) emission spectra of selected narrow bandwidth quantum dot emitters. Top row: Carbonized polymer dots (CPD's): absorption and emission spectra [53]. Second row: triangular Carbon quantum dots. The normalized UV-vis absorption **(C)** and photoluminescence **(D)** spectra of B-, G-, Y-, and R-NBE-T-CQDs, respectively [54]. Third row: CsPbCl3 nanocrystals. **(E)** optical absorption and PL spectra (excitation wavelength 350 nm) (inset: corresponding photograph under UV light), and **(F)** time-resolved photoluminescence (TRPL) decay curve of the synthesized $PhPOCl_2$-based $CsPbCl_3$ nanocrystals [55]. Bottom row: **(G)** Sketch of a compact nanowire-based spectrometer [56] and **(H)** of the overall scheme for the positioning of the different emission-wavelength specific nandots. The histogram to the right corresponds to the spectral intensity, and thus the depth profile, of the shower initiated by the incoming particle. Figures by permission.

nanodots at wavelengths that are not absorbed by any of the subsequent nanodots, and whose spectral intensity provides a proxy for the shower energy deposit at the depth corresponding to a specific nanodot emission wavelength. Furthermore, both onset and decay of photoluminescence is at the nanosecond timescale, as shown for $CsPbCl_3$ nanocrystals in **Figure 2**, providing for excellent shower timing and, possibly, even determining the temporal evolution of the shower itself.

In order to extract the shower profile from the spectral intensity distribution, the photodetector must be able to resolve the intensities and timings of the individual spectral lines. Very recently, compact moderate-spectral-resolution or narrowband spectrometers based on nanowires [56] or nanodots [57, 58] have been produced or are under active development; more traditional, albeit bulkier, alternatives based on Bragg spectrometers or prismatic structures, coupled to photodiode arrays are also imaginable.

2.2 Low Dimensional Materials for Gaseous Detectors

Gaseous detectors are widely used as large area detection systems in HEP experiments owing to their high gain factors, rate capabilities and compatibility with harsh radiation environments. MicroPattern Gaseous Detectors (MPGDs) feature good granularity and are employed as tracking detectors as well as for the readout of Time Projection Chambers (TPCs) among other applications. State-of-the-art MPGDs achieve high spatial resolution, energy resolution and specific developments of precise timing MPGDs have demonstrated <25 ps timing resolution [59].

To enhance the performance of MPGDs in view of future applications, low-dimensional materials are considered for different aspects of detectors: they may be used to tailor the primary charge production process, protect sensitive photocathodes in harsh environments or improve the performance of the amplification stage.

2.2.1 Enhancement of Charge Conversion in Low Dimensional Materials

While gaseous detectors conventionally rely on primary ionisation of gas in the conversion region by incident radiation, solid conversion layers or photocathodes offer a number of advantages by generating primary electrons in a well defined location allowing significantly improved timing resolution. The efficiency and spectral response of conversion layers or photocathodes also directly defines the sensitivity of the detector.

Low dimensional materials and nanostructures can offer new approaches to implement performant conversion layers for radiation detectors and may offer both increased quantum yield as well as access to specific ranges of sensitivity to incident radiation. This can range from exploiting nano-scale geometries to increase the surface available for absorption and photoelectron emission to engineering low dimensional structures to enhance photoemission by resonant processes. Significant improvements of efficiency over thin film or bulk materials as well as tunable work function have been demonstrated in systems such as nanostructured plasmonic surfaces [60], single-wall carbon nanotubes [61], Mg nanodots [62] or graphene layers [63], to name but a few. While some systems target broadband response for versatile detectors with wide spectral sensitivity ranges, other materials offer highly selective and tunable response making them promising conversion layers for particle identification methods. We will suggest possible applications for nanostructures as charge conversion and photocathode layers and highlight their potential for novel radiation detectors.

A key application for photocathodes in gaseous radiation detectors is their use in precise timing detectors where primary charge production needs to occur at a specific time in a well defined location. By converting Cherenkov light from a radiator to primary electrons with a semi-transparent CsI photocathode and using a Micromegas-based amplification stage, the PICOSEC Micromegas detectors have achieved <25 ps timing precision [59]. The efficiency of the photocathode directly translates to achievable timing resolution and while metallic or other robust photocathodes would offer resistance against environmental effects and ageing during prolonged operation, their QE is typically too low to be suitable for this detection concept.

An enhancement of photocathode QE by resonant processes in low dimensional structures could offer a possibility to overcome this limitation. Studies on nanodots of different sizes suggest that a significant enhancement of photocathode QE may be achieved by a discretisation of energy levels arising from quantum confinement [62]. Enhancement factors as high as 38 have been shown for Mg nanodots with diameters of 52.2 nm compared to Mg thin films with a strong dependance on dimensions [62]. While the efficiency of metallic photocathodes enhanced by nanostructuring would still be below the QE of semiconductor photocathodes it may be attractive to profit from the robustness of such structures.

Resonance-enhanced multiphoton photoemission processes have also been observed in single-wall carbon nanotubes [61] along with ultrafast emission timescales. In addition to enhanced QE, the high anisotropy of nanotubes also leads to a dependance of their response to different polarisations of incident light as well as a modification of absorption spectra for different geometries, thus providing a high level of flexibility in tuning detector response. In addition to sensitivity enhancements and optimisation to specific wavelength ranges, particle identification (PID) methods can profit from the selectivity of the response of nanomaterials. Tuning the band gap or work function has been demonstrated in a variety of systems, including modifications of graphene layer work function by UV exposure or plasma treatment [63] or the engineering of nanophotonic crystals to cover specific spectral ranges [64]. While nanophotonic crystals provide enhanced sensitivity for narrow spectral ranges and are thus selective for specific particle momenta in PID detectors, stacks of 1D photonic crystals of different periodicities can simultaneously offer high sensitivity to different particle momenta and may allow for efficient PID even in high particle flux environments [65]. Being selective to discrete bands of particle momenta can be a significant advantage in

mitigating pileup and preserving PID capabilities in high rate experiments.

2.2.2 Graphene or Other 2D Materials as Photocathodes Encapsulant

The use of photocathodes in gaseous detectors is advantageous in the improvement of both time and spatial resolutions. Semiconducting photocathodes such as CsI provide high quantum efficiency (QE) in the UV range but have a limited lifetime due to 1) environmental condition such as humidity and 2) ion bombardment in gaseous radiation detectors. Surface coatings with 2D materials may enhance the lifetime by blocking incident ions while also modifying surface work function thus increasing QE. Theoretical studies [66, 67], based on ab initio density functional calculations, have been explored showing how a hBN layer on top of alkali-based semiconductive photocathodes should decrease the work function.

Ongoing promising studies demonstrate an increased operational lifetime by encapsulation of semiconductive [68, 69] as well as metallic [70] (i.e., Cu) photocathodes with few layers of graphene (from two up to eight layers) while lowering the QE. The decreased QE is mainly attributed to issues during the transfer process and graphene quality and achieved experimental values are approximately one order of magnitude below theoretical predictions [68].

2.2.3 Tailoring Microscopic Transport Processes

In addition to applications for charge conversion layers and photocathodes, atomically thin layers may be exploited to optimise the operation of gaseous detectors and tailor microscopic transport processes of charges. Gaseous detectors suffer from the back flow of positive ions created during charge amplification to the drift region which can lead to significant distortions of electric fields impacting subsequent events and is of particular importance in gaseous TPCs. Graphene has previously been proposed as selective filter, which could suppress the ion back flow fraction while permitting electrons to pass [71]. Graphene is the thinnest 2D material in nature with single atom thickness composed of sp^2 hybridized carbon atoms arranged with a honeycomb symmetry. Thanks to its electrical and optical properties it has been used in various applications including ultrafast photodetectors and FETs. Despite being one atom thick, graphene is impermeable to atoms [72] and its mechanical properties as elastic properties and intrinsic breaking strength of free standing layers make it the strongest known material [73]. As shown in [72], carbon's π-orbitals are delocalized and thus do not allow even to the smallest molecules to pass through the layer. At the same time, graphene is expected to be transparent to electrons traversing the sheet due to its low electron density $n \sim 10^{12}/cm^2$ in a perpendicular direction, which should hold even for low electron energies in the range of 5–10 eV. Since in gaseous detectors the mean energy of primary electrons cannot be significantly higher than 10 eV because it is limited by the electrons mean free path between subsequent interactions with atoms/molecules of the gas, a good low energy electron transmission through the graphene layer is important in order to apply graphene as a filter for ion back flow.

Freely suspended single or few layer graphene membranes on top of tens of micrometer large holes in Gaseous Electron Multipliers (GEMs) may block ions while permitting electrons to pass and participate in avalanche multiplication. First evaluations of this approach were limited by defects of transferred graphene layers, which degraded electron transparency. Exploiting advances in the growth and transfer of graphene, ion and electron transparency of graphene membranes is being explored and may be used to develop low ion back flow detectors based on GEM technology.

Several theoretical and experimental works were carried out in recent years, investigating electron beam transmission through graphene, by measuring the transparency as a function of the incident electrons' energy [74–78]. The results of these studies are often contradictory, especially in the very low energy electron range around 5–20 eV. The transverse electron transmission coefficient through graphene is usually measured in vacuum where layers show high transparency almost close to unity to electrons with energies ranging from tens of keV up to 300 keV. These energy values are about three to four orders of magnitude higher than the energies in gaseous detectors and are commonly used for transmission electron microscopy (TEM) where graphene can be used as a sample support for TEM measurements [74, 75]. Transparency for electron energies in the low energy range of tens up to hundreds of eV can be evaluated by the use of electronvolt transmission electron microscopy. Graphene shows a good transparency of about 60–70% for electron energies from 40 to 50 eV up to 100 eV [76, 77]. For low electron energies below 15 eV, experimental results have shown discrepancies which can result from the strong dependence on graphene quality at these low energies ranges [76, 77]. In particular for electrons with energies below 10 eV there is no systematic investigation to date but promising results for this low-energy region, which suggest that the transmission coefficient of electrons with energy less than 10 eV can achieve as high as 99% transparency [78]. In addition to selective ion filtering, freestanding graphene membranes with transparency to primary electrons may also be used to physically separate drift and amplification regions of the detectors working as gas separator [79–81] and profit from additional flexibility in the choice of gas mixtures optimised for high conversion efficiency in the drift region and suitable mixtures for high electron amplification factors.

3 NANOENGINEERED SEMICONDUCTING DEVICES

3.1 Chromatic Tracking

The position resolution of existing semiconductor pixel or microstrip detectors lies in the region of 10 μm, well suited to the current generation of high energy physics detectors. For future collider experiments, such as FCC-ee, individual hit resolution must be of the order of 3 μm [82] in order to allow achieving a combined vertex resolution of around 5 μm.

Particularly, improved position resolution in the first detection plane (first scattering plane) after the primary vertex would allow improved secondary vertex determination, better particle discrimination, and higher jet flavor identification probability. While ultra-thin silicon microstrip and pixel detectors are a very effective approach to address this challenge, their resolution remains limited to the few μm level.

The scintillation properties of quantum dots (**section 2**), together with their O (10 nm) dimensions, suggest a possibility of improving the hit resolution of these innermost detection planes even further. Tests of InAs/GaAs quantum dots with α particles have provided evidence for a significant photo-electron yield (3×10^4 photoelectrons per 1 MeV of deposited energy) and very rapid luminescence (emission time 0.3 ~0.6 ns) [83]. Furthermore, due to the mismatch between the differences between the base layer lattice and that of the InAs quantum dots, these self-assemble: in [83], a quantum dot density of about 3.5×10^{10} cm^2, with lateral dimensions of 14 ~18 nm and heights of 5 ~6 nm was achieved.

It is conceivable to cover the surface of a silicon microstrip or pixel detector, but also of thin intercalated dedicated planes, with thin layers of light-emitting material in form of such semi-conducting quantum dots or quantum wells. While the functionality of existing quantum dots is not subject to external power sources, it may be interesting to consider the possibility that they could be coupled to the power distribution system of the silicon (or GaAs) detectors, perhaps leading to enhanced control over their dynamic properties. Moreover, it might be possible to do this in such a manner that each strip of a silicon microstrip detector is subdivided into a limited number of sub-micron wide bands. Self-assembly of nanodots [84] would appear to ensure that each microstrip zone would contain identical periodically-spaced nanodots, and thus nanodots producing light of the same frequency. If it is possible to affect this self-assembly process, one could effectively introduce a periodic, long-range modulation of the nanodot dimensions within each zone, and thus of the emission wavelengths. While no technology has yet been developed that would allow effectively controlling and modulating the growth of nanodots at the atomic scale over large distances, some possible directions to investigate could be through the moiré-modulated substrate interaction discussed in [84], through interferometric laser annealing of the deposited layer [85] or through careful choice of the lattice mis-match between substrate and few atomic-layer thick additional layer. These could then lead to e.g. periodic changes in the chemical composition of the nanodots [86]).

If feasible, this would result in several distinct sub-bands (or at least a continuous variation of the dimensions of the formed nanodots) on the inter-strip or pixel scale, each concentrating nanodots of a specific size, corresponding to emission of light at a zone-specific frequency range [54]. Detecting the frequency of the light emitted by these nanodots in coincidence with the strip signal would then allow uniquely identifying which band the charged particle traversed, effectively resulting in a sub-micron position resolution. Also here, narrow-band photospectrometers (as in **section 2.1.3**) will be required.

To enhance and control the photon yield in such layers, an alternative to the passively emitting quantum dots is provided by epitaxially grown intermediate structures between quantum wells and quantum dots [87] which combine the light yield of quantum dots with the active control of quantum wells.

Similar to the above approach, thin layers coated with layer-specific nanodots, thus resulting in light of a layer-specific frequency, can be intercalated between tracking layers. If the emission spectrum of these is chosen in the mid-IR, then the emitted scintillation light will be able to pass through any intermediate silicon-based tracking layers, silicon being mostly transparent for infrared wavelengths between 1 μm and 10μ. A spatially and spectrally resolving IR camera positioned outside of the tracking layers would thus also be able to detect the emitted IR photons, adding spatially (and possibly even temporally) resolved hits to any tracks. This approach can be carried out also inside of the innermost tracking layer by an appropriate coating of the beam pipe at the heart of collider experiments. Naturally, care will need to be taken to minimize absorption through services or silicon-tracker internal structures, by e.g., use of ultrathin (few nm) metallic layers or of IR-transparent conductors.

3.2 Quantum Cascade (Active) Scintillators

While in most widely used scintillators, minimum ionizing particles excite electrons from the valence band into the conduction band, which then emit fluorescence light due to transitions between energy band levels (inorganic scintillators) or through molecular excitation and relaxation processes (organic scintillators), these processes can not be actively enhanced (except for static doping with activators), nor electronically controlled. In other words, there is no possibility of "tuning", of "switching on" (or "off"), of "priming" the properties of the scintillating material dynamically, nor of benefiting from quantum effects beyond those occurring within the existing crystal lattices or the molecular constituents of the scintillators. In recent years, a number of authors have explored novel types of "custom-built" scintillators, whereby the composition, the structure or the surface of the scintillating material is controlled at the nm scale, with significant potential for dramatic improvements in light yield and temporal resolution [88, 89]. But also these nanostructured scintillators, such as those proposed for chromatic tracking above, as essentially passive devices: the scintillation light they emit is produced by spontaneous de-excitation of the nm-sized structures excited by the passage of a charged particle or the interaction with high energy photons, and its frequency is defined by the chemical composition, geometry and dimensions of the emitting structure.

In contrast, active components should allow—in principle—tuning both frequency and intensity of the emitted light. It is thus tempting to consider systems that would either allow "priming" a detector shortly prior to the passage of charged particles of interest, or that would allow "triggering" optical transitions from excited states after the passage of particles, depending on the intended use. Photo-emitting nano-structures are of wide industrial interest are consequently undergoing very active development. Of particular interest here is F-band photoluminescence (see e.g., [90] for a review).

Of particular interest here are quantum cascade lasers (QCL), whose series of wells are carefully tailored to allow a coherent process of photon emission, triggered by the successive tunneling steps of an electron from the central potential well into subsequent wells at lower potential, to take place [91]. A modified QCL with a small number of wells (potentially a single one) would provide the possibility of functioning not as an "amplifier" for the initial electron of a QCL that is converted into a large number of monochromatic photons through the multiple stages of the cascade, but rather as a single stage "converter" of the large number of electrons produced through ionization by the passage of a m.i.p. through the QCL's central well into the same number of (monochromatic) photons as the electrons tunnel through the single stage cascade structure.

Detection of the photons alone would only result in such modified QCL structures behaving like "active scintillators"; in order for them to function additionally as trackers, a correlation between their position and the frequency of the light produced by each individual active element is required. The tunability of QCL's can provide this correlation, either by fine-tuning the dimensions or by precise control of the voltages applied to each QCL; in either case, under realistic conditions, only a limited number of differentiable frequencies could be produced. This scheme would thus only work in concert with additional trackers that can disambiguate the specific QCL from which a photon was emitted from the subset of identical QCL's. Readout of such photonic trackers would, as indicated above, be simplified if the emitted photons were in the infrared regime, as is typical of QCL's, as the photodetectors could then be remote to the emission region.

4 NANOPHOTONICS, METAMATERIALS, AND PLASMONICS

4.1 Nanophotonic Cherenkov Detectors

Nanophotonics, metamaterials and the engineering of plasmon modes in nanostructured materials lies somewhat at the border of quantum sensors, but given that it too represents engineering of materials at the nanometer scale and relies on quantum effects brought forth by the interactions between small numbers of atomic or molecular systems, it makes sense to briefly refer in this paper to opportunities appearing due to recent developments in this field. In addition to the surface treatment of scintillators [92] touched upon in **section 2.1.2**, and which is being pursued to enhance the light yield in the case of photonic crystal scintillators [93, 94] or plasmonic scintillators [95, 96], recent work on Cerenkov light generation also points towards potentially interesting possibilities.

The material limitations (constraints imposed by the use of low refractive index materials and the concomitant low photon yield, or, in the case of dielectrics, very similar Cerenkov angles for high energy particles) in Cherenkov detectors can be partly overcome by using modern concepts from nanophotonics and metamaterials: longitudinal plasmon modes in nanometallic layered materials [97] can allow achieving continuously-tunable enhanced Cherenkov radiation. In another approach [98] relying on the Brewster effet, structures built of 1D photonic crystals of different periodicities but identical constituent materials, form a broadband angular filter. While this Brewster-Cerenkov scheme is significantly more sensitive to particle velocity than approaches based on the standard Cerenkov angle (albeit at the price of a reduced photon flux), the particle to be identified must travel parallel (within 0.5°) to the surface of this structure, a limitation that will require appropriate detector designs.

This directional limitation is however also an invitation to consider using appropriately-constructed Cerenkov detectors as elements of the charged particle tracker. The nanoscale sensitivity of surface Dyakonov-Cerenkov radiation induced by the motion of charged particles in birefringent crystals [99] (radiation emission is greatly enhanced when the particle is within 200 nm of the surface of the birefringent material) makes it possible to consider very high spatial resolution tracking detectors based on the detection of the Dyakonov surface waves. However, the emission of secondaries through the interaction of the primary within the high-density birefringent material and the concomitant energy loss (which limits the photon yield) present serious challenges towards a concrete realization of such a detector. Nevertheless, it is encouraging that approaches based on engineering scintillator or crystal structures at the nanometer scale might open up completely new functionalities or detector families.

5 NITROGEN-VACANCY DIAMONDS

Defects in diamonds, which are the reason for the variety in coloration of the jewels, have been studied for decades. One of the most studied is the Nitrogen-Vacancy (NV) defect. Its spin state can be optically prepared and be read out [100–102] using the photoluminescence properties of the crystal [103]. The observed quantum properties at room temperature [104] have attracted intense interest and find applications in the areas of quantum information [105], quantum manipulation [106] and quantum sensing with unprecedented accuracy of absolute nanoscale-resolution measurements of magnetic and electric fields [107], spin [108], strain [109] or temperature [110].

The color center can be in a negative (NV^-) or neutral (NV^0) state [111]. The NV^- center has a detectable magnetic resonance associated with its ground and excited levels, which is not the case for the NV^0 center. For this reason, the NV^- system is usually preferred. The optical transitions of the defects are far from the conduction and valence band levels of the diamond [112]. The NV^- defect can be spin polarized using optical excitation, allowing a precise control of its state. This property allows one to use them as polarimeters. However, their use for direct detection requires to have the probe in the atomic vicinity of the sensed particle: the use of electronic spin reporters on the surface of the diamond allows to read at nanometric distances the spin of single particles [113].

The applications that could potentially incorporate nitrogen vacancies in diamonds benefit from the following intrinsic properties: the vacancies constitute a self-calibrated instrument

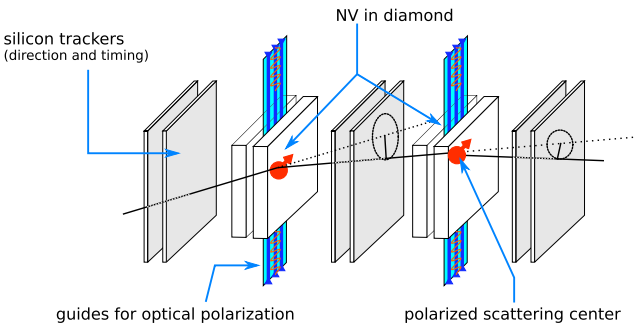


FIGURE 3 | Conceptual scheme of a charged particle tracked with incorporated polarimeter based on NV's in diamond as polarizable scattering medium. The polarization is measured particle-by-particle by using the right-left asymmetry in the scattering probabilities. Two planes of tracking detectors, such as silicon pixel or microstrip detectors, are needed to determine the particle direction before and after the polarized scattering plane. The polarization of the diamonds can be achieved using optical means.

based on known optical transitions, exhibit fast optical spin-polarization, sensitivity to static and dynamic magnetic fields, long coherence times and coupling to neighboring nuclear and electronic spins, and are reliable in handling and manufacturing. Here, we propose two schemes for tracking and polarimetry using NV's in diamond as active medium of a new kind of particle detectors. The first scheme introduces the optical polarization of the NV center as the active scattering medium which deflects particles depending on the incoming particle spin orientation; the second scheme is that of an active target, prepared to be operated in close vicinity of the reaction in order to sense the remaining low energy fragments after the collisions have taken place.

5.1 Polarimetry and Tracking: NV in Diamond Arrays as a Polar Tracker

The measurement of observables that can be related to the spin orientation of a particle provides an additional degree of control for understanding the underlying physical processes. In this sense, the internal structure of the nucleon, nucleus, the origin of the spin of hadrons or the spin properties of the deconfined fluids produced in collisions of heavy ions could be explored using particle trackers enhanced with intercalated polarized thin scattering planes for measurement of left-right azimuthal asymmetries along the ϕ angle of the particles impinging on these polarimeters.

Currently, spin physics is limited to facilities with polarized beams [114] and targets [115] or to particles such as the self-resolving weak decays, such as of the Λ baryon [116]. Extending the availability of measurements to other hadrons could bring new insights into their structure and interactions or those from whose decay they stem.

The intrinsic optical polarization properties of the NV defects in diamond can be suitable for construction of thin polarized layers with them. Charged particles undergoing elastic scattering with the polarized centers will have a small anisotropy in the left-right scattering direction. By measuring the process several times, a probabilistic estimate of the polarization of the particle can be reconstructed according to the scheme in **Figure 3**. The simplest setup would consist of a series of thin pairs of silicons or detectors with similar position sensitivity providing high pointing accuracy and a reduced probability of scattering interleaved with NV's in diamond planes embedded in guides enabling the polarization of the defects. These defects can be efficiently polarized [102] and slightly modify the helicity-dependent scattering direction probability of a charged particle crossing the polar tracker, thus providing access to determining single particle spin orientations.

The probability of scattering in a polarized atom is directly proportional to defect abundance in the diamond. The density of defects is one of the parameters that is actively being optimized as the sensitivity as a magnetometer scales with the square root of its number [117]. Several existing approaches using modified deposition schemes [118] or creation of defects with laser radiation [119] can achieve densities of $10^{16-18}\,\mathrm{cm}^{-3}$. In spite of the possibility of locally enhancing this density by another two to three orders of magnitude through hyperpolarization [102] (polarizing the atoms surrounding the defects), an additional increase by two orders of magnitude in the defect abundance (or the size of the locally polarized region surrounding them) will nevertheless be required before such helicity trackers can be realistically contemplated.

5.2 Polarimetry and Tracking: NV in Diamond as Active Targets for Sensing Heavy Fragments

In the previous scheme we have incorporated in the measuring scheme only the optical polarization property of the NV's in diamond. Diamonds are known for having a high tolerance to ionizing radiation. This makes them suitable beam detectors [120]. Diamonds with defects could however be also used as part of an active target. The high sensitivity of the defects to the presence and spin of even single particles [121] can be used to sense the production and the spin of the remaining highly

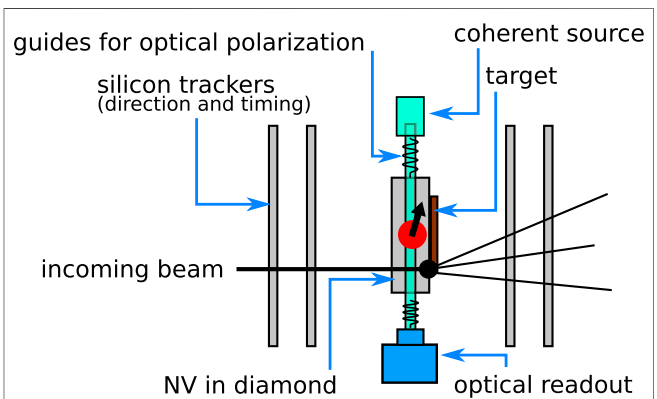


FIGURE 4 | Scheme of an active target based on polarized NV's in diamond used to sense the remaining heavy and slow spectator fragments. The sketch shows the silicon trackers, the target material and the optoelectronic system used to prepare the diamond for the measurement and to perform the readout.

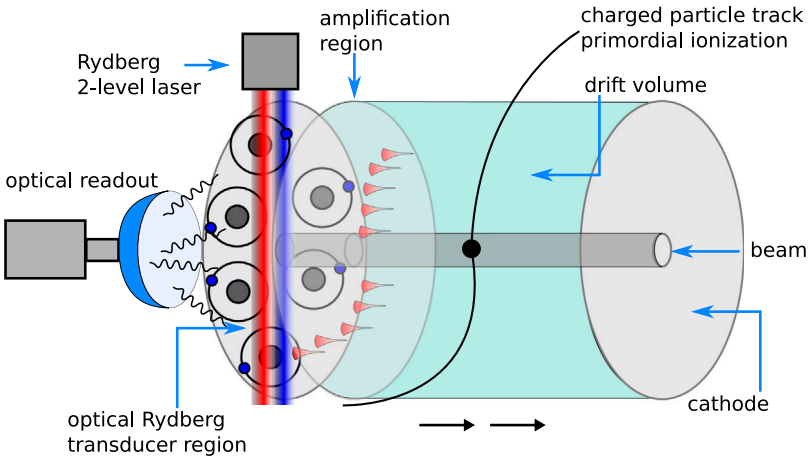


FIGURE 5 | Scheme of a projection chamber with a Rydberg transducer of the EM pulses produced in the amplification region and which are upconverted into the visible domain for optical readout. The amplification region furthermore produces an enhanced number of avalanche electrons due to the low ionization threshold of the Rydberg atoms. The atoms are excited by a two-level laser system. The rest of the chamber follows a classical design with a magnetic field to sense the Lorentz force and the homogeneous electric field to transport the electrons and ions from the primordial ionization region to the amplification region.

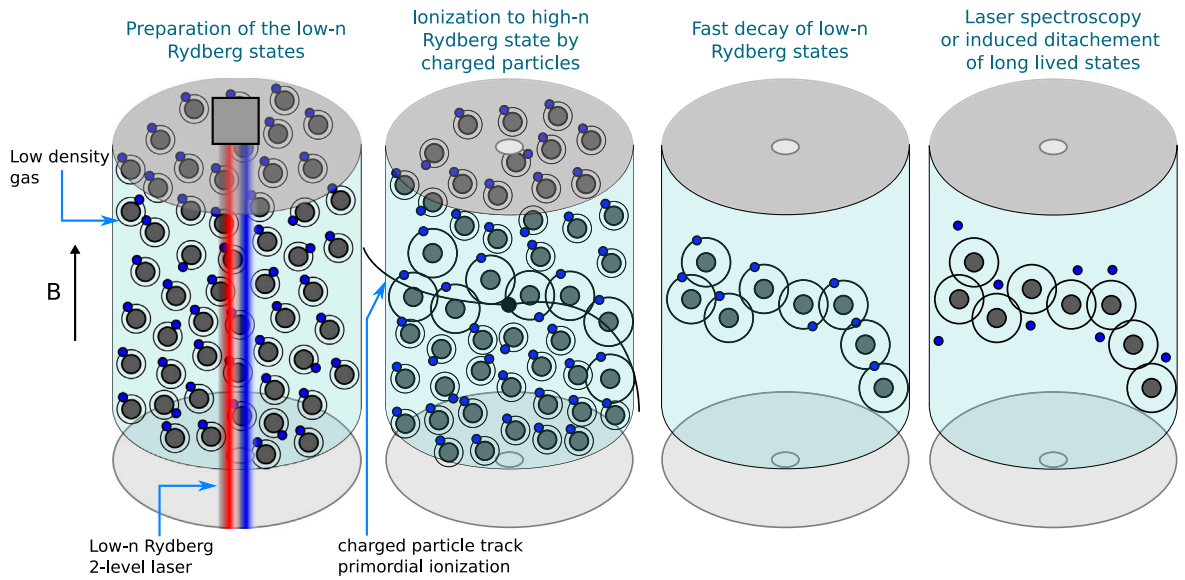


FIGURE 6 | Scheme of a tracking chamber using Rydberg atoms as sensing medium. The measurement is performed in four steps. First, via two-level laser excitation, the atoms in the primary ionization region are brought to a low-n Rydberg state synchronously with the collision time. In the second step, the charged particles produced in the reaction ionize or excite the Rydberg atoms along their flight path. After that, the short lived states decay back to the ground state and the highly excited Rydberg states can be read out using opto-electronic means, such as electromagnetic induced transparency. Alternatively, the electrons can be photo-detached and detected as an effective locally increased ionization cloud using standard techniques.

charged ions (HCI), left after induced fission happens in the interface between the polarized diamond and the target.

The scheme of such an active target is shown in **Figure 4**. The beam particle hits the target after crossing the first (silicon) tracking section which provides accurate directional and impact location at the position of the diamond target. There, the beam particle interacts with the target producing high momentum particles which are projected in the forward direction. Remnants of the target nucleus might survive and remain in the vicinity of the target with low kinetic energy. The thin silicon trackers provide additional information needed to constraint the locus of the interaction and the vertex and momenta of outgoing particles. The optical readout, which can be synchronized at ns-level with the optical control of the diamond provides background free

measurements of the participants and spectators of the collisions. Such a design could also facilitate the measurement of the collision centrality in fixed-target experiments.

6 MANIPULATION OF ATOMS

In recent years, the usage of Rydberg atoms as a key element in many applications has been explored [122]. Their features, related to the precise control of the quantum state, their long lifetime and their size has stimulated their use in quantum computing [123], as sensitive volume [124] or as transducers of electromagnetic radiation of different wavelength [125–127]. The first approaches to produce atomic Rydberg states used collisions, making precise manipulations impossible. Replacing this approach by a laser-based excitation scheme [128] however results in precise control of the accessed energy levels. Rydberg states can be very large, more than 10^4 times the size of ground state atoms, when they are in high n-states, with n being the principle quantum number. As is the case for many of their properties, their size scales with increasing n, specifically with n^2. Their lifetime, after including all possible decay modes, scales as n^3, being of the order of 10^{-5} s for n ~ 50, making them perfect candidates for long interaction time exposures. Another source of sensitivity is the reduction of the energy needed to strip an electron from a highly excited Rydberg state as compared to starting from ground state atoms. In general, the required energy decreases with n^{-2}, making them exceptionally sensitive to small changes happening in the medium, be it external electromagnetic fields or the presence of other atomic species or neighbouring Rydberg atoms.

6.1 Optical Tracking TPC's

High precision charged particle tracking in large volumes requires complex systems, sometimes with more than half a million read-out channels [129]. In addition to collecting charge, recording detector information using electro-optical means is also possible by detecting the fluorescence light produced in the avalanche amplification region [130, 131] of gas-based detectors. This approach has the advantage of providing high accuracy tracking of complex events at a reduced cost [132, 133]. Enhancing tracking detectors with atoms far from the ground state can be interesting in the readout, multiplication or in the ionization regions.

The high sensitivity to electric and magnetic fields of Rydberg atoms makes them ideal candidates for optical transducers [127] for sensing the electron avalanches in the multiplication zone of a projection chamber detector. Sensors based on Rydberg states are not only sensitive to the presence of the fields but also to the incoming direction [134], making it possible to record complex pictures in the visible domain of the amplification region. The scheme of such an optical tracking detector is shown in **Figure 5**. The scheme shows the typical design of a projection chamber with the large drift volume where the primordial ionization is created along the path of the charged particles. A constant electric field guides the electrons towards the amplification region. Here, a strong electric field accelerates them and electron avalanches are produced. The fast electromagnetic signals in the GHz domain accompanying the electron avalanches can be efficiently transformed into the optical domain using Rydberg states [126]. The excitation of the gas atoms can be done using a two-level laser system tuned to the desired level with sub-ns time synchronization if systems designed for quantum experiments as the Sinara/ARTIQ open hardware and software are used [135, 136]. The highly excited Rydberg atoms in the amplification region also play a second role: their very low and adjustable ionization threshold ensures that electrons in the avalanche can easily ionize them, thus leading to an enhanced avalanche signal.

Alternatively, the concept can be extended to the ionization and drift volume region. If the sensitive volume of the detector can be brought to a highly-excited electronic state, then the immediate benefit is the decrease of the ionization threshold of the medium. This in turn can result in a higher effective ionization along the charged particle's trajectories, compensating the requirement of a very low gas density to avoid self-ionization (a consequence of the size of Rydberg atoms and their interactions). This would both allow and require a decrease of the current thickness of ~1% in radiation length [137] by 2 or 3 orders of magnitude, thus reducing the multiple-scattering and energy loss and allowing both the tracking of very low momentum particles and an increase in the momentum resolution and pointing accuracy.

The scheme of a Rydberg Tracking Chamber (RTC) is shown in **Figure 6**. The optical readout of the signals left by ionizing radiation can be done in 4 steps. In the first one, the low density gas is excited to a low-n Rydberg state with lifetime of 100 ns using a two-level laser system. This process happens synchronously with the expected beam arrival such that the products of the collision find the ionization medium in an excited state, which is the second step. As the excitation cross section dominates over the ionization [138], many low-n Rydberg states can be transferred to high-n states. In the third step, the low-n Rydberg atoms de-excite by emitting photons. However, those atoms that have been excited further have an extended lifetime which can reach μs. These highly excited states can be read out in the last step directly using the electromagnetic induced transparency [139, 140] or the atoms can be photo-detached by electric fields and the resulting supernumerary (with respect to standard ionization) electrons or ions can be detected by conventional means simultaneously with the electrons or ions generated along the same trajectory by standard ionization.

7 CONCLUSION

While most of the ideas proposed in this paper are very speculative, and much exploratory and developmental work will be required to ascertain their feasibility or usefulness in the context of high energy physics detectors, it is our hope that they could be seen as incentives for further exploration: the problems they address require solutions which these or other quantum sensors may be able to contribute to. It is also clear that

these proposed approaches may be neither the optimal, nor the only quantum sensing approaches to some of the challenges of future high energy physics detectors.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

All authors contributed to the paper, writing different sections and correcting the complete document. Specifically, EA, IF, and HH wrote sections on scintilators and nanodots; FB and GO wrote sections related to 2-dimensional materials; GK wrote sections on polarimetry and TPC's; MD wrote the remaining sections.

FUNDING

Research was funded by Warsaw University of Technology within the Excellence Initiative: Research University (IDUB) programme and the IDUB-POB-FWEiTE-1 project grant. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 801110. It reflects only the authors' view; the EU agency is not responsible for any use that may be made of the information it contains.

ACKNOWLEDGMENTS

The work on low dimensional materials for gaseous detectors is conducted within the activities of the CERN EP-DT-DD Gaseous Detector Development (GDD) team and FB and GO would like to thanks and recognise the contribution of the group of Camilla Coletti at Istituto Italiano di Tecnologia (IIT). EA would like to thank her colleagues of Crystal Clear Collaboration.

REFERENCES

1. Safronova M, Budker De. Focus on Quantum Sensors for New-Physics Discoveries. *Quan Sci. Technol.* (2022) 6:040401. Available at: https://iopscience.iop.org/journal/2058-9565/page/Focus-on-Quantum-Sensors-for-New-Physics-Discoveries. [Accessed June 10, 2022].
2. Irastorza IG, Redondo J. New Experimental Approaches in the Search for Axion-like Particles. *Prog Part Nucl Phys* (2018) 102:89–159. doi:10.1016/j.ppnp.2018.05.003
3. Backes KM, Palken DA, Kenany SA, Brubaker BM, Cahn SB, Droster A, et al. A Quantum Enhanced Search for Dark Matter Axions. *Nature* (2021) 590:238–42. doi:10.1038/s41586-021-03226-7
4. Geraci AA, Bradley C, Gao D, Weinstein J, Derevianko A. Searching for Ultralight Dark Matter with Optical Cavities. *Phys Rev Lett* (2019) 123(Jul):031304. doi:10.1103/PhysRevLett.123.031304
5. Dujardin C, Auffray E, Bourret-Courchesne E, Dorenbos P, Lecoq P, Nikl M, et al. Needs, Trends, and Advances in Inorganic Scintillators. *IEEE Trans Nucl Sci* (2018) 65(8):1977–97. doi:10.1109/tns.2018.2840160
6. Brunner SE, Schaart DR. BGO as a Hybrid Scintillator/Cherenkov Radiator for Cost-Effective Time-Of-Flight PET. *Phys Med Biol* (2017) 62:4421–39. doi:10.1088/1361-6560/aa6a49
7. Gundacker S, Martinez Turtos R, Kratochwil N, Pots RH, Paganoni M, Lecoq P, et al. Experimental Time Resolution Limits of Modern SiPM's and TOF-PET Detectors Exploring Different Scintillators and Cherenkov Emission. *Phys Med Biol* (2020) 65(2):025001. doi:10.1088/1361-6560/ab63b4
8. Kratochwil N, Gundacker S, Lecoq P, Auffray E. Pushing Cherenkov PET with BGO via Coincidence Time Resolution Classification and Correction. *Phys Med Biol* (2020) 65:115004. doi:10.1088/1361-6560/ab87f9
9. Van Eijk CWE. Cross-luminescence. *J Lumin* (1994) 60-61:936–41. doi:10.1016/0022-2313(94)90316-6
10. Pots E, Auffray R, Gundacker S Exploiting Cross-Luminescence in BaF_2 for Ultrafast Timing Applications Using Deep-Ultraviolet Sensitive hpk Silicon Photomultipliers. *Front Phys* (2020)8:4821. doi:10.3389/fphy.2020.592875
11. Gundacker S, Pots RH, Nepomnyashchikh A, Radzhabov E, Shendrik R, Omelkov S, et al. Vacuum Ultraviolet Silicon Photomultipliers Applied to BaF2 Cross-Luminescence Detection for High-Rate Ultrafast Timing Applications. *Phys Med Biol* (2021) 66:114002. doi:10.1088/1361-6560/abf476
12. Omelkov SI, Nagirnyi V, Vasil'ev AN, Kirm M. New Features of Hot Intraband Luminescence for Fast Timing. *J Lumin* (2016) 176:309–17. doi:10.1016/j.jlumin.2016.03.039
13. Létant SE, Wang T-F. Semiconductor Quantum Dot Scintillation under γ-Ray Irradiation. *Nano Lett* (2006) 6:2877–80. doi:10.1021/nl0620942
14. Turtos RM, Gundacker S, Polovitsyn A, Christodoulou S, Salomoni M, Auffray E, et al. Ultrafast Emission from Colloidal Nanocrystals under Pulsed X-ray Excitation. *J Inst* (2016) 11:P10015. doi:10.1088/1748-0221/11/10/p10015
15. Procházková L, Tomáš G, Václav C, Vítězslav J, Martin N. Fabrication of Highly Efficient ZnO Nanoscintillators. *Opt Mater* (2015) 47:67–71. doi:10.1016/j.optmat.2015.07.001
16. Hospodková A, Martin N, Oliva P, Jiří O, Petr B, Dalibor P, et al. InGaN/GaN Multiple Quantum Well for Fast Scintillation Application: Radioluminescence and Photoluminescence Study. *Nanotechnology* (2014) 25:455501. doi:10.1088/0957-4484/25/45/455501
17. Čuba KV, Brik MG, Mihóková E, Martinez Turtos R, Lecoq P, Auffray E, et al. On the Structure, Synthesis, and Characterization of Ultrafast Blue-Emitting CsPbBr3 Nanoplatelets. *APL Mater* (2019) 7:011104. doi:10.1063/1.5079300
18. Efros AL, Brus LE. Nanocrystal Quantum Dots: From Discovery to Modern Development. *ACS nano* (2021) 15:6192–210. doi:10.1021/acsnano.1c01399
19. Pietryga JM, Park YS, Lim J, Fidler AF, Bae WK, Brovelli S, et al. Spectroscopic and Device Aspects of Nanocrystal Quantum Dots. *Chem Rev* (2016) 116:10513–622. doi:10.1021/acs.chemrev.6b00169
20. Edvinsson T. Optical Quantum Confinement and Photocatalytic Properties in Two-, One- and Zero-Dimensional Nanostructures. *R Soc Open Sci* (2018) 5:180387. doi:10.1098/rsos.180387
21. Ooba H. Synthesis of Unique High Quality Fluorescence Quantum Dots for the Biochemical Measurements. *AIST TODAY* (2006) 6:26–7. Available at: https://www.aist.go.jp/Portals/0/resource_images/aist_e/aist_today/2006_21/pdf/2006_21_p22.pdf. [Accessed June 10, 2022].
22. Fontes A. Quantum Dots in Biomedical Researchs. In: *Biomedical Engineering - Technical Applications in Medicine*. London, United Kingdom: IntechOpen. (2012). Available at: https://www.intechopen.com/chapters/38785. [Accessed June 10, 2022].
23. Klimov VI. Spectral and Dynamical Properties of Multiexcitons in Semiconductor Nanocrystals. *Annu Rev Phys Chem* (2007) 58:635–73. doi:10.1146/annurev.physchem.58.032806.104537
24. Klimov VI. Multicarrier Interactions in Semiconductor Nanocrystals in Relation to the Phenomena of Auger Recombination and Carrier Multiplication. *Annu Rev Condens Matter Phys* (2014) 5:285–316. doi:10.1146/annurev-conmatphys-031113-133900
25. Korzh B, Zhao Q, Allmaras J, Simone F, Travis MA, Eric AB, et al. Demonstration of Sub-3 Ps Temporal Resolution with a Superconducting Nanowire Single-Photon Detector. *Nat Photon* (2020) 14:250–5.

26. Onaru K, Hara K, Harada D, Wada S, Nakamura K, Unno Y. Study of Time Resolution of Low-Gain Avalanche Detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2021) 985: 164664. doi:10.1016/j.nima.2020.164664
27. Grim JQ, Christodoulou S, Di Stasio F, Krahne R, Cingolani R, Manna L, et al. Continuous-wave Biexciton Lasing at Room Temperature Using Solution-Processed Quantum wells. *Nat Nanotech* (2014) 9:891–5. doi:10.1038/nnano.2014.213
28. Christodoulou S, Vaccaro G, Pinchetti V, De Donato F, Grim JQ, Casu A, et al. Synthesis of Highly Luminescent Wurtzite CdSe/CdS Giant-Shell Nanocrystals Using a Fast Continuous Injection Route. *J Mater Chem C* (2014) 2:3439. doi:10.1039/c4tc00280f
29. Meng Z, Mahler B, Houel J, Kulzer F, Ledoux G, Vasil'ev A, et al. Perspectives for CdSe/CdS Spherical Quantum wells as Rapid-Response Nano-Scintillators. *Nanoscale* (2021) 13:19578–86. doi:10.1039/d1nr04781g
30. Liu C, Li Z, Hajagos TJ, Kishpaugh D, Chen DY, Pei Q. Transparent Ultra-high-loading Quantum Dot/Polymer Nanocomposite Monolith for Gamma Scintillation. *ACS nano* (2017) 11:6422–30. doi:10.1021/acsnano.7b02923
31. Procházková L, Cuba V, Mrazek J, Beitlerova A, Jary V, Nikl M. Preparation of Zn(Cd)O:Ga–SiO2 Composite Scintillating Materials. *Radiat Meas* (2016) 90:59–63. doi:10.1016/j.radmeas.2015.12.040
32. Buresova H, Lenka P, Rosana MT, Vítězslav J, Eva M, Alena B, et al. Preparation and Luminescence Properties of ZnO:Ga – Polystyrene Composite Scintillator. *Opt Express* (2016) 24:15289–98. doi:10.1364/OE.24.015289
33. Turtos RM, Gundacker S, Lucchini MT, Procházková L, Čuba V, Burešová H, et al. Timing Performance of ZnO:Ga Nanopowder Composite Scintillators. *Phys Status Solidi RRL* (2016) 10:843–7. doi:10.1002/pssr.201600288
34. Hubáček T, Hospodková A, Kuldová K, Oswald J, Pangrác J, Jarý V, et al. Advancement toward Ultra-thick and Bright InGaN/GaN Structures with a High Number of QWs. *CrystEngComm* (2019) 21:356–62. doi:10.1039/C8CE01830H Available at: https://pubs.rsc.org/en/content/articlelanding/2019/ce/c8ce01830h/unauth. [Accessed June 10, 2022].
35. Tocci G, Leonida GA, Petra K, Baffigi F, Lorenzo F, Luca L, et al. InGaN/GaN Multiple Quantum Well for Superfast Scintillation Application: Photoluminescence Measurements of the Picosecond Rise Time and Excitation Density Effect. *J Lumin* (2019) 208:119–24. doi:10.1016/j.jlumin.2018.12.034
36. Jary V, Hospodkova A, Hubacek T, Hajek F, Blazek K, Nikl M. Optical Properties of InGaN/GaN Multiple Quantum Well Structures Grown on GaN and Sapphire Substrates. *IEEE Trans Nucl Sci* (2020) 67:974–7. doi:10.1109/tns.2020.2985666
37. Decka K, Jan K, František H, Petr P, Vladimir B, Eva M. Scintillation Response Enhancement in Nanocrystalline lead Halide Perovskite Thin Films on Scintillating Wafers. *Nanomaterials* (2022) 12(1):14. doi:10.3390/nano12010014
38. Gandini M, Villa I, Beretta M, Gotti C, Imran M, Carulli F, et al. Efficient, Fast and Reabsorption-free Perovskite Nanocrystal-Based Sensitized Plastic Scintillators. *Nat Nanotechnol* (2020) 15:462–8. doi:10.1038/s41565-020-0683-8
39. Silicon. High Transmission Silicon for Infrared Optical Applications (2019). Available at: https://www.sil-tronix-st.com/en/news/High-transmission-silicon-for-infrared-optical-applications. [Accessed June 10, 2022].
40. Lecoq P, Metamaterials for Novel X- or γ-ray Detector Designs, 2008 IEEE Nuclear Science Symposium Conference Record. (2008) 1405–9.
41. Lecoq P. TICAL: 4D Total absorptionTime Imaging CALorimeter, ERC grant 338953 (TICAL) (2014). Available at: https://cordis.europa.eu/project/id/338953/fr. [Accessed June 10, 2022].
42. Turtos RM, Gundacker S, Auffray E, Lecoq P. Towards a Metamaterial Approach for Fast Timing in PET: Experimental Proof-Of-Concept. *Phys Med Biol* (2019) 64(18):185018. doi:10.1088/1361-6560/ab18b3
43. Turtos R, Gundacker S, Omelkov S. On the Use of Cdse Scintillating Nanoplatelets as Time Taggers for High-Energy Gamma Detection. *npj 2D Mater Appl* (2019) 3(1):185018. doi:10.1038/s41699-019-0120-8
44. Tomanová K, Suchá A, Mihóková E, Procházková L, Jakubec I, Turtos RM, et al. CsPbBr3 Thin Films on LYSO:Ce Substrates. *IEEE TRANSACTIONS NUCLEAR SCIENCE* (2020) 67(6):933–7. doi:10.1109/TNS.2020.2978581 Available at: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9025189. [Accessed June 10, 2022].
45. Oktyabrsky S, Yakimov M, Tokranov V, Murat P. Integrated Semiconductor Quantum Dot Scintillation Detector: Ultimate Limit for Speed and Light Yield. *IEEE Trans Nucl Sci* (2016) 63:656–63. doi:10.1109/tns.2015.2502426
46. Dropiewski K, Minns A, Yakimov M, Tokranov V, Murat P, Oktyabrsky S. Optical Properties of InAs Quantum Dots/GaAs Waveguides for Ultra-fast Scintillators. *J Lumin* (2020) 220:116952. doi:10.1016/j.jlumin.2019.116952
47. Tam AK, Boyraz O, Unangst J, Nazareta P, Schreuder M, Nilsson M. Quantum-dot Doped Polymeric Scintillation Material for Radiation Detection. *Radiat Measurements* (2018) 111:27–34. doi:10.1016/j.radmeas.2018.02.008
48. Delage M-È, Lecavalier M-È, Larivière D, Nì Allen C, Beaulieu L. Characterization of a Binary System Composed of Luminescent Quantum Dots for Liquid Scintillation. *Phys Med Biol* (2018) 63(17):175012. doi:10.1088/1361-6560/aad9c3
49. Delage MÈ, Lecavalier MÈ, Larivière D, Allen CN, Beaulieu L. Dosimetric Properties of Colloidal Quantum Dot-Based Systems for Scintillation Dosimetry. *Phys Med Biol* (2018) 64:095027. doi:10.1088/1361-6560/ab109b
50. Beretta M, Amirkhani A, Brofferio C, Brovelli S, Buonanno L. The ESQUIRE Project: Quantum Dots as Scintillation Detectors. *IL NUOVO CIMENTO* (2019) 42C:188. doi:10.1393/ncc/i2019-19188-4 Available at: https://www.sif.it/riviste/sif/ncc/econtents/2019/042/04/article/43. [Accessed June 10, 2022]
51. AIDAinnova. AIDAinnova Project (2021). Available at: https://aidainnova.web.cern.ch/newsletter. [Accessed June 10, 2022].
52. Moulson M, NanoCal, Blue Sky Project AIDAinnova, Private Communication.
53. Liu J, Geng Y, Li D, Yao H, Huo Z, Li Y, et al. Deep Red Emissive Carbonized Polymer Dots with Unprecedented Narrow Full Width at Half Maximum. *Adv Mater* (2020) 32:1906641. doi:10.1002/adma.201906641
54. Yuan F, Yuan T, Sui L, Wang Z, Xi Z, Li Y, et al. Engineering Triangular Carbon Quantum Dots with Unprecedented Narrow Bandwidth Emission for Multicolored LEDs. *Nat Commun* (2018) 9:2249. doi:10.1038/s41467-018-04635-5
55. Zhang C, Wan Q, Ono LK, Liu Y, Zheng W, Zhang Q, et al. Narrow-Band Violet-Light-Emitting Diodes Based on Stable Cesium Lead Chloride Perovskite Nanocrystals. *ACS Energ Lett.* (2021) 6:3545–54. doi:10.1021/acsenergylett.1c01380
56. Cheng R, Zou C-L, Guo X, Wang S, Han X, Tang HX. Broadband On-Chip Single-Photon Spectrometer. *Nat Commun* (2019) 10:4104. doi:10.1038/s41467-019-12149-x
57. Miao J, Zhang F. Recent Progress on Highly Sensitive Perovskite Photodetectors. *J Mater Chem C* (2019) 7:1741–91. doi:10.1039/C8TC06089D
58. Zhou J, Luo J, Rong X, Wei P, Molokeev MS, Huang Y, et al. Lead-Free Perovskite Derivative Cs 2 SnCl 6− X Br X Single Crystals for Narrowband Photodetectors. *Adv Opt Mater* (2019) 7:1900139. doi:10.1002/adom.201900139
59. Iguaz F, Bortfeldt J, Brunbauer F, David C, Desforge D, Fanourakis G, et al. Charged Particle Timing at Sub-25 Picosecond Precision: The PICOSEC Detection Concept. *Nucl Instruments Methods Phy. Res Sect A* (2019) 936: 515–18.
60. Polyakov A, Senft C, Thompson KF, Feng J, Cabrini S, Schuck PJ, et al. Plasmon-enhanced Photocathode for High Brightness and High Repetition Rate X-ray Sources. *Phys Rev Lett* (2013) 110(Feb):076802. doi:10.1103/PhysRevLett.110.076802
61. Green ME, Bas DA, Yao HY, Gengler JJ, Headrick RJ, Back TC, et al. Bright and Ultrafast Photoelectron Emission from Aligned Single-Wall Carbon Nanotubes through Multiphoton Exciton Resonance. *Nano Lett* (2019) 19: 158–64. doi:10.1021/acs.nanolett.8b03564
62. Choi HJ, Raghunathan A, Groves TR. Size Dependent Enhancement of Photoelectron Emission Quantum Efficiencies from Magnesium Dots. *J Vac Sci Technol B, Nanotechnol Microelectron Mater Process Meas Phenom* (2012) 30:06F605. doi:10.1116/1.4766883
63. Naghdi S, Sanchez-Arriaga G, Rhee KY. Tuning the Work Function of Graphene toward Application as Anode and Cathode. *J Alloys Compounds* (2019) 805:1117–34. doi:10.1016/j.jallcom.2019.07.187

64. Lin X, Easo S, Shen Y, Chen H, Zhang B, Joannopoulos JD, et al. *Controlling Cherenkov Angles with Resonance Transition Radiation". Nat Phys* (2018) 14: 816–21. doi:10.1038/s41567-018-0138-4
65. Lin X, Hu H, Easo S, Yang Y, Shen Y, Yin K, et al. A Brewster Route to Cherenkov Detectors. *Nat Commun* (2021) 12. arXiv:2107. doi:10.1038/s41467-021-25822-x
66. Wang G, Yang P, Moody NA, Batista ER. Overcoming the Quantum Efficiency-Lifetime Tradeoff of Photocathodes by Coating with Atomically Thin Two-Dimensional Nanomaterials. *Npj 2d Mater Appl* (2018) 2:17. doi:10.1038/s41699-018-0062-6
67. Wang G, Yang P, Batista ER. Computational Screening of Two-Dimensional Coatings for Semiconducting Photocathodes. *Phys Rev Mater* (2020) 4: 024001. doi:10.1103/physrevmaterials.4.024001
68. Yamaguchi H, Liu F, DeFazio J, Narvaez Villarrubia CW, Finkenstadt D, Shabaev A, et al. Active Bialkali Photocathodes on Free-Standing Graphene Substrates. *Npj 2d Mater Appl* (2017) 1:12. doi:10.1038/s41699-017-0014-6
69. Liu F, Guo L, DeFazio J, Pavlenko V, Yamamoto M, Moody NA, et al. Photoemission from Bialkali Photocathodes through an Atomically Thin Protection Layer. *ACS Appl Mater Inter* (2022) 14(1):1710–7. PMID: 34935342. doi:10.1021/acsami.1c19393
70. Liu F, Moody NA, Jensen KL, Pavlenko V, Narvaez Villarrubia CW, Mohite AD, et al. Single Layer Graphene Protective Gas Barrier for Copper Photocathodes. *Appl Phys Lett* (2017) 110(4):041607. doi:10.1063/1.4974738
71. Franchino S, Gonzalez-Diaz D, Hall-Wilton R, Jackman RB, Muller H, Nguyen TT, et al. Charge Transfer Properties through Graphene for Applications in Gaseous Detectors. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2016) 824:571–4. doi:10.1016/j.nima.2015.11.077
72. Berry V. Impermeability of Graphene and its Applications. *Carbon* (2013) 62: 1–10. doi:10.1016/j.carbon.2013.05.052
73. Lee C, Wei X, Kysar JW, Hone J. Measurement of the Elastic Properties and Intrinsic Strength of Monolayer Graphene. *Science* (2008) 321:385–8. doi:10.1126/science.1157996
74. Sinha S, Warner JH. Recent Progress in Using Graphene as an Ultrathin Transparent Support for Transmission Electron Microscopy. *Small Structures* (2021) 2(4):2000049. doi:10.1002/sstr.202000049
75. Meyer JC, Geim AK, Katsnelson MI, Novoselov KS, Booth TJ, Roth S. The Structure of Suspended Graphene Sheets. *Nature* (2007) 446:60–3. doi:10.1038/nature05545
76. Hassink G, Wanke R, Rastegar I, Braun W, Stephanos C, Herlinger P, et al. Transparency of Graphene for Low-Energy Electrons Measured in a Vacuum-Triode Setup. *APL Mater* (2015) 3(7):076106. doi:10.1063/1.4927406
77. Il'ichev EA, Kuleshov AE, Migunov DM, Nabiev RM, Petrukhin GN, Rychkov GS, et al. Studying the Transparency of Graphene for Low-Energy Electrons. *Tech Phys Lett* (2018) 44:848–51. doi:10.1134/S1063785018090201
78. Srisonphan S, Kim M, Kim HK. Space Charge Neutralization by Electron-Transparent Suspended Graphene. *Sci Rep* (2014) 4:3764. doi:10.1038/srep03764
79. Huang S, Dakhchoune M, Luo W, Oveisi E, He G, Rezaei M, et al. Single-layer Graphene Membranes by Crack-free Transfer for Gas Mixture Separation. *Nat Commun* (2018) 9:2632. doi:10.1038/s41467-018-04904-3
80. Sun PZ, Yang Q, Kuang WJ, Stebunov YV, Xiong WQ, Yu J, et al. Limits on Gas Impermeability of Graphene. *Nature* (2020) 579:229–32. doi:10.1038/s41586-020-2070-x
81. Sun PZ, Yagmurcukardes M, Zhang R, Kuang WJ, Lozada-Hidalgo M, Liu BL, et al. Exponentially Selective Molecular Sieving through Angstrom Pores. *Nat Commun* (2021) 12:7170. doi:10.1038/s41467-021-27347-9
82. Bacchetta N, Collins P, Riedler P. Tracking and Vertex Detectors at FCC-Ee. *Eur Phys J Plus* (2022) 137:231. doi:10.1140/epjp/s13360-021-02323-w
83. Minns A, Dropiewski K, Yakimov M, Tokranov V, Hedges M, Murat P, et al. Parameters of Fast and High-Yield InAs/GaAs Quantum Dot Semiconductor Scintillator. *MRS Adv* (2021) 6:297–302. doi:10.1557/s43580-021-00019-y
84. Camilli L, Jørgensen JH, Tersoff J, Stoot AC, Balog R, Cassidy A, et al. Self-assembly of Ordered Graphene Nanodot Arrays. *Nat Commun* (2017) 8:47. doi:10.1038/s41467-017-00042-4
85. Zheng M, Yu M, Liu Y, Skomski R, Liou SH, Sellmyer DJ, et al. Magnetic Nanodot Arrays Produced by Direct Laser Interference Lithography. *Appl Phys Lett* (2001) 79:2606–8. doi:10.1063/1.1409948
86. Maximov MV, Nadtochiy AM, Mintairov SA, Kalyuzhnyy NA, Kryzhanovskaya NV, Moiseev EI, et al. Light Emitting Devices Based on Quantum Well-Dots. *Appl Sci* (2020) 10(3):1038. doi:10.3390/app10031038
87. Maximov MV, Nadtochiy AM, Mintairov SA, Kalyuzhnyy NA, Kryzhanovskaya NV, Moiseev EI, et al. Light Emitting Devices Based on Quantum Well-Dots. *Appl Sci* (2020) 10(3). doi:10.3390/app10031038
88. Knapitsch A, Auffray E, Fabjan CW, Leclercq J-L, Letartre X, Mazurczyk R, et al. Results of Photonic crystal Enhanced Light Extraction on Heavy Inorganic Scintillators. *IEEE Trans Nucl Sci* (2012) 59(5):2334–9. doi:10.1109/tns.2012.2184556
89. Roques-Carmes C, Rivera N, Ghorashi A, Kooi SE, Yang Y, Lin Z, et al. *A General Framework for Scintillation in Nanophotonics* (2021). arXiv preprint arXiv:2110.11492.
90. Canham L. Introductory Lecture: Origins and Applications of Efficient Visible Photoluminescence from Silicon-Based Nanostructures. *Faraday Discuss* (2020) 222:10–81. doi:10.1039/d0fd00018c
91. Faist J, Capasso F, Sivco DL, Sirtori C, Hutchinson AL, Cho AY. Quantum cascade Laser. *Science* (1994) 264:553–6. doi:10.1126/science.264.5158.553
92. Knapitsch A, Lecoq P. Review on Photonic crystal Coatings for Scintillators. *Int J Mod Phys A* (2014) 29:1430070. doi:10.1142/S0217751X14300701
93. Salomoni M, Pots R, Auffray E, Lecoq P. Enhancing Light Extraction of Inorganic Scintillators Using Photonic Crystals. *Crystals* (2018) 8(2). doi:10.3390/cryst8020078
94. Kurman Y, Shultzman A, Segal O, Pick A, Kaminer I. Photonic-Crystal Scintillators: Molding the Flow of Light to Enhance X-Ray and γ-Ray Detection. *Phys Rev Lett* (2020) 125(Jul):040801. doi:10.1103/PhysRevLett.125.040801
95. Liu B, Ouyang X. Optical Improvements of Plastic Scintillators by Nanophotonics. *Plast Scintillators* (2021) 140. doi:10.1007/978-3-030-73488-6_9
96. García-Lojo D, Núñez Sánchez S, Gómez-Graña S, Marek Grzelczak M, Pastoriza-Santos I, Pérez-Juste J, et al. Plasmonic Supercrystals. *Acc Chem Res* (2019) 52(7):1855–64. doi:10.1021/acs.accounts.9b00213
97. Günay M, Chuang Y-L, Tasgin ME. Continuously-tunable Cherenkov-Radiation-Based Detectors via Plasmon index Control. *Nanophotonics* (2020) 9(6):1479–89. doi:10.1515/nanoph-2020-0081
98. Lin X, Hu H, Easo S, Yang Y, Shen Y, Yin K, et al. A Brewster Route to Cherenkov Detectors. *Nat Commun* (2021) 12:5554. doi:10.103810.1038/s41467-021-25822-x
99. Hu H, Lin X, Wong LJ, Yang Q, Liu D, Zhang B, et al. Surface Dyakonov-Cherenkov Radiation. *eLight* (2022) 2:2. doi:10.1186/s43593-021-00009-5
100. Jelezko F, Gaebel T, Popa I, Domhan M, Gruber A, Wrachtrup J. Observation of Coherent Oscillation of a Single Nuclear Spin and Realization of a Two-Qubit Conditional Quantum Gate. *Phys Rev Lett* (2004) 93:130501. doi:10.1103/PhysRevLett.93.130501
101. Jelezko F, Gaebel T, Popa I, Gruber A, Wrachtrup J. Observation of Coherent Oscillations in a Single Electron Spin. *Phys Rev Lett* (2004) 92:076401. doi:10.1103/PhysRevLett.92.076401
102. Álvarez GA, Bretschneider CO, Fischer R, London P, Kanda H, Onoda S, et al. Local And Bulk ^{13}C Hyperpolarization in Nitrogen-Vacancy-Centred Diamonds at Variable fields and Orientations. *Nat Commun* (2015) 6(Sept):8456. arXiv:1412.8635 [quant-ph].
103. Schirhagl R, Chang K, Loretz M, Degen CL. Nitrogen-Vacancy Centers in Diamond: Nanoscale Sensors for Physics and Biology. *Annu Rev Phys Chem* (2014) 65(1):83–105. PMID: 24274702. doi:10.1146/annurev-physchem-040513-103659
104. Dutt MVG, Childress L, Jiang L, Togan E, Maze J, Jelezko F, et al. Quantum Register Based on Individual Electronic and Nuclear Spin Qubits in Diamond. *Science* (2007) 316:1312–6. doi:10.1126/science.1139831
105. Dolde F, Jakobi I, Naydenov B, Zhao N, Pezzagna S, Trautmann C, et al. Room-temperature Entanglement between Single Defect Spins in diamond. *Nat Phys* (2013) 9(Feb):139–43. doi:10.1038/nphys2545
106. Jelezko F, Wrachtrup J. Single Defect Centres in diamond: A Review. *Phys Stat Sol (A)* (2006) 203(13):3207–25. doi:10.1002/pssa.200671403

107. Budker D, Romalis M. Optical Magnetometry. *Nat Phys* (2007) 3(Apr): 227–34. doi:10.1038/nphys566
108. Müller C, Kong X, Cai J-M, Melentijević K, Stacey AD, Markham ML, et al. Nuclear Magnetic Resonance Spectroscopy with Single Spin Sensitivity. *Nat Commun* (2014) 5. doi:10.1021/acs.accounts.9b00213
109. Doherty MW, Struzhkin VV, Simpson DA, McGuinness LP, Meng Y, Stacey A, et al. Electronic Properties and Metrology Applications of the DiamondNV−Center under Pressure. *Phys Rev Lett* (2014) 112: doi:10.1103/PhysRevLett.112.047601
110. Acosta VM, Bauch E, Ledbetter MP, Waxman A, Bouchard LS, Budker D. Temperature Dependence of the Nitrogen-Vacancy Magnetic Resonance in Diamond. *Phys Rev Lett* (2010) 104(Feb):070801. doi:10.1103/PhysRevLett.104.070801
111. Doherty MW, Manson NB, Delaney P, Jelezko F, Wrachtrup J, Hollenberg LCL. The Nitrogen-Vacancy Colour centre in diamond. *Phys Rep* (2013) 528(Jul):1–45. doi:10.1016/j.physrep.2013.02.001
112. Loubser JHN, Wyk JAv. Electron Spin Resonance in the Study of diamond. *Rep Prog Phys* (1978) 41(aug):1201–48. doi:10.1088/0034-4885/41/8/002
113. Sushkov AO, Lovchinsky I, Chisholm N, Walsworth RL, Park H, Lukin MD. Magnetic Resonance Detection of Individual Proton Spins Using Quantum Reporters. *Phys Rev Lett* (2014) 113(Nov):197601. doi:10.1103/PhysRevLett.113.197601
114. Bunce G, Saito N, Soffer J, Vogelsang W. Prospects for Spin Physics at RHIC. *Annu Rev Nucl Part Sci* (2000) 50:525–75. arXiv:hep-ph/0007218. doi:10.1146/annurev.nucl.50.1.525
115. Crabb DG, Meyer W. Solid Polarized Targets for Nuclear and Particle Physics Experiments. *Annu Rev Nucl Part Sci* (1997) 47:67–109. doi:10.1146/annurev.nucl.47.1.67
116. Kharzeev DE, Liao J, Voloshin SA, Wang G. Chiral Magnetic and Vortical Effects in High-Energy Nuclear Collisions-A Status Report. *Prog Part Nucl Phys* (2016) 88:1–28. arXiv:1511.04050 [hep-ph]. doi:10.1016/j.ppnp.2016.01.001
117. Rondin L, Tetienne J-P, Hingant T, Roch J-F, Maletinsky P, Jacques V. Magnetometry with Nitrogen-Vacancy Defects in diamond. *Rep Prog Phys* (2014) 77(May):056503. doi:10.1088/0034-4885/77/5/056503
118. Tallaire A, Brinza O, Huillery P, Delord T, Pellet-Mary C, Staacke R, et al. High NV Density in a Pink CVD diamond Grown with N2O Addition. *Carbon* (2020) 170:421–9. doi:10.1016/j.carbon.2020.08.048
119. Kurita T, Shimotsuma Y, Fujiwara M, Fujie M, Mizuochi N, Shimizu M, et al. Direct Writing of High-Density Nitrogen-Vacancy Centers inside diamond by Femtosecond Laser Irradiation. *Appl Phys Lett* (2021) 118(21):214001. doi:10.1063/5.0049953
120. Pietraszko J, Galatyuk T, Grilj V, Koenig W, Spataro S, Träger M. Radiation Damage in Single crystal CVD diamond Material Investigated with a High Current Relativistic 197Au Beam. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2014) 763:1–5. doi:10.1016/j.nima.2014.06.006
121. Sasaki K, Watanabe H, Sumiya H, Itoh KM, Abe E. Detection and Control of Single Proton Spins in a Thin Layer of diamond Grown by Chemical Vapor Deposition. *Appl Phys Lett* (2020) 117(11):114002. doi:10.1063/5.0016196
122. Adams CS, Pritchard JD, Shaffer JP. Rydberg Atom Quantum Technologies. *J Phys B: Mol Opt Phys* (2019) 53:012002. doi:10.1088/1361-6455/ab52ef
123. Wu X, Liang X, Tian Y, Yang F, Chen C, Liu Y-C, et al. A Concise Review of Rydberg Atom Based Quantum Computation and Quantum Simulation*. *Chin Phys. B* (2021) 30:020305. doi:10.1088/1674-1056/abd76f
124. Schmidt J, Münzenmaier Y, Kaspar P, Schalberger P, Baur H, Löw R, et al. An Optogalvanic Gas Sensor Based on Rydberg Excitations. *J Phys B: Mol Opt Phys* (2020) 53:094001. doi:10.1088/1361-6455/ab728e
125. Osterwalder A, Merkt F. Using High Rydberg States as Electric Field Sensors. *Phys Rev Lett* (1999) 82(Mar):1831–4. doi:10.1103/physrevlett.82.1831
126. Sedlacek JA, Schwettmann A, Kübler H, Löw R, Pfau T, Shaffer JP. Microwave Electrometry with Rydberg Atoms in a Vapour Cell Using Bright Atomic Resonances. *Nat Phys* (2012) 8(Nov):819–24. doi:10.1038/nphys2423
127. Wade CG, Šibalić N, de Melo NR, Kondo JM, Adams CS, Weatherill KJ. Real-time Near-Field Terahertz Imaging with Atomic Optical Fluorescence. *Nat Photon* (2016) 11(Nov):40–3. doi:10.1038/nphoton.2016.214
128. Gallagher TF. Rydberg Atoms. In: *Cambridge Monographs on Atomic, Molecular and Chemical Physics*. Cambridge University Press (1994).
129. Adolfsson J, Ahmed M, Aiola S, Alme J, Alt T, Amend W. The Upgrade of the ALICE TPC with GEMs and Continuous Readout. *JINST* (2021) 1603: P03022. arXiv:2012.09518 [physics.ins-det]. doi:10.1088/1748-0221/16/03/P03022
130. Charpak G, Fabre J-P, Sauli F, Suzuki M, Dominik W. AN OPTICAL, PROPORTIONAL, CONTINUOUSLY OPERATING AVALANCHE CHAMBER. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1987) 258:177–84. doi:10.1016/0168-9002(87)90054-4
131. Fonte P, Breskin A, Charpak G, Dominik W, Sauli F. Beam Test of an Imaging High-Density Projection Chamber. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1989) 283:658–64. doi:10.1016/0168-9002(89)91436-8
132. Gai M, Ahmed MW, Stave SC, Zimmerman WR, Breskin A, Bromberger B, et al. An Optical Readout TPC (O-TPC) for Studies in Nuclear Astrophysics with Gamma-ray Beams at HIγS1. *J Inst* (2010) 5:P12004. arXiv:1101.1940 [physics.ins-det]. doi:10.1088/1748-0221/5/12/p12004
133. Phan NS, Lauer RJ, Lee ER, Loomba D, Matthews JAJ, Miller EH. GEM-based TPC with CCD Imaging for Directional Dark Matter Detection. *Astroparticle Phys* (2016) 84:82–96. arXiv:1510.02170 [physics.ins-det]. doi:10.1016/j.astropartphys.2016.08.006
134. Robinson AK, Prajapati N, Senic D, Simons MT, Holloway CL. Determining the Angle-Of-Arrival of a Radio-Frequency Source with a Rydberg Atom-Based Sensor. *Appl Phys Lett* (2021) 118(11):114001. doi:10.1063/5.0045601
135. Kasprowicz G, Kulik P, Gaska M, Przywozki T, Pozniak K, Jarosinski J, et al. ARTIQ and Sinara: Open Software and Hardware Stacks for Quantum Physics. In: *2020. OSA Quantum 2.0 Conference* (2020). QTu8B.14. doi:10.1364/quantum.2020.qtu8b.14
136. Bourdeauducq S, Jördens R, Zotov P, Britton J, Slichter D, Leibrandt D, et al. *ARTIQ 1.0* (2016). doi:10.5281/zenodo.51303
137. Anderson M, Berkovitz J, Betts W, Bossingham R, Bieser F, Brown R, et al. The Star Time Projection Chamber: A Unique Tool for Studying High Multiplicity Events at RHIC. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2003) 499:659–78. arXiv:nucl-ex/0301015. doi:10.1016/s0168-9002(02)01964-2
138. Vrinceanu D. Electron Impact Ionization of Rydberg Atoms. *Phys Rev A* (2005) 72:022722. doi:10.1103/physreva.72.022722
139. Mohapatra AK, Jackson TR, Adams CS. Coherent Optical Detection of Highly Excited Rydberg States Using Electromagnetically Induced Transparency. *Phys Rev Lett* (2007) 98: doi:10.1103/PhysRevLett.98.113003
140. Mauger S, Millen J, Jones MPA. Spectroscopy of Strontium Rydberg States Using Electromagnetically Induced Transparency. *J Phys B: Mol Opt Phys* (2007) 40:F319–F325. doi:10.1088/0953-4075/40/22/f03

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Doser, Auffray, Brunbauer, Frank, Hillemanns, Orlandini and Kornakov. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

REVIEW
published: 27 June 2022
doi: 10.3389/fphy.2022.920520

Challenges of Future Accelerators for Particle Physics Research

Stephen Gourlay[1]*, *Tor Raubenheimer*[2] and *Vladimir Shiltsev*[3]

[1]Lawrence Berkeley National Laboratory (Retired), Berkeley, CA, United States, [2]SLAC National Accelerator Laboratory, Menlo Park, CA, United States, [3]Fermi National Accelerator Laboratory, Batavia, IL, United States

For over half a century, high-energy particle accelerators have been a major enabling technology for particle and nuclear physics research as well as sources of X-rays for photon science research in material science, chemistry and biology. Particle accelerators for energy and intensity Frontier research in particle and nuclear physics continuously push the accelerator community to invent ways to increase the energy and improve the performance of accelerators, reduce their cost, and make them more power efficient. The accelerator community has demonstrated imagination and creativity in developing a plethora of future accelerator ideas and proposals. The technical maturity of the proposed facilities ranges from shovel-ready to those that are still largely conceptual. At this time, over 100 contributed papers have been submitted to the *Accelerator Frontier* of the US particle physics decadal community planning exercise known as *Snowmass'2021*. These papers cover a broad spectrum of topics: beam physics and accelerator education, accelerators for neutrinos, colliders for Electroweak/Higgs studies and multi-TeV energies, accelerators for *Physics Beyond Colliders* and rare processes, advanced accelerator concepts, and accelerator technology for Radio Frequency cavities (RF), magnets, targets and sources. This paper provides an overview of the present state of accelerators for particle physics and gives a brief description of some of the major facilities that have been proposed, their perceived advantages and some of the remaining challenges.

Keywords: accelerators, colliders, beam physics, magnets, particle sources, RF acceleration, plasma acceleration

OPEN ACCESS

Edited by:
Patrick Le Du, Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), France

Reviewed by:
Michelangelo Mangano, European Organization for Nuclear Research (CERN), Switzerland
Marco Incagli, Istituto Nazionale di Fisica Nucleare—sezione di Pisa, Italy

***Correspondence:**
Stephen Gourlay
sagourlay@lbl.gov

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *14 April 2022*
Accepted: *27 May 2022*
Published: *27 June 2022*

Citation:
Gourlay S, Raubenheimer T and Shiltsev V (2022) Challenges of Future Accelerators for Particle Physics Research. Front. Phys. 10:920520. doi: 10.3389/fphy.2022.920520

1 INTRODUCTION

There are more than 30,000 particle accelerators in operation around the world. Most of them use low energy beams ($\ll$ 1 MeV) and are used in industry [1]. Less numerous are high energy accelerators for research where there are just over a hundred, many of which have been recently constructed for photon science. Particle physics requires pushing the accelerated beams to the highest possible energies [2] and to the highest possible intensities [3]. Below, we briefly consider the most actively developing accelerator projects, such as the high energy lepton and hadron colliders and accelerators for neutrino studies and rare processes searches, and outline the status and progress in accelerator beam physics as well as in the core accelerator technologies - magnets, Radio Frequency cavities (RF), plasma, targets and sources. There are about two dozen energy Frontier colliders that complement or exceed the LHC in their discovery potential. Among them is the three TeV center-of-mass (CoM) CLIC option (100 MV/m accelerating gradient, 50 km long), two 100 km circumference *pp* colliders: the SPPC in China (75 TeV CoM, based on 12 T Superconducting (SC) magnets) and the FCC-hh at CERN (100 TeV CoM, 16 T SC

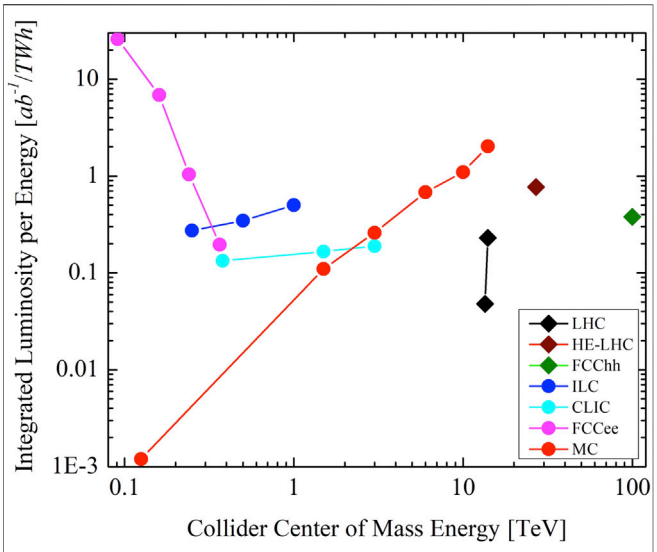


FIGURE 1 | Energy efficiency of present and future colliders. Annual integrated luminosity per Terawatt-hour of electric power consumption as a function of the centre-of-mass energy. The LHC—both present and expected after its high-luminosity upgrade (black diamonds) — is contrasted with a variety of proposed particle colliders: the Muon Collider (MC, red circles), the Future Circular electron—positron Collider (FCC-ee, magenta circles) assuming experiments at two collision points, the International Linear Collider (ILC, blue circles), the Compact Linear Collider (CLIC, cyan circles), the High Energy LHC (HE-LHC, magenta diamonds), and the Future Circular proton–proton Collider (FCC-hh, green diamonds). For lepton and hadron colliders operating at the same COM energy, the energy reach of the former is typically larger, by factors that depend, process by process, on the nature of the produced particles and their interactions.

magnets) [4], and a 10–14 TeV CoM $\mu^+\mu^-$ collider (10–14 km circumference, 16 T SC magnets) [5].

2 ENERGY FRONTIER COLLIDERS

2.1 Lepton Colliders

Understanding the unification of electromagnetism with the weak interactions is one of the cornerstones of the Standard Model of particle physics and the Higgs is the mechanism that breaks the electroweak symmetry and gives particles their mass. Thus, understanding the electroweak/Higgs sector is a major focus of particle physics. At present, there are as many as eight Higgs/ElectroWeak factories under consideration: e^+e^- colliders such as the CepC in China and FCC-ee at CERN, both about 100 km circumference, which require O (100 MW) RF systems to sustain high luminosity [6]; or an 11 km long CLIC (CERN) two-beam normal-conducting RF linear accelerator with an average gradient of 72 MV/m [7]; or the 21 km long International Linear Collider (ILC) based on super-conducting RF (SRF) linacs with an average gradient of 31.5 MV/m [8]. Besides technical feasibility and the cost, the most critical requirements for a post-LHC energy Frontier collider include the CoM energy reach, the required AC power consumption (see **Figure 1**), and the required duration and scale of the R&D effort to reach construction readiness - see the discussion in [2; **Table 1**].

2.1.1 International Linear Collider (ILC)

The ILC is the most technologically mature of the proposed next-generation e^+e^- colliders - see [8]. The initial phase of the project has a CoM energy of 250 GeV for precision studies of the Higgs boson, a major goal in collider physics. As a linear machine it can operate at higher or lower energies. For example, the initial 21 km length can be extended to reach the threshold for the production of a top antitop quark pair ($t\bar{t}$) as well as for the associated production of a Higgs boson with a $t\bar{t}$ with CoM energies up to 500 GeV with a length of 31 km. The site allows an upgrade to one TeV by extending the SRF linac lengths and, with improvements to the SRF cavities, CoM energies of three to four TeV may be possible.

At 250 GeV, the primary enabling technology is Superconducting RF (SRF) cavities operating at an average gradient of 31.5 MV/m. The European XFEL in Hamburg Germany is a multi-billion dollar project that provides a 10%-scale demonstration of the ILC acceleration systems with over 750 SRF cavities operating at an average of 23 MV/m and producing a 17.5 GeV electron beam for the X-ray Free Electron laser (FEL). A Technical Design Report (TDR) for the ILC has been completed and the project is considered to be "shovel ready." At this point, the R&D focus is on reducing the cost of the cryomodules. The remaining technical challenges are: improvement of the positron source, achieving the nanometer-scale spot size and stability at the interaction point (IP), and optimizing the damping ring injection and extraction systems.

2.1.2 Compact Linear Collider (CLIC)

The Compact Linear Collider (CLIC) is a multi-TeV high-luminosity linear e^+e^- collider proposed by an international collaboration led by CERN - see [7]. The design is based on a staged approach that includes three CoM energies of 380 GeV, 1.5 TeV and three TeV. In contrast to the ILC, CLIC uses a novel two-beam acceleration technique with normal-conducting accelerating structures operating in the range of 70 MV/m to 100 MV/m. A CDR was produced for CLIC in 2012 and, while the design is less mature than the ILC, the CLIC design parameters are well understood and have been reproduced in beam tests indicating that the CLIC performance goals are achievable.

The main risks and uncertainties for CLIC will be in scaling from the small-scale low-power demonstrations to the km-scale high-power two-beam deceleration systems. Efforts to reduce power consumption are ongoing and new estimates show a significant reduction related to improvements in the X-band RF technology and klystron design. Like the ILC, other R&D includes improvement of the positron source, achieving the nanometer-scale spot size and stability at the interaction point (IP), and optimizing the damping ring injection and extraction systems. In general, the spot sizes and stability requirements are tighter for CLIC than for the ILC while the positron system requirements may be easier due to a lower number of e^+/sec and a higher macro-pulse repetition rate that eases requirements on a rotating target.

2.1.3 Future Circular e^+e^- Collider (FCC-ee)

The Future Circular Collider (FCC) is a proposed international collider complex located near Geneva Switzerland. It is based on the same successful staging strategy used for the Large Electron-Positron collider (LEP) and the Large Hadron Collider (LHC). The approximately 100 km tunnel would initially house the FCC-ee e^+e^- collider that would would offer a broad physics reach operating at four different CoM energies: the Z pole, WW threshold, the ZH production peak and the $t\bar{t}$ threshold. The tunnel would eventually house the FCC-hh, a 100 TeV proton-proton collider - see [4]. A Conceptual Design Report (CDR) for the FCC complex was completed and published in 2019 [9–11] and a Siting and Feasibility Study for the FCC-ee [12,13] will be completed in 2025, making it one of the more advanced proposals.

The main technologies for the FCC-ee are well-developed. The technology R&D is focused on incremental improvements aimed mainly at further optimizing electrical efficiency, obtaining the required diagnostic precision, and achieving the target performance in terms of beam current and luminosity. Optimization is also desired to improve the performance of the positron source [14]. The greatest challenge facing the facility is the difficulty and cost of tunneling in the Geneva area that the ongoing Feasibility Study is aimed at addressing. The proposed schedule allows for the start of tunnel construction in the early 2030s with first collisions in the early 2040s.

2.1.4 Circular Electron Positron Collider (CepC)

Proposed by chinese scientists in 2012, the CepC is an international scientific project hosted by China to build a 240 GeV circular e^+e^- collider in an approximately 100 km tunnel - see [6]. Similar to the FCC, the tunnel for the CepC would eventually be used for a Super Proton-Proton Collider (SPPC). A CDR for the CepC was released in November 2018. The TDR is planned for completion at the end of 2022 followed by work on an Engineering Design Report (EDR) that will look at the detailed engineering design of components, site selection and preparations for industrialization.

Like the FCC-ee, the technological basis for the design is well-understood and the R&D focus is now on improving performance of the RF klystrons, SRF cavities, high precision magnets, and vacuum systems. The proposed schedule calls for first collisions in the mid-2030s.

2.1.5 Cool Copper Collider (C^3)

The C^3 is a linear collider concept based on recent innovations in the technology of cold copper cavities using distributed coupling that allows for increased accelerator performance and better optimization - see [15]. Operation at liquid nitrogen temperature substantially increases the RF efficiency and gradient of normal-conducting copper cavities. For a 250 GeV CoM energy the accelerator is 8 km and, with additional innovation, 550 GeV could be reached using the same footprint. A GeV-scale demonstration facility is proposed to provide input for a TDR. It would include three cryomodules operating at 70, 120 and 170 MeV/m to test the RF design. A gradient of 155 MeV/m would be needed for the 550 GeV upgrade. Other collider subsystems can be based on those developed for CLIC and ILC. Areas of technical focus are development of an ultra-low emittance polarized electron gun which would benefit all linear collider designs, optimization of the RF structures and reducing the cost of the RF sources.

2.1.6 Energy Recovery and Recirculating Linacs (ERLs and RLAs)

Energy Recovery Linacs (ERL) and Circular Colliders are an alternative approach to high energy electron-positron colliders with the aim to significantly reduce beam energy losses and consequently, power consumption. There are two proposed configurations. A circular e^+e^- Collider with two 100 km storage rings using Energy-Recovery Linacs (CERC) [16,17] or two large linear colliders with damping rings, Linear Energy Recovery Linac Collider (ReLiC) [18]. Starting as a Higgs factory they have the capability of achieving CoM energies up to 600 GeV. The energy as well as the particles are recycled in this scheme and make fully polarized electron and positron beams possible. A large fraction of the energy of the used beams is recovered by decelerating them. The beams are then reinjected into a damping ring where they are cooled and reused. Beam that is lost during the recovery process is replaced via a linear injector into the damping rings. An alternative concept for an SRF ERL-based linear e^+e^- collider has been recently proposed in [19].

ERL's are an ongoing technological development. A number of compact ERL facilities and demonstrators have been constructed. The highest circulating Continuous Wave (CW) power was achieved in the IR FEL ERL at Jefferson Laboratory which operated with 8.5 mA at 150 MeV. A global development program for the ERL technology is discussed in [20].

2.1.7 X-Ray Free Electron Laser Compton Collider (XCC)

The XCC concept combines the cold copper distributed coupling technology of C^3 with X-ray FELs [21] to create a γ-γ collider. The linac accelerates electron bunches with a gradient of 70 MeV/m until 31 GeV is reached, at which point alternate bunches are diverted to an X-ray FEL to produce circular polarized 1 keV X-rays using a helical undulator. The electron bunches remaining in the linac are accelerated to 62.8 GeV through a final focus system to the e^-e^- interaction points (IP). The 62.8 GeV electrons then collide with the focused X-ray laser light from the opposite X-ray FEL, producing 62.5 GeV photons with are then collided at 125 GeV CoM in the primary IP. The number of Higgs' produced in such a machine would be comparable to the ILC, but backgrounds need to be studied in more detail.

Relative to some other e^+e^- Higgs factory proposals the XCC requires two additional beamlines and collision points and requires significant improvement to X-ray FEL technology. However, if a high brightness polarized RF gun can be developed, damping rings would not be required and the beam energy is half of that required for e^+e^- Higgs factories, raising the possibility of significant reduction in cost, a common obstacle for large-scale facilities. Gun development is challenging, but will benefit from experience with the LCLS-II-HE SRF gun when it turns on.

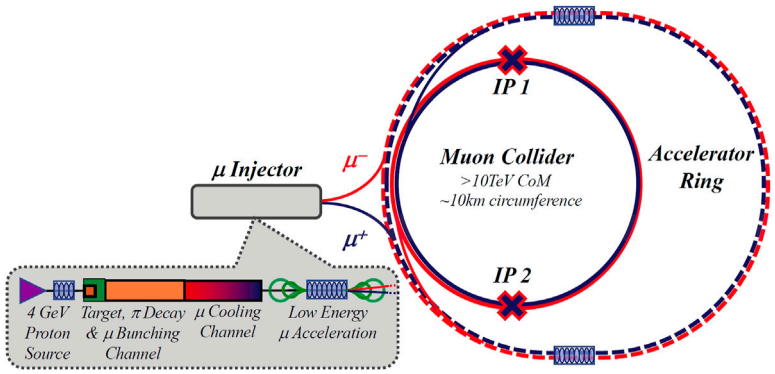


FIGURE 2 | A conceptual scheme for the muon collider [22].

2.1.8 Muon Colliders

The muon collider has great potential to extend the energy of lepton colliders by taking advantage of the strong suppression of synchrotron radiation from muons relative to electrons, though the finite lifetime of the muon is a critical issue. This allows for efficient acceleration in rings and a more compact RF system - see **Figure 2** - as well as a better defined collision energy. For example, the energy consumption of a 10 TeV CoM energy muon collider is estimated to be lower than CLIC at 3 TeV - see [5]. As a ring, a muon collider may be able to provide luminosity to two detectors.

There are a number of technological challenges that need to be addressed to be able to take advantage of the large energy and luminosity reach. To this end, an International Muon Collider Collaboration (IMCC) was formed based on a recommendation in the update of the European Strategy for Particle Physics [22]. The initial focus is on 10 TeV with an integrated luminosity of 10 ab^{-1}.

The main critical enabling technologies are 6D cooling of the muon beams, development of high field solenoids and accelerator magnets, very fast ramping magnets, high power targets, a proton driver, and operating with very strong collective effects. There has been recent progress in most of these areas demonstrating feasibility, but in the near term, a robust R&D program will be needed to bring the technology to the same level as linear or circular lepton colliders [20].

The scale of the R&D program can be expected to be a significant fraction of the final facility cost, much as for the linear collider program, where a large fraction of a B$ has been invested in dedicated international test facilities beyond the Stanford Linear Collider. Assuming sufficient funding to support a technically limited program and successful development of key technologies, it may be feasible to start colliding beams in the mid 2040s.

2.1.9 Fermilab Site Fillers

Domestically, the US is fully engaged in the Long Baseline Neutrino Facility (LBNF) and the Deep Underground Neutrino Experiment (DUNE). However, given the long time-frame needed to plan and build large accelerator facilities, preparation for major new projects needs to start soon. The Snowmass decadal strategic planning process for the US Particle Physics community (snowmass21.org) is an opportunity to consider possible options that could potentially be built in the US, and in particular, on the Fermilab site [23]. Given the current situation with the ILC and strong community support for a Higgs factory, it may be time to reconsider the option of hosting it in the US.

Other future collider concepts of intermediate-scale that might feasibly be built on the Fermilab site are also being considered:

- About 7 km long C^3 (250 GeV up to 550 GeV);
- Linear colliders utilizing high gradient SRF, standing wave or travelling wave structures (250–500 GeV);
- 16 km circumference circular e^+e^- collider with (90–240 GeV);
- Same circumference proton-proton collider (24–27 TeV);
- A staged muon collider from a Higgs factory at 125 GeV up to 8–10 TeV.

Each of these options will require varying levels of R&D to produce a CDR by the time of the next Snowmass a decade from now. To accomplish this, it is proposed that the U.S. establish an integrated future colliders R&D program in the DOE Office of High Energy Physics (OHEP) to carry out feasibility studies and collaboratively engage in projects proposed abroad.

2.2 Hadron and Hadron/Lepton Colliders

2.2.1 Future Circular Collider - Hadron/Hadron (FCC-hh)

Proposed as a second phase of the FCC program after FCC-ee, the FCC-hh is a proton-proton collider aimed at increasing the physics reach by an order of magnitude beyond the LHC [24]. Following the same strategy used for LEP and LHC, the accelerators will utilize the same 100 km-scale tunnel. It is one of the relatively mature proposals, and a conceptual design report was published in 2019 [10]. One of the major enabling technologies are the 16 T superconducting magnets required to reach the target energy of 100 TeV. This field level does not currently exist but programs in the US and EU/CERN are actively

pursuing the R&D. Experience from the construction of high gradient quadrupoles for the LHC high luminosity (HL-LHC) upgrade will serve as a launching point for further development.

Attaining the desired luminosity of $30{\cdot}10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ will be challenging. However, the HL-LHC will be an opportunity to gain considerable experience. Other challenges are related to the size and number of components, and increased number of injector rings, adding considerably to the overall complexity. Crab cavities, necessary for compensating the crossing angle at the interaction points, are still an untested system in hadron colliders, but the HL-LHC will provide an opportunity to develop this technology.

2.2.2 Super Proton-Proton Collider (SPPC)

The SPPC is proposed as the second phase of the CepC-SPPC, sharing the same tunnel in a scheme parallel to the FCC-ee/hh [25]. It is planned to operate at a CoM energy of up to 125 TeV in the final stage using 20T magnets with an intermediate stage at 75 TeV using 12T magnets.

High-field magnets are the key enabling technology for future hadron colliders. A unique feature of the magnet technology is the proposed use of iron-based high temperature superconductor (HTS). This material is still in the R&D stage, but if successful, has the potential for significant cost savings over currently available HTS materials. The project is currently in the pre-CDR phase. In addition to the high field magnets, synchrotron radiation power, luminosity and site power are a consideration in the overall design [26].

3 HIGH INTENSITY ACCELERATORS FOR NEUTRINO RESEARCH AND RARE PROCESSES

The leading accelerator-based facilities for high energy neutrino research are *superbeams* based on a conventional beam dump technique: an intense high energy proton beam is directed onto a thick nuclear target producing mostly pions and kaons, that are captured by focusing magnetic horns in order to obtain a well directed beam of same charge secondaries. High-energy neutrino beams are products of the decays of charged pions and kaons in a long decay channel [3]. Superbeams sources operate with proton beam intensities close to the mechanical stability limit of the primary targets which is at present *O* (1 MW). The most powerful accelerators for neutrino research to date are the rapid cycling synchrotron facility J-PARC in Japan which has reached 515 kW of 30 GeV proton beam power and the Fermilab Main Injector delivering up to 862 kW of 120 GeV protons on target. These facilities support neutrino oscillations research programs at the SuperK experiment (295 km from J-PARC) and MINOS (810 km from Fermilab), correspondingly.

The needs of neutrino physics call for the next generation, higher-power, megawatt and multi-MW-class superbeams facilities. Average proton beam power on the neutrino target scales with the beam energy E_b, number of particle per pulse N_{ppp} and cycle time T_{cycle} as $P_b = (E_b N_{ppp})/T_{cycle}$. Corresponding upgrades of the RF system and magnet power supply ramping rate $1/T_{cycle}$ have been initiated at J-PARC, while Fermilab has started construction of an 800 MeV SRF H- PIP-II linac (Proton Improvement Plan-II) that will help to boost N_{ppp} and the Main Injector beam power to above 1.2 MW, and considers further facility upgrades to get to 2.4 MW - see **Figure 3**.

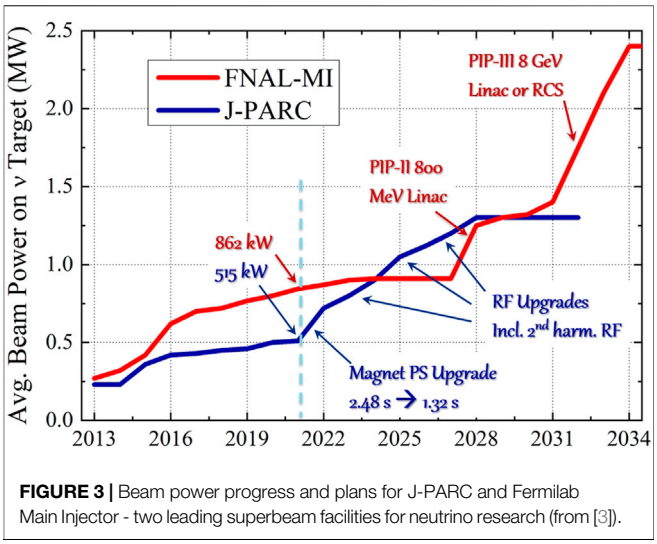


FIGURE 3 | Beam power progress and plans for J-PARC and Fermilab Main Injector - two leading superbeam facilities for neutrino research (from [3]).

Many existing and planned high intensity accelerators can be effectively used for fixed target experiments complementary to high-energy-Frontier colliders [27]. For example, a recent CERN study of the *Physics Beyond Colliders* (PBC) [28] resulted in numerous accelerator-based proposals, ranging from a gamma-factories [29] to explorations of the dark sector, to precision measurements of either strongly interacting processes or light, feebly interacting particles at beam dump facilities, such as, e.g., the SHiP experiment [30] at the SPS North Area [31]. A similar study is underway in the US as part of the Snowmass process. It considers a variety of potentially available beams ranging from 800 MeV to 120 GeV protons [32–34] to multi-GeV electrons [35].

4 ACCELERATOR TECHNOLOGIES AND ADVANCED ACCELERATORS

The cost of large accelerators is set by the scale (energy, length, power) and technology. Typically, accelerator components (Normal Conducting (NC) or/and Superconducting (SC) magnets and RF systems) account for 50 ± 10% of the total cost, while the civil construction takes 35 ± 15%, and power production, delivery and distribution technology adds the remaining 15 ± 10% [36]. While the last two parts are mostly determined by industry, the magnet and RF technologies are linchpins in the progress of accelerators.

4.1 Superconducting Magnet R&D

Superconducting magnets are required for many of the proposed facilities, in some cases with operational parameters well beyond current state-of-the-art, such as muon colliders or next

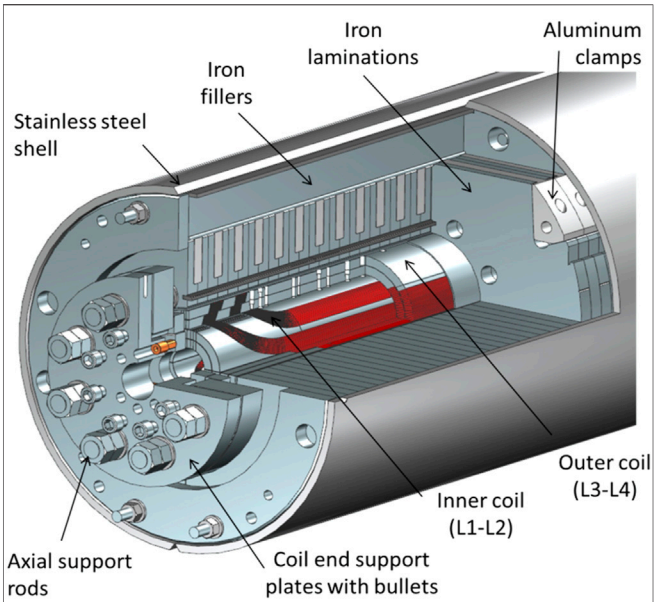


FIGURE 4 | 15 T Nb_3Sn magnet. [Zlobin].

generation high energy hadron colliders. This need was recognized by the last P5 process and the US Department of Energy-Office of High Energy Physics initiated the US Magnet Development Program (MDP) [37,38], a general R&D program that pulls together US HEP magnet research groups at DOE laboratories and Universities under a common collaboration, with focused mission and goals. Another similar collaborative effort, the High Field Magnet (HFM) program has recently been organized in the EU [39]. Superconducting Nb-Ti magnets needed for colliders operate with fields up to 8.3 T (in the LHC). Dipole fields in excess of 16 T (up to 24 T) are needed, and solenoids, primarily for the muon collider, up to 50 T [40]. A recently tested US MDP prototype Nb_3Sn magnet has reached 14.5 T [41] - see **Figure 4**. Fields above 16 T require the use of high temperature superconductors (HTS); Bi-2212, REBCO and Fe-based either in lieu of or combined in a hybrid scheme with Nb_3Sn. The fastest rapid cycling magnets (required for a muon collider) show up to 300 T/s ramping rates in the HTS-based superconducting magnet test at Fermilab [42].

4.2 RF Technology R&D

The highest gradient large-scale NC RF system is the 28 MV/m linac of the SwissFEL at the PSI (Switzerland). Up to 100 MV/m accelerating beam gradients were achieved in the CLIC 12 GHz structures at the CERN test facility, while up to 150 MV/m gradients were reported in the first test of short 11.4 GHz NC structures cooled to 77 K at SLAC [15] - **Figure 5**. As for the SC RF, the largest scale accelerator to date is the 17.5 GeV 1.3 GHz pulsed linac of the European XFEL at DESY that has achieved an average beam gradient of roughly 23 MV/m with a nominal Superconducting RF (SRF) cavity quality factor of $Q_0 \simeq 1.4 \cdot 10^{10}$ at 2 K. The full ILC specification on the beam acceleration gradient of 31.5 MV/m has been demonstrated at the FNAL FAST facility [43]. Recent advances in SRF cavity processing such as *nitrogen doping* allow further improvement of the quality factor to $(3–6) \cdot 10^{10}$ (hence, reducing the required cryogenic cooling power) and an initial cryomodule for the LCLS-II-HE, operated at 25 MV/m CM with a $Q_0 = 3 \cdot 10^{10}$. Further R&D aims for 50 MV/m gradients in 1.3 GHz structures [44]. An active ongoing accelerator R&D program for future multi-MW proton beams includes development of more efficient power supplies with capacitive energy storage and recovery, more economical RF power sources such as 80% efficient klystrons, magnetrons, and solid-state RF sources (compared to current ~ 55%) see for example [45,46].

4.3 Targets and Sources R&D

The next generation of high power targets for future accelerators will use more complex geometries, novel materials, and new concepts (like flowing granular materials, Ammigan et al. [47]). Under active development are advanced numerical approaches that satisfy the physical design requirements of reliable beam targets [48]. In parallel, there is work to be done on radiation hardened beam instrumentation [49], development of irradiation methods for high power targets and irradiation facilities [50].

There are challenges in high-intensity high-brightness beam sources for future accelerators. They are particularly formidable for sources of beams with special characteristics, such as polarized electrons and ions or ultra-small emittances, and for tertiary and secondary particles, such as muons and positrons. High intensity positron sources, which are critical for future e^+e^- linear colliders, may be simplified using advanced accelerator concepts and photon-driven schemes that can potentially outperform conventional e^+ production techniques [51].

4.4 Advanced Acceleration Methods

4.4.1 Acceleration in Plasma

Electric fields due to charge separation in plasma can sustain > GV/m gradients and have enormous promise for accelerator technology - see, e.g., [52]. To date, several experiments demonstrated O (1–10 GeV) acceleration over 0.1–10 m long plasma channels.

Plasma waves can be effectively excited by short and powerful external drivers. Three effective ways to excite plasma have been

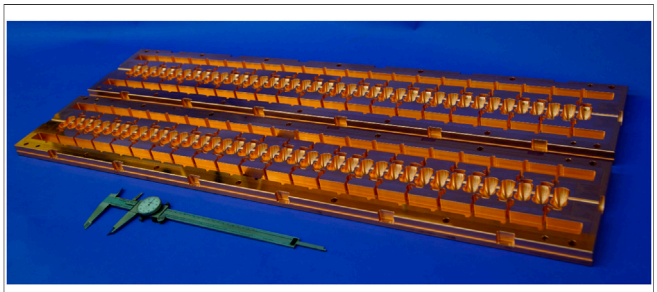


FIGURE 5 | Both halves of the C^3 prototype structure prior to braze. The 1 m structure consists of 40 cavities. A rf manifold that runs parallel to the structure feeds 20 cavities on each side. The structure operates at 5.712 GHz - from [45].

demonstrated in the past 2 decades: 1) by intense electron bunches (9 GeV acceleration over 1.3 m at 10^{17}cm^{-3}) at the SLAC FACET facility [53]; 2) short laser pulses resulted in an electron beam energy gain of up to 7.8 GeV over 20 cm in a few 10^{17}cm^{-3} plasma at the LBNL BELLA facility [54]; and 3) by self-modulated high energy proton bunches (2 GeV over 10 m at 10^{15}cm^{-3}) at the CERN AWAKE experiment [55]. In principle, plasma-based linear accelerators can be employed in high energy e^+e^- colliders, but a significant long-term R&D effort is required to address many critical issues, such as acceleration of positrons, beamstrahlung, staging of multiple plasma acceleration cells, power efficiency, emittance control, jitter and scattering in plasma, etc [2,20].

There are three remarkable recent developments on the way to practical plasma-based beam accelerators: 1) EuPRAXIA—a 569 MEuro European plasma accelerator project proposal, supported by 50 institutions from 15 countries, and aiming for 5 GeV electron beam acceleration and development of plasma-based FELs—has been included in the ESRFI 10–20 year roadmap [56]; 2) the laser wakefield accelerator at SIOM/CAS in Shanghai (China) has achieved an outstanding quality of the accelerated 0.5 GeV electron beam (produced by a 200 TW laser exciting plasma in a 6 mm He gas jet) sufficient for initiation of the FEL generation of 27 nm light in the downstream undulators—making it the first demonstration of the plasma-based FEL [57]; and 3) a similar free electron lasing demonstration at the LNF-INFN with a beam-driven plasma accelerator [58].

4.4.2 Acceleration in Structures

Relativistic electrons or positrons passing near a material boundary produce wakefields when the particle velocity exceeds the Cherenkov radiation condition. The longitudinal wakes can be used to accelerate an appropriately phased trailing beam [59] or can be extracted by a high efficiency RF coupler as a high power RF source [60]. There has been significant progress on another approach that utilizes laser-driven microstructure accelerators [61].

Structure wakefield acceleration (SWFA) has an advantage for application to linear colliders because the structures naturally accelerate electrons and positrons and are expected to have smaller challenges preserving the beam emittance than plasma accelerators. Recent progress in development has increased the gradient to 300 MV/m and wakefield power generation to 500 MW [62]. The next proposed steps are to design, construct and test a laboratory-scale SWFA module and incorporate it into a SWFA-based facility design with the intent of implementing a dedicated test facility as a scalable demonstration of the technology aimed at an energy Frontier machine.

5 BEAM PHYSICS ADVANCES

Accelerator and beam physics (ABP), the science of charged particle beams, is an essential aspect of designing and building the next Frontier accelerators. A robust and scientifically challenging program in accelerator and beam physics is critical for the field of particle physics to be productive and successful. Major ABP challenges aim at improving the reliability, performance, safety, and cost reduction of future accelerators and push the envelope of beam intensity, beam quality, beam measurement and control, and development of methods to model and predict beam behavior.

There are many notable recent developments in the physics of beams - see. e.g. [2]. Here we will present those in the fields of beam cooling and artificial intelligence and machine learning (AI/ML) and then summarize other developments across the field.

TABLE 1 | HEP Collider proposals: near-term, medium-term, and longer-term.

Collider	Species	Nominal c.m. Energy (Range), TeV	Luminosity per IP at Nominal c.m.e 10^{34} cm^{-2}s^{-1}
FCCee	e^+e^-	0.24 (0.09–0.37)	8.5 (28.9)
CEPC	e^+e^-	0.24 (0.09–0.24)	8.3 (16.6)
ILC	e^+e^-	0.25 (0.09–3)	1.4
CCC	e^+e^-	0.25 (0.25–0.55)	1.3
CLIC	e^+e^-	0.38 (0.09–3)	1.5
CERC	e^+e^-	0.24 (0.09–0.6)	78
MC-Higgs	$\mu^+\mu^-$	0.13	0.01
LHeC	e^-p	1.3	1
High-Energy ILC	e^+e^-	3 (1–3)	6.1
High-Energy CCC	e^+e^-	3 (1–3)	6.0
High-Energy CLIC	e^+e^-	3 (1–3)	5.9
Muon Collider	$\mu^+\mu^-$	10 (3–14)	20
FCChh	pp	100	30
SPPC	pp	75 (75–150)	10
Laser-Driven WFA-LC	e^+e^-	3 (1–15)	10
Beam-Driven WFA-LC	e^+e^-	3 (1–14)	6
Structure WFA-LC	e^+e^-	3 (1–15)	5.9

5.1 Beam Cooling

Beam cooling refers to the process of increasing the beam phase space density (ideally, 6D and loss-less) or, equivalently, beam emittance reduction. Widely used methods include synchrotron radiation damping, electron, stochastic and laser cooling. Over the past few years, we have seen several novel cooling schemes experimentally demonstrated at operational accelerators. A true breakthrough was the demonstration of the *ionization cooling* of 140 MeV/c muons at the MICE experiment at RAL (United Kingdom)—some 10% beam emittance reduction was observed in a single pass through the cooling section

[63]. In 2020, *"bunched" electron beam cooling* of ions in RHIC ($\gamma \sim 5$)—remarkable by the pioneering use of high quality bunched electron beams from an electron beam RF photoinjector gun (before, only DC electron accelerators were used with limited capability to get to very high energies)—was demonstrated at BNL [64]. Earlier this year another outstanding result was reported by the Fermilab team which has successfully carried out a proof-of-principle experiment on the *optical stochastic cooling* of 100 MeV electrons in the IOTA ring in which the use of undulator magnets - instead of electrostatic pickups in traditional stochastic cooling setups - allowed to expand the feedback system bandwidth by several orders of magnitude to a THz range [65]. Finally, the proof-of-principle tests of a novel *coherent electron cooling* of the record-high energy 26 GeV/*u* ions have started this year at BNL [66].

5.2 Artificial Intelligence and Machine Learning for Accelerators

Efforts have been ramping-up in recent years to use Artificial Intelligence and Machine Learning (AI/ML) to enhance the performance of accelerators and beamlines [67,68]. There is significant promise of applications of AI/ML in beam diagnostics, controls, and modeling [69]. Opportunity exists in broadening AI/ML methods for early detection of a broad range of accelerator component or subsystem failures [70] and for optimization of advanced numerical simulations through identification of the most promising combinations of parameters thereby reducing the total number of required simulations ([71]).

5.3 Other Beam Physics

A recent community exercise summed up the needs and directions of future developments in ABP [72]. Among many, those include issues related to beam loss control in high-intensity high-power accelerators (space-charge effects, instabilities, collimation, electron lens compensation, integrable optics, etc) which require innovative approaches, theoretical and experimental studies (at, e.g., the IOTA ring [73], and operational accelerators in the US and abroad) and validated computer models/codes. A key challenge would be to reduce particle losses (dN/N) at a faster rate than increases in achieved beam intensity (power) (N) [3].

Future circular and linear e^+e^- colliders (FCC-ee, CepC, gamma-gamma Higgs factory) require collision optimization studies including 3D beam size flip-flop from the beam-beam effect, polarization and Interaction Region (IR) design; pico-meter vertical emittance preservation techniques in high-charge circular colliders with strong focusing IR, detector solenoids, and beam-beam effects; end-to-end emittance preservation simulations for linear colliders should be augmented with experimental tests of the beam-based alignment techniques in the presence of realistic external noise sources; plasma-lens-based final focus and beam transport system design [74,75].

Required ABP explorations for future hadron colliders include efficient collimation techniques [76], electron lenses for Landau damping and collimation [77], dynamic aperture optimization methods to make possible new integrable optics solutions [78], and studies to obtain lower emittances from new particle sources for injecting beams in high-bunch-charge colliders.

6 SUMMARY

Accelerator science and technology has advanced significantly in the decade since the last Snowmass/Particle Physics Project Prioritization Panel (P5) strategic planning process. Advances in super-conducting and normal-conducting RF cavities, high field magnets, particle sources, targets, and advanced acceleration technologies as well as advances in the modeling and understanding of beam physics have allowed the development of many new or improved proposals for future accelerators to advance particle physics. Current activities such as the High Luminosity LHC upgrade, SuperKEKB, PIP-II, the ESS, the EIC and FAIR also serve as platforms for these future accelerators.

Thus, the ongoing Snowmass and P5 strategic planning process will have the opportunity and challenge of a variety of proposals to consider in developing a strategy for the future of particle physics. These proposals vary widely in terms of physics potential, cost, size, maturity and executable time-frame. However, with the support of a strong international accelerator and particle physics community, we are confident that there are multiple paths that will deliver a multi-decadal program in accelerator-based particle physics.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This manuscript has been authored with support from the Department of Energy under Contract No. DE-AC02-07CH11359 with the Fermi Research Alliance, LLC and Contract No. DE-AC0276SF00515 with the SLAC National Accelerator Laboratory.

ACKNOWLEDGMENTS

We greatly appreciate discussions with Pushpa Bhat (FNAL), Dmitri Denisov (BNL), and Meenakshi Narain (Brown University) within the framework of the APS *Snowmass'21* HEP planning process.

REFERENCES

1. Shiltsev V. Particle Beams behind Physics Discoveries. *Phys Today* (2020) 73. doi:10.1063/pt.3.4452
2. Shiltsev V, Zimmermann F. Modern and Future Colliders. *Rev Mod Phys* (2021) 93:015006. doi:10.1103/revmodphys.93.015006
3. Shiltsev V. Superbeams and Neutrino Factories - Two Paths to Intense Accelerator-Based Neutrino Beams. *Mod Phys Lett A* (2020) 35:2030005. doi:10.1142/s0217732320300050
4. Benedikt M, Zimmermann F. The Physics and Technology of the Future Circular Collider. *Nat Rev Phys* (2019) 1:238–40. doi:10.1038/s42254-019-0048-0
5. Long KR, Lucchesi D, Palmer MA, Pastrone N, Schulte D, Shiltsev V. Muon Colliders to Expand Frontiers of Particle Physics. *Nat Phys* (2021) 17:289–92. doi:10.1038/s41567-020-01130-x
6. Lou X. The Circular Electron Positron Collider. *Nat Rev Phys* (2019) 1:232–4. doi:10.1038/s42254-019-0047-1
7. Stapnes S. The Compact Linear Collider. *Nat Rev Phys* (2019) 1:235–7. doi:10.1038/s42254-019-0051-5
8. Michizono S. The International Linear Collider. *Nat Rev Phys* (2019) 1:244–5. doi:10.1038/s42254-019-0044-4
9. The FCC Collaboration. FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1. *Eur Phys J C* (2019) 79.
10. The FCC Collaboration. FCC-hh: The Hadron Collider. *Eur Phys J Spec Top* (2019) 228.
11. The FCC Collaboration. FCC-ee: The Lepton Collider. *Eur Phys J Spec Top* (2019) 228.
12. CERN Council. Organisational Structure of the FCC Feasibility Study. In: *Restricted CERN Council - Two-Hundred-And-Third Session* (2021). CERN/SPC/1155/Rev.2.
13. CERN Council. Main Deliverables and Timeline of the FCC Feasibility Study. In: *Restricted CERN Council - Two-Hundred-And-Third Session* (2021). CERN/SPC/1161.
14. Agapov I, Benedikt M, Blondel A, Boscolo M, Brunner O, Chamizo Llatas M, et al. Future Circular Lepton Collider FCC-Ee: Overview and Status. *arXiv: 2203.08310 [physics.acc-ph]* (2022).
15. Bai M, Barklow T, Bartoldus R, Breidenbach M, Grenier P, Huang Z, et al. C^3: A "Cool" Route to the Higgs Boson and beyond. *arXiv preprint arXiv: 2110.15800* (2021).
16. Litvinenko VN, Roser T, Chamizo-Llatas M. High-energy high-luminosity e+e– collider using energy-recovery linacs. *Phys Lett B* (2020) 804:135394. doi:10.1016/j.physletb.2020.135394
17. Litvinenko VN, Bachhawat N, Chamizo-Llatas M, Meot F, Roser T. CERC - circular e^+e^- collider using Energy-Recovery Linac. *arXiv preprint arXiv: 2203.07358* (2022).
18. Litvinenko VN, Bachhawat N, Chamizo-Llatas M, Jing Y, Meot F, Petrushina I, et al. The ReLiC: Recycling linear e^+e^- collider. *arXiv preprint arXiv: 2203.06476* (2022).
19. Telnov VI. A high-luminosity superconducting twin e+e- linear collider with energy recovery. *J Inst* (2021) 16:P12025. doi:10.1088/1748-0221/16/12/p12025
20. Adolphsen C, Angal-Kalinin D, Arndt T, Arnold M, Assmann R, Auchmann B, et al. European Strategy for Particle Physics Accelerator R&D Roadmap. *Tech Rep* (2022). CERN-2022-001.
21. Barklow T, Dong S, Emma C, Duris J, Huang Z, Naji A, et al. XCC: An X-ray FEL-Based γγ Collider Higgs Factory. *arXiv preprint arXiv:2203.08484* (2022).
22. European Strategy Group. *2020 Update of the European Strategy for Particle Physics (Brochure)* (2020). CERN-ESU-015.
23. Bhat P, Jindariani S, Ambrosio G, Apollinari G, Belomestnykh S, Bross A, et al. Future Collider Options for the US. *arXiv preprint arXiv:2203.08088* (2022).
24. Benedikt M, Chance A, Dalena B, Denisov D, Giovannozzi M, Gutleber J, et al. Future Circular Hadron Collider FCC-Hh: Overview and Status. *arXiv preprint arXiv:2203.07804* (2022).
25. Tang J, Zhang Y, Xu Q, Gao J, Lou X, Wang Y. Snowmass 2021 white Paper AF4-SPPC. *arXiv preprint arXiv:2203.07987* (2022).
26. Tang J. Design Concept for a Future Super Proton-Proton Collider. *Front Phys* (2022) 2022:138. doi:10.3389/fphy.2022.828878
27. Jaeckel J, Lamont M, Vallée C. The Quest for New Physics with the Physics beyond Colliders Programme. *Nat Phys* (2020) 16:393–401. doi:10.1038/s41567-020-0838-4
28. Beacham J, Burrage C, Curtin D, De Roeck A, Evans J, Feng JL, et al. Physics beyond Colliders at CERN: beyond the Standard Model Working Group Report. *J Phys G: Nucl Part Phys* (2019) 47:010501. doi:10.1088/1361-6471/ab4cd2
29. Budker D, Berengut JC, Flambaum VV, Gorchtein M, Jin J, Karbstein F, et al. Expanding Nuclear Physics Horizons with the Gamma Factory. *Annalen der Physik* (2022) 534:2100284. doi:10.1002/andp.202100284
30. Ahdida C, Albanese R, Alexandrov A, Anokhina A, Aoki S, Arduini G, et al. Sensitivity of the SHiP experiment to Heavy Neutral Leptons. *J High Energ Phys* (2019) 2019:1–20.
31. Ahdida C, Alia R, Arduini G, Arnalich A, Avigni P, Bardou F, et al. SPS Beam Dump Facility–Comprehensive Design Study. *arXiv preprint arXiv:1912.06356* (2019).
32. Pellico W, Bhat C, Eldred J, Johnstone C, Johnstone J, Seiya K, et al. FNAL PIP-II Accumulator Ring. *arXiv preprint arXiv:2203.07339* (2022).
33. Toups M, Van de Water R, Batell B, Brice S, deNiverville P, Eldred J, et al. SBN-BD: O(10 GeV) Proton Beam Dump at Fermilab's PIP-II Linac. *arXiv preprint arXiv:2203.08102* (2022).
34. Apyan A, Batell B, Berlin A, Blinov N, Chaharom C, Cuadra S, et al. DarkQuest: A Dark Sector Upgrade to SpinQuest at the 120 GeV Fermilab Main Injector. *arXiv preprint arXiv:2203.08322* (2022).
35. Akesson T, Blinov N, Brand-Baugher L, Bravo C, Bryngemark LK, Butti P, et al. Current Status and Future Prospects for the Light Dark Matter eXperiment. *arXiv preprint arXiv:2203.08192* (2022).
36. Shiltsev V. A Phenomenological Cost Model for High Energy Particle Accelerators. *J Inst* (2014) 9:T07002. doi:10.1088/1748-0221/9/07/t07002
37. Gourlay SA, Prestemon SO, Zlobin AV, Cooley L, Larbalestier D. The US Magnet Development Program Plan. In: *Tech. Rep.* Berkeley, CA, USA: Lawrence Berkeley National Lab.(LBNL (2016).
38. Ambrosio G, Amm K, Anerella M, Apollinari G, Arbelaez D, Auchmann B, et al. A Strategic Approach to advance Magnet Technology for Next Generation Colliders. *arXiv preprint arXiv:2203.13985* (2022).
39. Védrine P, Garcia-Tabarès L, Auchmann B, Ballarino A, Baudouy B, Bottura L, et al. High Field Magnet Development for HEP in Europe: a Proposal from LDG HFM Expert Panel. *arXiv preprint arXiv:2203.08054* (2022).
40. Alexahin YI, Barzi E, Gianfelice-Wendt E, Kapin V, Kashikhin V, Mokhov N, et al. Critical Problems of Energy Frontier Muon Colliders: Optics, Magnets and Radiation. *arXiv preprint arXiv:2203.10431* (2022).
41. Zlobin A, Turrioni D, Krave S, Stoynev S, Caspi S, Carmichael J, et al. Reassembly and Test of High-Field Nb_3 Sn Dipole Demonstrator MDPCT1. *IEEE Trans Appl Supercond* (2021) 31:4000506.
42. Piekarz H, Claypool B, Hays S, Kufer M, Shiltsev V, Zlobin A, et al. Fast Cycling HTS Based Superconducting Accelerator Magnets: Feasibility Study and Readiness Demonstration Program Driven by Neutrino Physics and Muon Collider Needs. *arXiv preprint arXiv:2203.06253* (2022).
43. Broemmelsiek D, Chase B, Edstrom D, Harms E, Leibfritz J, Nagaitsev S, et al. Record High-Gradient SRF Beam Acceleration at Fermilab. *New J Phys* (2018) 20:113018. doi:10.1088/1367-2630/aaec57
44. Grassellino A, Romanenko A, Trenikhina Y, Checchin M, Martinello M, Melnychuk OS, et al. Unprecedented Quality Factors at Accelerating Gradients up to 45 MVm−1in Niobium Superconducting Resonators via Low Temperature Nitrogen Infusion. *Supercond Sci Technol* (2017) 30:094004. doi:10.1088/1361-6668/aa7afe
45. Nanni EA, Breidenbach M, Vernieri C, Belomestnykh S, Bhat P, Nagaitsev S, et al. C^3 Demonstration Research and Development Plan. *arXiv preprint arXiv: 2203.09076* (2022).
46. Ives RL, Read M, Bui T, Marsden D, Collins G, Chase B, et al. High Efficiency, Low Cost, RF Sources for Accelerators and Colliders. *arXiv preprint arXiv: 2203.12043* (2022).
47. Ammigan K, Bidhar S, Pellemoine F, Pronskikh V, Pushka D, Yonehara K, et al. Novel Materials and Concepts for Next-Generation High Power Target Applications. *arXiv preprint arXiv:2203.08357* (2022).
48. Barbier C, Bidhar S, Calviani M, Dooling J, Gao J, Jacques A, et al. Modeling Needs for High Power Targets. *arXiv preprint arXiv:2203.04714* (2022).

49. Yonehara K. Radiation Hardened Beam Instrumentations for Multi-Mega-Watt Beam Facilities. *arXiv preprint arXiv:2203.06024* (2022).
50. Pellemoine F, Barbier C, Sun Y, Ammigan K, Bidhar S, Zwaska B, et al. Irradiation Facilities and Irradiation Methods for High Power Targets. *arXiv preprint arXiv:2203.08239* (2022).
51. Chaikovska I, Chehab R, Kubytskyi V, Ogur S, Ushakov A, Variola A, et al. Positron Sources: from Conventional to Advanced Accelerator Concepts-Based Colliders. *arXiv preprint arXiv:2202.04939* (2022). doi:10.1088/1748-0221/17/05/p05015
52. Ferrario M, Assmann R. Advanced Accelerator Concepts. *arXiv preprint arXiv: 2103.10843* (2021).
53. Litos M, Adli E, Allen JM, An W, Clarke CI, Corde S, et al. 9 GeV Energy Gain in a Beam-Driven Plasma wakefield Accelerator. *Plasma Phys Control Fusion* (2016) 58:034017. doi:10.1088/0741-3335/58/3/034017
54. Gonsalves AJ, Nakamura K, Daniels J, Benedetti C, Pieronek C, De Raadt TCH, et al. Petawatt Laser Guiding and Electron Beam Acceleration to 8 GeV in a Laser-Heated Capillary Discharge Waveguide. *Phys Rev Lett* (2019) 122: 084801. doi:10.1103/PhysRevLett.122.084801
55. Adli E, Ahuja A, Apsimon O, Apsimon R, Bachmann A-M, Barrientos D, et al. Acceleration of Electrons in the Plasma wakefield of a Proton bunch. *Nature* (2018) 561:363–7. doi:10.1038/s41586-018-0485-4
56. Assmann R, Weikum M, Akhter T, Alesini D, Alexandrova A, Anania M, et al. Eupraxia Conceptual Design Report. *Eur Phys J Spec Top* (2020) 229: 3675–4284.
57. Wang W, Feng K, Ke L, Yu C, Xu Y, Qi R, et al. Free-electron Lasing at 27 Nanometres Based on a Laser wakefield Accelerator. *Nature* (2021) 595: 516–20. doi:10.1038/s41586-021-03678-x
58. Shpakov V, Anania M, Behtouei M, Bellaveglia M, Biagioni A, Cesarini M, et al. First Emittance Measurement of the Beam-Driven Plasma wakefield Accelerated Electron Beam. *Phys Rev Acc Beams* (2021) 24:051301. doi:10.1103/physrevaccelbeams.24.051301
59. Gai W, Schoessow P, Cole B, Konecny R, Norem J, Rosenzweig J, et al. Experimental Demonstration of Wake-Field Effects in Dielectric Structures. *Phys Rev Lett* (1988) 61:2756–8. doi:10.1103/physrevlett.61.2756
60. Gao F, Conde M, Gai W, Jing C, Konecny R, Liu W, et al. Design and Testing of a 7.8 GHz Power Extractor Using a Cylindrical Dielectric-Loaded Waveguide. *Phys Rev Spec Topics-Accelerators Beams* (2008) 11:041301. doi:10.1103/physrevstab.11.041301
61. England R, Filippetto D, Torrisi G, Bacci A, Della Valle G, Mascali D, et al. Laser-driven Structure-Based Accelerators. *arXiv preprint arXiv:2203.08981* (2022).
62. Jing C, Power J, Shao J, Ha G, Piot P, Lu X, et al. Continuous and Coordinated Efforts of Structure wakefield Acceleration (SWFA) Development for an Energy Frontier Machine. *arXiv preprint arXiv:2203.08275* (2022).
63. MICE collaboration. Demonstration of Cooling by the Muon Ionization Cooling experiment. *Nature* (2020) 578:53–9. doi:10.1038/s41586-020-1958-9
64. Fedotov AV, Altinbas Z, Belomestnykh S, Ben-Zvi I, Blaskiewicz M, Brennan M, et al. Experimental Demonstration of Hadron Beam Cooling Using Radio-Frequency Accelerated Electron Bunches. *Phys Rev Lett* (2020) 124:084801. doi:10.1103/PhysRevLett.124.084801
65. Jarvis J, Lebedev V, Romanov A, Broemmelsiek D, Carlson K, Chattopadhyay S, et al. First Experimental Demonstration of Optical Stochastic Cooling. *arXiv preprint arXiv:2203.08899* (2022).
66. Wang G, Litvinenko V, Pinayev I, Jing Y, Kayran D, Ma J, et al. Summary of the CeC experiment in RHIC Run 21 [(No. CA/AP/656; BNL-222354-2021-TECH]. In: *Tech. Rep*. Upton, NY, USA: Brookhaven National Lab.(BNL (2021).
67. Duris J, Kennedy D, Hanuka A, Shtalenkova J, Edelen A, Baxevanis P, et al. Bayesian Optimization of a Free-Electron Laser. *Phys Rev Lett* (2020) 124: 124801. doi:10.1103/physrevlett.124.124801
68. Arpaia P, Azzopardi G, Blanc F, Bregliozzi G, Buffat X, Coyle L, et al. Machine Learning for Beam Dynamics Studies at the CERN Large Hadron Collider. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2021) 985:164652. doi:10.1016/j.nima.2020.164652
69. Edelen AL, Biedron SG, Chase BE, Edstrom D, Milton SV, Stabile P. Neural Networks for Modeling and Control of Particle Accelerators. *IEEE Trans Nucl Sci* (2016) 63:878–97. doi:10.1109/TNS.2016.2543203
70. Edelen A, Biedron S, Milton S, Edelen J. First Steps toward Incorporating Image Based Diagnostics into Particle Accelerator Control Systems Using Convolutional Neural Networks. *arXiv preprint arXiv:1612.05662* (2016).
71. Koser D, Waites L, Winklehner D, Frey M, Adelmann A, Conrad J. Input Beam Matching and Beam Dynamics Design Optimization of the IsoDAR RFQ Using Statistical and Machine Learning Techniques. *arXiv preprint arXiv: 2112.02579* (2021).
72. Nagaitsev S, Huang Z, Power J, Vay J-L, Piot P, Spentzouris L, et al. Accelerator and Beam Physics Research Goals and Opportunities. *arXiv preprint arXiv: 2101.04107* (2021). doi:10.2172/1764152
73. Antipov S, Broemmelsiek D, Bruhwiler D, Edstrom D, Harms E, Lebedev V, et al. IOTA (Integrable Optics Test Accelerator): Facility and Experimental Beam Physics Program. *J Inst* (2017) 12:T03002. doi:10.1088/1748-0221/12/03/t03002
74. Ariniello R, Doss C, Hunt-Stone K, Cary J, Litos M. Transverse Beam Dynamics in a Plasma Density Ramp. *Phys Rev Acc Beams* (2019) 22: 041304. doi:10.1103/physrevaccelbeams.22.041304
75. Lindstrøm CA. Staging of Plasma-wakefield Accelerators. *Phys Rev Acc Beams* (2021) 24:014801.
76. Scandale W, Taratin AM. Channeling and Volume Reflection of High-Energy Charged Particles in Short Bent Crystals. Crystal Assisted Collimation of the Accelerator Beam Halo. *Phys Rep* (2019) 815:1–107. doi:10.1016/j.physrep.2019.04.003
77. Shiltsev V. Electron Lenses: Historical Overview and Outlook. *J Inst* (2021) 16: P03039. doi:10.1088/1748-0221/16/03/p03039
78. Danilov V, Nagaitsev S. Nonlinear Accelerator Lattices with One and Two Analytic Invariants. *Phys Rev Spec Topics-Accelerators Beams* (2010) 13: 084002. doi:10.1103/physrevstab.13.084002

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Gourlay, Raubenheimer and Shiltsev. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

ORIGINAL RESEARCH
published: 19 July 2022
doi: 10.3389/fphy.2022.909205

Tau Lepton Identification With Graph Neural Networks at Future Electron–Positron Colliders

Stefano Giagu[1]*, *Luca Torresi*[2]* and *Matteo Di Filippo*[1]

[1]Department of Physics, Sapienza Università di Roma and INFN Sezione di Roma, Rome, Italy, [2]Institut für Theoretische Informatik, Karlsruhe Institute of Technology, Karlsruhe, Germany

Efficient and accurate reconstruction and identification of tau lepton decays plays a crucial role in the program of measurements and searches under the study for the future high-energy particle colliders. Leveraging recent advances in machine learning algorithms, which have dramatically improved the state of the art in visual object recognition, we have developed novel tau identification methods that are able to classify tau decays in leptons and hadrons and to discriminate them against QCD jets. We present the methodology and the results of the application at the interesting use case of the IDEA dual-readout calorimeter detector concept proposed for the future FCC-ee electron–positron collider.

Keywords: Bayesian inference, graph neural networks, calorimetry, tau identification, dual-readout, IDEA calorimeter, FCC, point-based model

OPEN ACCESS

Edited by:
Patrizia Azzi,
National Institute of Nuclear Physics of Padova, Italy

Reviewed by:
Frank Krauss,
Durham University, United Kingdom
Christian Veelken,
National Institute of Chemical Physics and Biophysics (NICPB) Tallinn, Estonia

***Correspondence:**
Stefano Giagu
stefano.giagu@uniroma1.it
Luca Torresi
luca.torresi@kit.edu

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *31 March 2022*
Accepted: *13 June 2022*
Published: *19 July 2022*

Citation:
Giagu S, Torresi L and Di Filippo M (2022) Tau Lepton Identification With Graph Neural Networks at Future Electron–Positron Colliders. Front. Phys. 10:909205. doi: 10.3389/fphy.2022.909205

1 INTRODUCTION

The Future Circular Collider (FCC) [8] is a proposed design for a new research infrastructure that will host a 100 km particle accelerator, in order to extend the research currently being conducted at the LHC, once the high-luminosity phase (HL-LHC) reaches its conclusion in 2040 [24]. In the first phase as an electron–positron collider, FCC-ee is designed to deliver the highest possible statistics and ultimate precision of Z, W, and Higgs bosons and the top quark. FCC-ee is expected to produce $\sim 10^{12}$ events $Z \to \tau\tau$ around the Z pole, the largest sample of $\tau\tau$ events foreseen at any lepton collider [12]. Several detector concepts have been studied to fully exploit the high energy and the great luminosity that new colliders will reach, like the IDEA (Innovative Detector for Electron–positron Accelerators) concept [1], that will provide an innovative calorimeter [2], based on the dual-readout method, that is expected to guarantee great performances in the event reconstruction.

Tau is the charged elementary particle belonging to the third lepton generation; its mass, around 15 times larger than the muon mass, makes the τ the only lepton that can decay into hadrons. These decay modes, unlike the ones originated from quarks and gluons, can be described and predicted by the weak interaction decay theory and quantum chromodynamics, since tau decays as a free, isolated particle [32]. Several experiments leverage this property in order to find discrepancies with the theory that could lead to new physics beyond the Standard Model, like the charged lepton flavor violation (cLFV) processes [7, 19], the violation of the lepton universality [12], or the tau polarization [6]. Tau provides an optimal channel for Higgs precision measurements, since a significant fraction of the SM Higgs boson decays into the di-τ channel. In particular, the branching ratio of $H \to \tau\tau$ is ~ 6.3 %, a factor 100 or 10 larger than the BR($H \to \mu\mu$) ~ 0.02 %, or the BR($H \to \gamma\gamma$) of ~ 0.22 %, used in the Higgs discovery [3, 5]. For these reasons, Tau identification and reconstruction can play a relevant role for new physics search in the Higgs sector at FCC-ee [4, 11, 13, 15, 21, 28, 30, 36].

In the last few years, deep learning (DL) algorithms have dramatically improved the state of the art in many fields, such as speech recognition, visual object recognition, and object detection, and have also been successfully applied to the analysis of data collected in high-energy physics experiments. In the jet tagging context, where the task is to identify the elementary particle that originated the jet, several machine learning methods have been explored with great results. Representation can be crucial to highlight peculiar characteristics of an object or a pattern; jets have been studied viewing them as two- or three-dimensional images, as sequences, trees, or graphs of particles. The image-based approach achieved high performances using convolutional neural networks (CNNs) developed in the computer vision area. For tau decay identifications, DL techniques have been used since 1992 at LEP [25]; however, complex multivariate analysis (MVA) using observably motivated by the QCD theory is usually preferred in the modern experiments [10, 22].

In this work, motivated by the ParticleNet architecture [35], developed for jet tagging, where a jet is viewed as an unordered set of particles (Particle Cloud) and the neural network acts on graphs dynamically created in the analysis, we extended it for the tau decay identification task. In the simulation, taus originate from Z bosons and decay in the IDEA detector, where all the particles are detected by the calorimeter. In our study, we use only calorimetric data, and each tau event has to be recognized and discriminated against jet events originated directly from Z. As the inner tracking and particle flow reconstruction for the IDEA detector is still a work-in-progress task, we have not included this information yet in the tau identification algorithm presented here. We have planned to include it, when ready, in a successive development of the algorithm.

2 MATERIALS AND METHODS

2.1 Dataset

All the experiments in this study were performed on simulations of Z boson decay events absorbed by the IDEA calorimeter. The IDEA detector is based on the dual-readout technology, a concept that has been extensively studied over 10 years of R&D by the DREAM/RD52 collaboration [39, 40]. It is composed of an ultra-light drift chamber with a low-mass superconducting solenoid coil, as main tracker, and a dual-readout fiber calorimeter for both hadronic and electromagnetic energy measurements. A more exhaustive description of the parts can be found in [1]. The calorimeter has a "barrel" geometry centered around the interaction point, whose inner length and diameter are 6 m, and the outer diameter measures 9 m. The endcaps have an inner radius of about 0.25 m, guaranteeing an angular acceptance above 100 mrad from the beam line. Scintillation and Cherenkov fibers of 1 mm diameter are embedded in a copper absorber material, displaced in a checkerboard-like structure at distances of 1.5 mm for a total of about 13 × 107 fibers [2].

Products of the electron and positron collision at the Z pole and the subsequent decay of Z bosons were simulated with Pythia, a standard tool for the generation of high-energy collisions [37], while the whole IDEA calorimeter was simulated in Geant4 [16], considering both the solenoid material and the magnetic field in the drift chamber. Each fiber of the calorimeter is readout with a dedicated silicon photo-multiplier (SiPM). Their electronic output was generated implementing a Hamamatsu SiPM sensor through the SimSiPM library [33], developed in the context of IDEA DRC Software. The dark count rate (DCR) and after-pulse noise were included as well in the signal generation. In simulation 7, main τ decay channels as well as the hadronic jet decay of the Z boson were considered, as summarized in **Table 1**, for a total of 5 k $Z \to qq$ events and 30 k $Z \to \tau\tau$ events overall.

TABLE 1 | Decay identification dataset. Leptonic decays in blue and hadronic ones in red.

Decay	Label eight-class	#Event
$\tau^- \to e^- \nu_e \nu_\tau$	0	5,000
$\tau^- \to \pi^- \nu_\tau$	1	5,000
$\tau^- \to \pi^0 \pi^- \nu_\tau$	2	5,000
$\tau^- \to \pi^0 \pi^0 \pi^- \nu_\tau$	3	5,000
$\tau^- \to \pi^- \pi^- \pi^+ \nu_\tau$	4	5,000
$\tau^- \to \pi^0 \pi^- \pi^- \pi^+ \nu_\tau$	5	5,000
$\tau^- \to \mu^- \nu_\mu \nu_\tau$	6	5,000
$Z \to q\bar{q} \to jet\,jet$	7	5,000

The process $e^- e^+ \to Z$ being simulated around the Z pole, the center of mass of the Z boson is still with respect to the interaction point and the two τs or quarks, produced in the decay generate showers in the calorimeter in a back-to-back configuration. This simplified setting allowed us to separate each of the two clusters of any given event by means of simple geometrical clustering methods, without using specific algorithms that take into account models of the physics of the process. Given that in each event there must be two particle showers and that these are in back-to-back configuration, we opted for a simple k-means. Having identified the centroid of each cluster, we extracted the active fibers inside a unit circle around it, $\sqrt{\Delta\theta^2 + \Delta\phi^2} < 1$. No fiducial volume has been applied; thus, the part of information of the clusters close to the edges of the calorimeter may have been cut out. We then considered each cluster as an independent data point to be classified.

Performing voxelization or similar operations to regularize the geometric structure of the data would result in the loss of spatial resolution, which would mean hinder a distinctive point of strength of the IDEA calorimeter. Thus, we directly formalized our data as a point cloud. **Figure 1** shows few examples of the cluster patterns we obtained.

Given a fiber fired in an event, the following features are collected:

- geometric information: in a first approximation, we used the distance in polar coordinates ($\Delta\theta$ e $\Delta\phi$) from the center of the cluster, which leads to a partial distortion of the patterns close to the endcaps of the calorimeter;
- energy information: total number of photo-electrons and total energy released in the fiber;

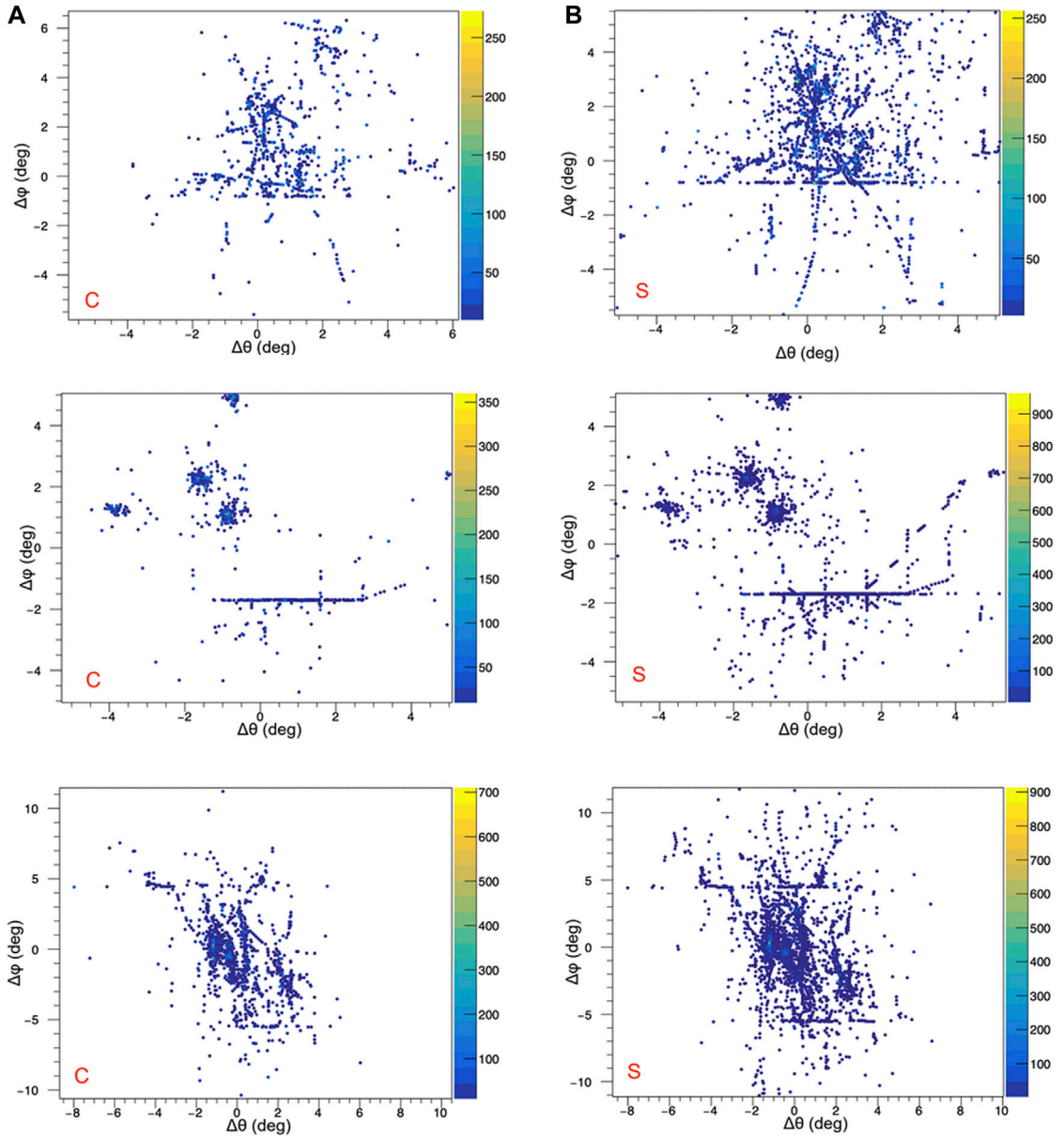


FIGURE 1 | Examples from our dataset. Columns represent Cherenkov **(A)** and scintillation **(B)** channels; rows represent events. From the top to bottom: QCD jet, $\tau^- \to \pi^0\pi^0\pi^-\nu_\tau$, and $\tau^- \to \pi^0\pi^-\pi^-\pi^+\nu_\tau$.

- fiber type: one-hot encoding of active fiber's type (Cherenkov or scintillating);
- SiPM electronic information: integral and peak of the signal, time of arrival, time over threshold, and time of peak.

A summary of the available features is reported in **Table 2**.

A nice advantage in using a point cloud representation is that incorporating additional information is straightforward, and it is sufficient to concatenate the additional vectors to the internal representation of each point. Thus, our method can simply be extended to include additional data coming from the calorimeter itself or from other detectors, for e.g., the inner tracker and the muon spectrometer.

2.2 Dynamic Graph CNN

The main property that a model has to ensure when operating on point clouds for classification tasks is permutation invariance, i.e., the output has to be constant regardless of the order in which the inputs are presented to it. In the most recent literature, various architectures can be found that ensure such a property and directly manipulate raw point cloud structured data, for e.g., [29, 34, 38, 41]. Point clouds inherently lack topological

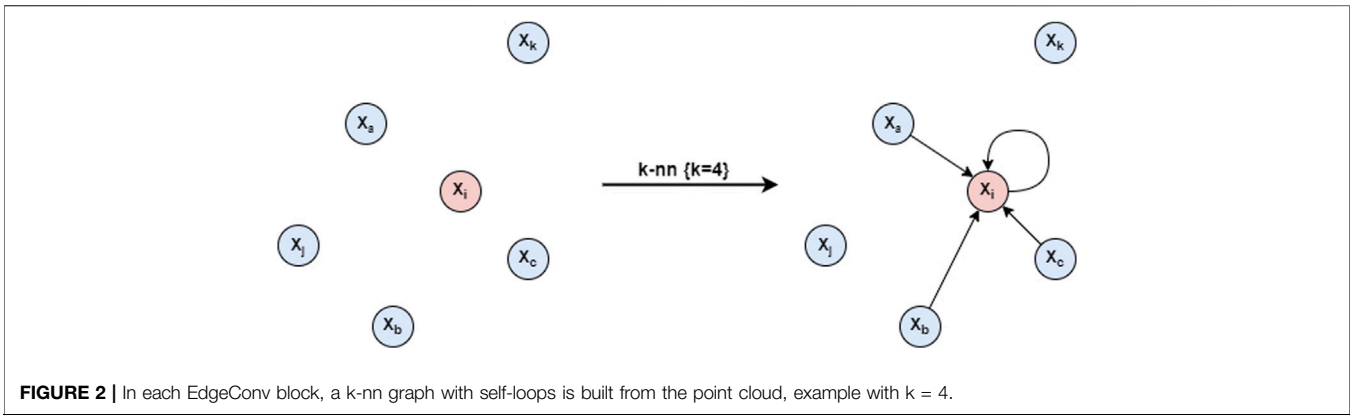


FIGURE 2 | In each EdgeConv block, a k-nn graph with self-loops is built from the point cloud, example with k = 4.

TABLE 2 | List of the observables of a fired fiber available in the dataset used in our study.

Input type	Observable	Description
Geometric information	$\Delta\theta$ $\Delta\phi$	Distance in polar coordinates from the left of the cluster
Energy information	#p.e. E_{tot}	Total number of photo-electrons in the fiber and its corresponding energy
Fiber type	S/C	Whether the fiber is scintillating or Cherenkov
SiPM information	Integral Peak ToA ToT ToP	Signal information from the silicon photomultiplier: integral and peak of the signal, time of arrival, time over threshold, and time of the peak

information; thus, a model able to recover topology can enrich their representation power. We opted for a dynamic graph convolutional neural network (DGCNN) [38], a graph-based model which explicitly takes advantages of local geometric structure.

The original module introduced in the aforementioned work is named EdgeConv, suitable for CNN-based high-level tasks on point clouds, including classification and segmentation. It is an operation that, while maintaining permutation equivariance, constructs from the point cloud of a local neighborhood graph and applies convolution-like operation on the edges connecting neighboring pair of points, in the spirit of graph neural networks [42, 43].

Given a point cloud $X = \{x_1, x_2, \ldots, x_N\}$, where each $x_i \in \mathbb{R}^F$, an EdgeConv block first encodes its geometrical structure building a k-nn-directed graph with self-loops, i.e., connections from a vertex to itself (see **Figure 2**). Then, for each $i - th$ node of the graph, it extracts its edge features e_{ij} by means of a nonlinear function h_Θ: $\mathbb{R}^F \times \mathbb{R}^F \rightarrow \mathbb{R}^{F'}$ with learnable parameters Θ, for e.g., an MLP neural network, in the formula: $e_{ij} = h_\Theta(x_i, x_j)$. Finally, the EdgeConv applies a channel-wise aggregation operation $\square$ on the edge features associated to each point, for e.g., max-pooling or a sum. Thus, the output of the layer is of the form:

$$x_i' = \square_{j \in \mathcal{N}(x_i)} h_\Theta\left(x_i, x_j\right). \tag{1}$$

In particular, in the original formulation, the operator is defined, channel-wise, as:

$$\begin{cases} e_{ij} = ReLU\left(\theta \cdot \left(x_j - x_i\right) + \phi \cdot x_i\right), \\ x_i' = max_{j \in \mathcal{N}(x_i)} e_{ij} \end{cases}. \tag{2}$$

Overall, given an F-dimensional point cloud with n points, EdgeConv produces an F'-dimensional point cloud with the same number of points. A basic scheme of the module is depicted in **Figure 3**.

The Edge block can be integrated into deep learning models to improve their performance. In the original DGCNN article, it is integrated into the basic version of PointNet [34]. Its resulting structure, shown in **Figure 4**, can be schematized with:

- a hierarchical feature extractor composed by a sequence of EdgeConv layers with skip-connections [20];
- an aggregator (max, avg, or sum) that produces a global feature for the whole point cloud;
- a classifier, implemented in the original article with MLPs.

The characteristic that differentiates DGCNN from other graph-based neural network architectures is that it does not make use of a fixed graph but rather dynamically updates it in each layer of the network. The k-neighborhood of each point may change from layer to layer, and thus proximity in feature space differs from proximity in the input. This leads to non-local diffusion of information throughout the point cloud and makes the receptive field as large as the diameter of the point cloud, while being sparse.

The DGCNN architecture requires a fixed number of points for each point cloud representing the events. The dataset studied in our experiments contains a highly variable number of active fibers per event, ranging from 150 to as much as 4,000. To tackle this problem, as already carried out in [35], the number of fibers considered for each event was fixed to a specific number n, which we treated as an hyperparameter to be tuned. In the case the active fibers were more than n, the ones with the lowest values of signal integral were discarded; if, instead, the number of active fibers was lower than n, a set of zero-valued artificial points is added in order

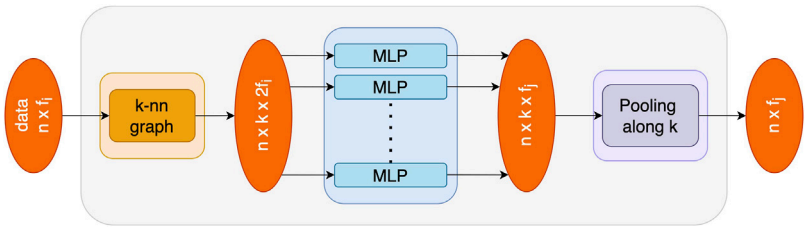


FIGURE 3 | Basic EdgeConv block, as introduced in [38].

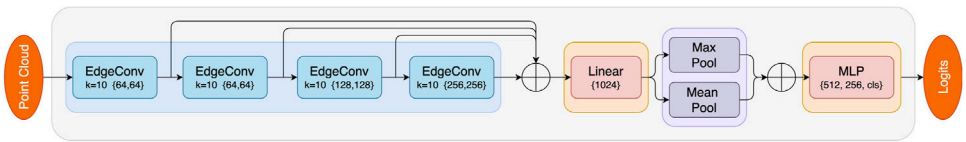


FIGURE 4 | DGCNN architecture for tau decay identification.

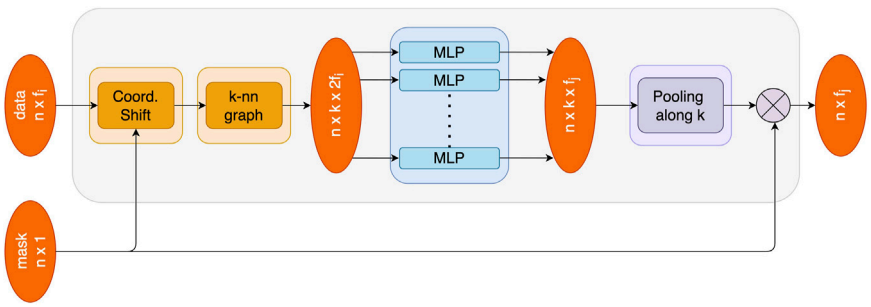


FIGURE 5 | EdgeConv block augmented with masks for synthetic datapoints.

to reach the correct number. We keep track of the possibly present artificial vectors by means of binary masks associated to each cluster. These are used to shift far from zero coordinates of the added vectors in such a way not to introduce artefacts when the k-nn graph is generated and to zero-out their computed features at the output of each EdgeConv block, as shown in **Figure 5**.

2.3 Bayesian Neural Networks

It is not straightforward to extract confidence measures on the predictions of neural networks. In particular, in classification tasks, we should not expect that the last Softmax layer probabilities reflect their ground truth correctness likelihood. It has in fact been showed that, despite the recent improvements in terms of accuracy, most modern neural networks are not well-calibrated [18].

When performing physics experiments, the evaluation of the uncertainty is an integral part of every measurement; a measured value without the corresponding uncertainty only provides partially useful physical information. For example, evaluating the measurement uncertainty is fundamental in scientific research to establish the validity limits of theories. Being able to assign a reliable confidence estimate on a prediction made by a deep learning model is therefore crucial in the context of physics experiments and one viable solution to attain this are Bayesian neural networks.

While in traditional neural networks, deterministic values are inferred for the model parameters in the training phase, these point estimates are replaced by probability distributions in the Bayesian approach. This is carried out extending standard networks with posterior inference, i.e., adding a prior distribution on the model parameters. From a statistician's point of view, this is equivalent to switching from a maximum likelihood estimation of traditional neural networks (or a maximum a posteriori estimation when regularization is used) to a Bayesian inference approach [26].

This approach allows to measure uncertainty and identify out-of-distribution inputs while regularizing the whole network and preventing overfitting of the training data [17]. Training a BNN can be interpreted as training an infinite ensemble of models with the same structure [44]. An estimate of the uncertainty of the BNN can be obtained by comparing the predictions of multiple

sampled model parameterizations: the uncertainty is low when there is agreement among different models, otherwise high.

Among the various approaches to infer the posterior distribution of the model parameters [9, 14, 23, 31], given the complexity of our network, we opted for a variational approach, namely, the *Bayes by Backprop* algorithm introduced originally in [9].

From our point of view, the advantages of this approach are that:

- the loss is still completely differentiable thanks to the reparameterization trick, thus enabling the use of the backpropagation algorithm;
- the number of parameters only doubles with respect to the corresponding non-Bayesian model, given that each weight w_i is approximated by a normal distribution with the mean μ_i and standard deviation σ_i;
- the training phase is fast, and any deterministic model can be easily extended to its Bayesian counterpart.

Following the Bayes by Backprop approach, we have designed a Bayesian DGCNN that shares essentially the same architecture of the DGCNN but with each weight replaced by a normal distribution described by the mean and the standard deviation as learnable parameters.

3 RESULTS

3.1 Decay Classification

The main task we want to solve is to correctly match single tau clusters reconstructed in the dual-readout calorimeter with the respective decay channel of the tau lepton and distinguishing them from clusters due to QCD jet events. As described in **Section 2.1**, tau and QCD jets used in this study are produced from Z decays. The classification was performed directly on the raw data collected from the IDEA calorimeter, treating each cluster of each event as an independent point cloud, with each point exhibiting the features listed in **Section 2.1**.

We balanced our dataset keeping 5,000 clusters for each decay channel, as shown in **Table 1**. This does not correspond to the true probabilities of each decay type to occur but ensures that our classifier is independent of any prior and bases its estimations only on patterns found in the calorimeter readout signals.

We treated most of the architectural parameters of our network as hyper-parameters to be optimized on the validation set. The resulting architecture consists in a feature extractor of four EdgeConv layers, each composed of two-layered MLPs with the following dimensions: \{64, 64\}, \{64, 64\}, \{128, 128\}, and \{256, 256\}. We composed the aggregator as an MLP of dimension 1024 followed by the parallel max and average pooling blocks. Finally, the classifier was implemented with a two-hidden layered MLP of dimension \{512, 256\} and an eight-class Softmax layer as output, for a total number of parameters in the order of 1.9M, in line with the other SOTA models for point-cloud classification. The complete network architecture is schematized in **Figure 4**.

As regularization, we introduced batch normalization layers after each EdgeConv and dropout layers before each linear layer in the MLP classifier. Various methods for data augmentation were investigated to improve robustness; a random sparsification of the input turned out to yield best results in terms of generalization, surprisingly enough with a relatively low retention probability (fibers zeroed out with probability ~0.7). Finally, we applied early stopping on the validation accuracy over a total of 200 epochs, enough to guarantee convergence.

We list here the other optimal parameters and the training setting: learning rate: E-03, optimizer: Adam, step scheduler with γ: 0.5 and milestones: [25, 50, 75], batch size: 64, dropout probability: 0.5, input dimension: 2000, k: 10 (for knn in EdgeConv), pooling: max (in EdgeConv), and activation functions: LeakyReLU with slope 0.1.

An assumed advantage of dual-readout technology is that the patterns generated in the calorimeter would be easily identifiable thanks to the different properties of the fibers that constitute it. To test the informativeness of the patterns and attest the gain in discriminability given by the information about the two different fiber types and as a first benchmark for our model, we trained the DGCNN using as node features the geometric information only, namely, the distances in polar coordinates from the cluster centroid of the fired fibers and the geometric information concatenated with the one-hot encoding of the fiber type; in both cases, no energetic information was provided to the model. We obtain an overall accuracy of 73.7% in the geometric only configuration and of 88.3% when adding the fiber type information; **Figure 6** compares the confusion matrices (normalized per row) obtained in these two settings. This result provides a first interesting conclusion: leveraging the differential information provided by the scintillation and Cherenkov fibers of the dual readout calorimeter significantly improves the tau identification performance compared to using geometric information alone. The neural network in fact can efficiently use the information coming from the dual-readout for particle identification and the information about the fiber type results particularly important to discriminate among the hadronic decays (rows 2–6).

We then tested performance improvement that can be achieved by adding as features for each fiber the quantities available from the readout electronic simulation: energy information as the integral and peak of the SiPM signal and temporal information, namely, the time of arrival (ToA), time of peak (ToP), and time over threshold (ToT). We experimented with different configurations, always including the geometrical and fiber type information, to have an estimate of feature importance. The results are summarized in **Table 3**, and confusion matrices on the test set are reported in **Figure 7**. The temporal information, more than the energy, seems particularly informative for the model. We speculate that this is due to the fact that temporal information can be interpreted by the model as a proxy measure of the depth in which the particles interact with the calorimeter, adding a third dimension to the absorption patterns. In this particular dataset, produced around the Z pole, the use of energy information could in principle induce in the model a bias over the distribution of the total energy per event. We plotted (**Figure 8**) such distributions for each decay mode both using the true labels and

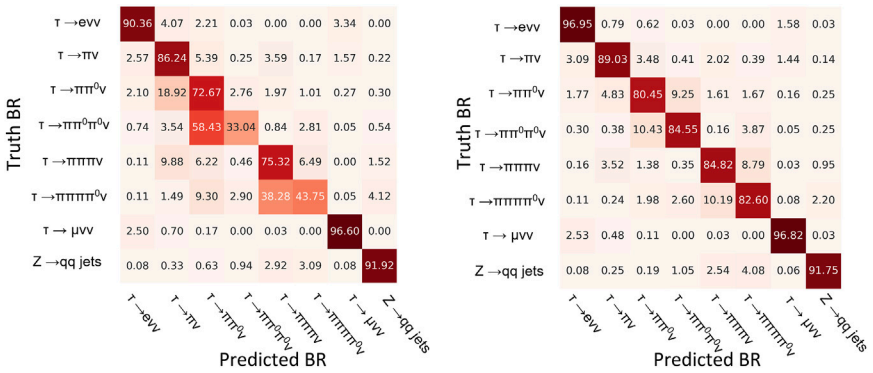


FIGURE 6 | Confusion matrices of DGCNN on the test dataset, using *geometric only* and *geometric+fiber type* information. Matrices are normalized per row.

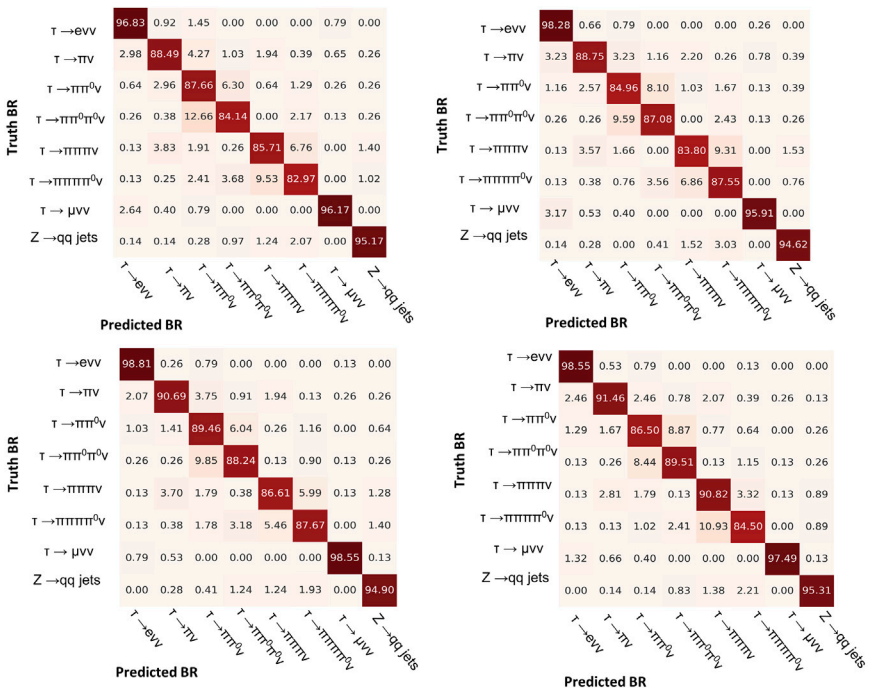


FIGURE 7 | Confusion matrices of DGCNN on the test dataset using as features, respectively, in clockwise order from the op left: only integral, only peak, only temporal, or whole SiPM information. In each model, geometric and fiber type information was also used. Matrices are normalized per row.

the labels predicted by our model, and we found no systematic variation between them. Ulterior experiments were performed in a setting with only three classes (leptonic and hadronic tau decays and quark jets), but we found no improvements with respect to the eight-class model performances relative to those decay modes.

3.2 Uncertainty Estimation With a Bayesian DGCNN

The architecture of our BNN is essentially the same as the model described in **Section 3.1**, with each weight replaced by a normal distribution defined by the trainable parameters (μ, σ), representing, respectively, the mean and the standard deviation of the distribution, so that the total number of parameters roughly doubles with respect to its deterministic counterpart.

TABLE 3 | Train and validation accuracies obtained using SiPM features. In addition to the SiPM features, both coordinates and fiber type information were fed to the model in all experiments.

SiPM feature	Train acc.	Validation acc.
Integral	0.917	0.877
Peak	0.921	0.889
Temporal	0.942	0.909
All	0.940	0.914

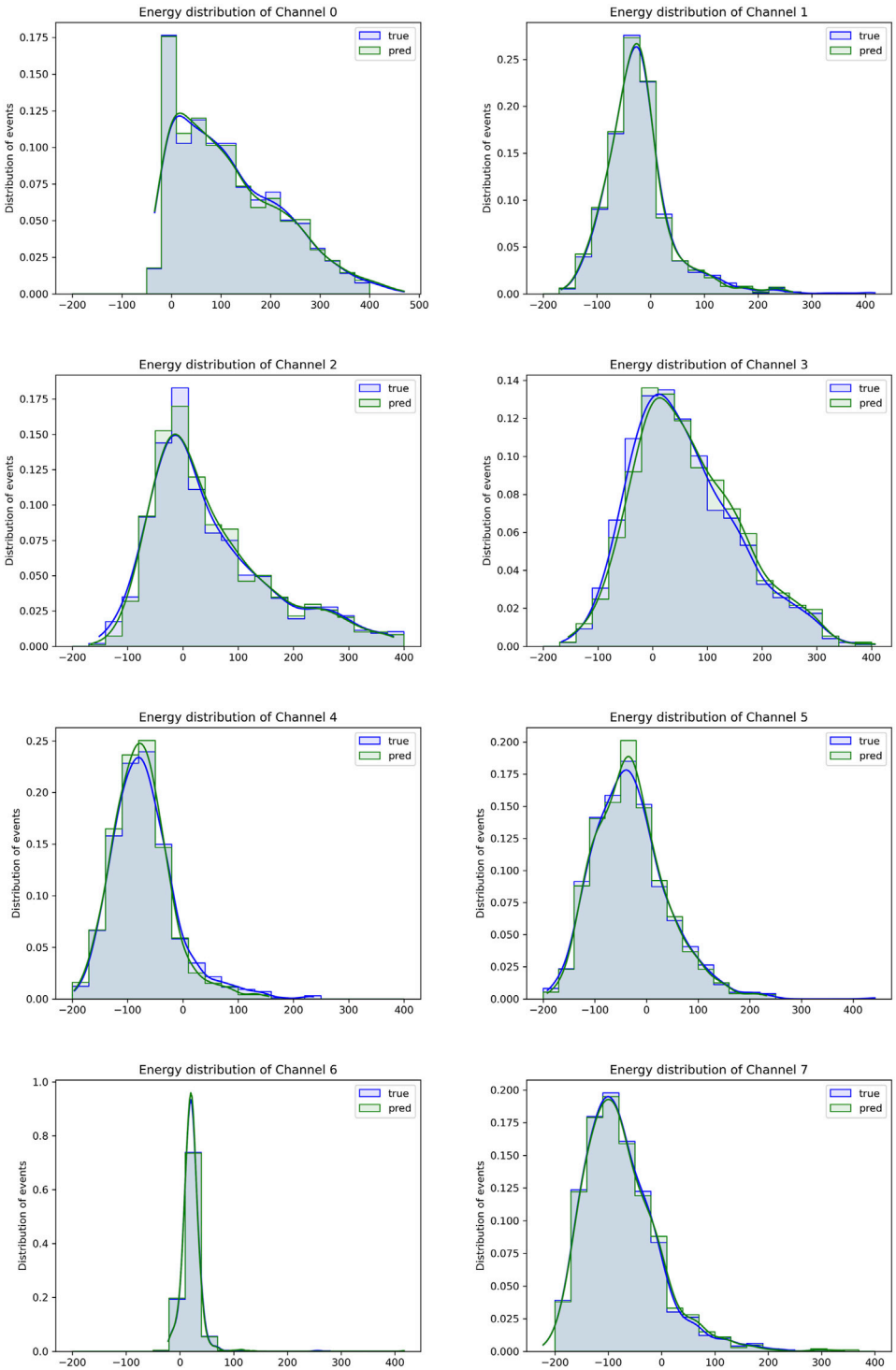


FIGURE 8 | Distribution of total energy per event for eight decay channels. Channels are referred to **Table 1**.

Frequently, when transposing a deterministic network to its Bayesian counterpart, only the weights of the last few layers of the network are treated as distributions, as this simplifies the implementation and the computational complexity of the problem. We found out instead that our model achieved better performance in terms of generalization converting to the Bayesian framework all convolutional and fully connected layers of the model and removing instead all dropout and batch normalization layers. The only other notable difference with the deterministic DGCNN model is the use of scaled

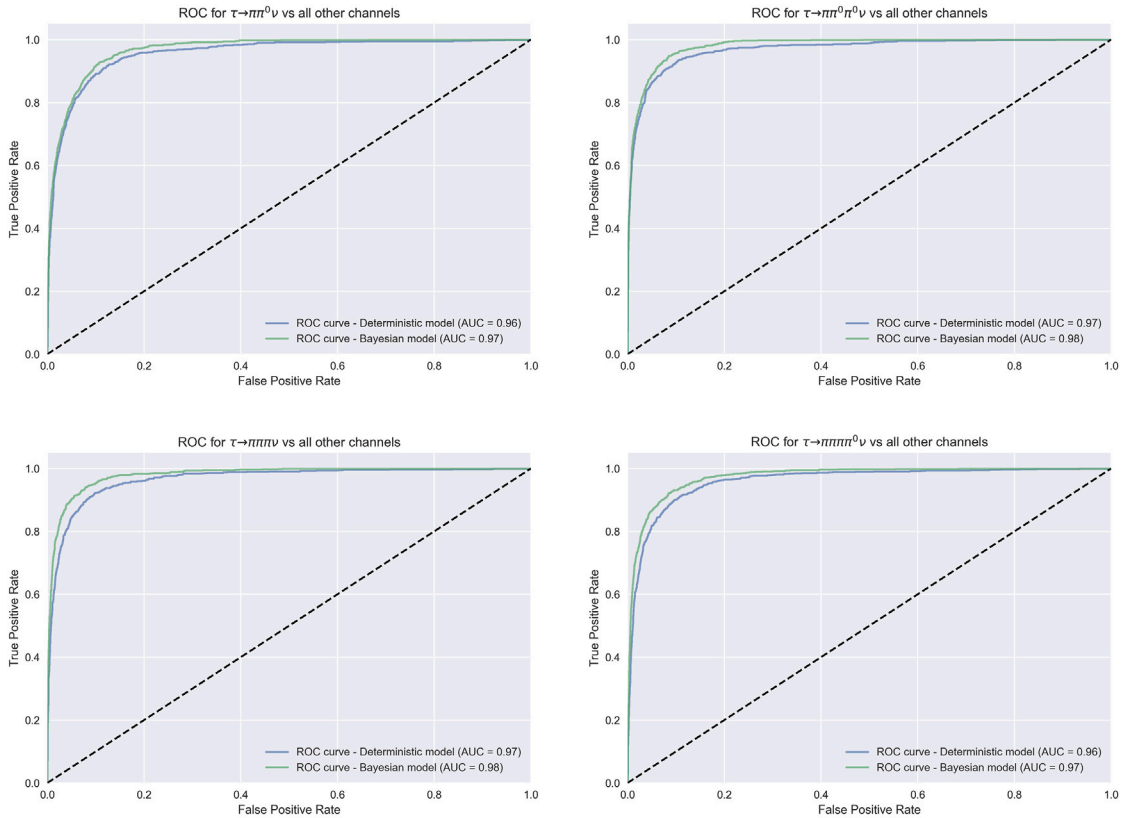


FIGURE 9 | ROC curves of the Bayesian DGCNN and point-estimate DGCNN for four classes of decay.

TABLE 4 | Bayesian DGCNN test accuracies by varying number of samples of network parameters and minimum confidence in the classification prediction.

Number of samples	Minimum confidence	Events considered (%)	Test accuracy
1	0.0	100	0.854
3			0.873
10			0.869
3	0.5	94.83	0.896
	0.7	80.33	0.943
	0.9	62.27	0.977
10	0.5	94.72	0.900
	0.7	79.82	0.947
	0.9	60.52	0.981

exponential linear units (SELUs) as an activation function, known to induce self-normalizing properties [27] in the neural networks improving stability and robustness.

We used the whole information as input features from the calorimeter simulation, being the configuration, which performed best in the deterministic case. Even if not strictly necessary to compute the KL divergence, the values of the parameters are sampled multiple times at each training step, to give more stability to the optimization process. Using three samples in each iteration provided a good trade-off between memory usage and performances.

In **Figure 9**, for the four most challenging classes of decay, we show improvements in terms of ROC curves and AUC of the Bayesian model over the deterministic one. **Table 4** shows the accuracies reached in the test phase for different number of samplings considering all the events and for different minimum confidence thresholds on the final prediction. We do not find much improvement in terms of overall accuracy and confidence of the model when using more than 10 samples at test time. In **Figure 10**, we compare confusion matrices on the test set applying no threshold or a 0.7 threshold on the minimum confidence of the prediction.

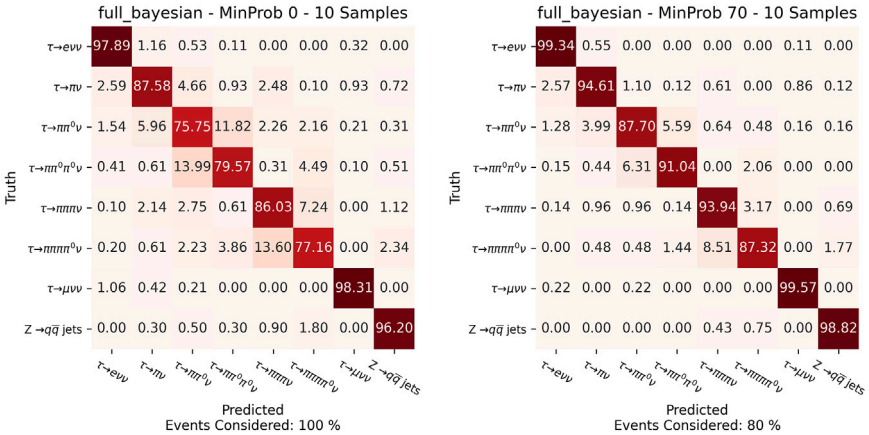


FIGURE 10 | Confusion matrices of the Bayesian DGCNN with 10 samples for each prediction without a threshold on minimum confidence and with a minimum confidence of 0.7. In the bottom, it is indicated the rate of events considered.

4 DISCUSSION

In this study, we propose a new technique based on the use of deep neural networks for the identification of tau lepton decays and show its performance when applied to the IDEA dual-readout calorimeter concept proposed for the future FCC-ee electron–positron collider. We have implemented a dynamic graph CNN architecture that takes in input only the raw information from the readout electronic of the calorimeter and leveraging the high granularity, and the different patterns produced by the dual readout is able to predict with excellent performance the specific tau lepton decay, discriminating also taus against signals produced by QCD jet events. An average accuracy of 91% in classifying different tau decays is obtained by using geometrical, energy, and time information from the calorimeter's fibers, while the accuracy for the specific decay modes ranges from 99% for the leptonic decays of the tau to 85–90% for the hadronic modes. The model is also able to discriminate with high-accuracy (> 95%) tau decays from jets from QCD. Even by using as input of the neural network only the information relative to the geometry and the type, scintillating or Cherenkov, of the firing fibers, the model is able to achieve average accuracies between 80 and 97% depending on the decay mode. The proposed algorithm is very flexible and can be easily extended in terms of additional tasks and additional input features, as for example, information from other detector systems. The ongoing work focuses on extending the capabilities of the algorithm by including information from other detector subsystem (inner tracking and muon spectrometer) and in identifying the individual contributions from the neutral and charged particles inside the tau and QCD jet clusters, to further improve the particle identification and particle flow reconstruction.

DATA AVAILABILITY STATEMENT

The datasets and the original code used for this study are available on request by contacting the corresponding authors.

AUTHOR CONTRIBUTIONS

SG proposed the original idea, coordinated the work, prepared the dataset, and performed the initial proof of concept implementation; LT implemented and trained the DGCNN model for the classification task and contributed to the data analysis; MD implemented the Bayes DGCNN model and contributed to the data analysis. All authors wrote the manuscript, contributed to the manuscript, and approved the submitted version.

FUNDING

This work was partially funded by the Istituto Nazionale di Fisica Nucleare, Sezione di Roma.

ACKNOWLEDGMENTS

We are greatly indebted to the IDEA dual-readout calorimeter concept group, especially the University and INFN Pavia and the University of Insubria groups, for the fundamental help in providing the simulated event samples used to train and test the algorithms and in understanding the specific features of the dual-readout calorimeter.

REFERENCES

1. Abada A, Abbrescia M, AbdusSalam SS, Abdyukhanov I, Fernandez JA, Abramov A, et al. Fcc-ee: the Lepton Collider. *Eur Phys J Spec Top* (2019) 228:261–623. doi:10.1140/epjst/e2019-900045-4
2. Antonello M, Caccia M, Ferrari R, Gaudio G, Pezzotti L, Polesello G, et al. Expected Performance of the Idea Dual-Readout Fully Projective Fiber Calorimeter. *J Inst* (2020) 15:C06015. doi:10.1088/1748-0221/15/06/c06015
3. ATLAS collaboration. Observation of a New Particle in the Search for the Standard Model Higgs Boson with the Atlas Detector at the Lhc. *Phys Lett B* (2012) 716:1–29. doi:10.1016/j.physletb.2012.08.020

4. Aad G, Abbott B, Abdallah J, Abdel Khalek S, Abdinov O, Aben R, et al. Identification and Energy Calibration of Hadronically Decaying Tau Leptons with the Atlas experiment in Pp Collisions at $\sqrt{s}$=8 Tev. *Eur Phys J* (20152015) C75:303. doi:10.1140/epjc/s10052-015-3500-z
5. Aad G, Abbott B, Abdallah J, Abdinov O, Aben R, AbouZeid OS, et al. Measurements of the Higgs Boson Production and Decay Rates and Coupling Strengths Using Pp Collision Data at $\sqrt{s}$=7 and 8 Tev in the Atlas experiment. *The Eur Phys J C* (2016) 76:1–51. doi:10.1140/epjc/s10052-015-3769-y
6. Aaboud M, Aad G, Abbott B, Abdinov O, Abeloos B, Abidi SH, et al. Measurement of τ Polarisation in $Z/\gamma^{\star} \rightarrow \tau\tau$ Decays in Proton–Proton Collisions at $\sqrt{s}$ = 8 Tev with the Atlas Detector. *Eur Phys J C* (2018) 78: 163. doi:10.1140/epjc/s10052-018-5619-1
7. Aubert B, Karyotakis Y, Lees J, Poireau V, Prencipe E, Prudent X, et al. Searches for lepton flavor violation in the decays $\tau^{\pm} \rightarrow e^{\pm}\gamma$ and $\tau^{\pm} \rightarrow \mu^{\pm}\gamma$. *Phys Rev Lett* (2010) 104:021802. doi:10.1103/PhysRevLett.104.021802
8. Benedikt M, Mertens V, Zimmermann F, Cerutti F, Otto T, Poole J, et al. Fcc-ee: The Lepton Collider: Future Circular Collider Conceptual Design Report Volume 2. *Eur Phys J Spec Top* (2018) 228:261–623. doi:10.1140/epjst/e2019-900045-4
9. Blundell C, Cornebise J, Kavukcuoglu K, Wierstra D. Weight Uncertainty in Neural Network. In: International Conference on Machine Learning (PMLR) (2015). p. 1613–22.
10. CMS Collaboration (2015). Reconstruction and Identification of Tau Lepton Decays to Hadrons and Tau Neutrino at Cms. *arXiv preprint arXiv: 1510.07488*. doi:10.1088/1748-0221/13/10/P10005
11. CMS collaboration. Performance of Reconstruction and Identification of τ Leptons Decaying to Hadrons and $\nu\tau$ in Pp Collisions at $\sqrt{s}$ = 13 Tev. *JINST* (2018) 13:P10005.
12. Dam M. Tau-lepton physics at the fcc-ee circular $e^{+}e^{-}$ collider. *Sci Post Phys Proc* (2019) 41. doi:10.21468/SciPostPhysProc.1.041
13. Elagin A, Safonov A. Likelihood-based Particle Flow Algorithm at Cdf for Accurate Energy Measurement and Identification of Hadronically Decaying Tau Leptons. In: Proceedings of 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research — PoS(ACAT2010), 093 (2010). doi:10.22323/1.093.0046
14. Gal Y, Ghahramani Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In: international conference on machine learning (PMLR) (2016). p. 1050–9.
15. Galea C. Tau Identification at D0. *Acta Phys.Polon.* (2007) B38:769–78.
16. GEANT CollaborationAgostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, et al. Geant4–a Simulation Toolkit. *Nucl Instrum Meth A* (2003) 506: 250–303. doi:10.1016/S0168-9002(03)01368-8
17. Goan E, Fookes C. Bayesian Neural Networks: An Introduction and Survey. In: *Case Studies in Applied Bayesian Data Science*. Springer (2020). p. 45–87. doi:10.1007/978-3-030-42553-1_3
18. Guo C, Pleiss G, Sun Y, Weinberger KQ. On Calibration of Modern Neural Networks. In: International Conference on Machine Learning (PMLR) (2017). p. 1321–30.
19. Hayasaka K, Inami K, Miyazaki Y, Arinstein K, Aulchenko V, Aushev T, et al. Search for Lepton-Flavor-Violating τ Decays into Three Leptons with 719 Million Produced τ+ τ- Pairs. *Phys Lett B* (2010) 687:139–43. doi:10.1016/j.physletb.2010.03.037
20. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (2016). p. 770–8. doi:10.1109/cvpr.2016.90
21. Heister et al. A. Measurement of the Tau Polarisation at Lep. *Eur Phys J C* (2001) 20:401–30. doi:10.1007/s100520100689
22. Heldmann M, Cavalli D. An Improved Tau-Identification for the ATLAS experiment. *Tech Rep* (2005).
23. Hernández-Lobato JM, Adams R. Probabilistic Backpropagation for Scalable Learning of Bayesian Neural Networks. In: International conference on machine learning (PMLR) (2015). p. 1861–9.
24. High Luminosity LHC Project (2021). The Hl-Lhc Project. [Online; accessed 1-February-2022]
25. Innocente V, Wang YF, Zhang ZP. Identification of Tau Decays Using a Neural Network. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (1992) 323:647–55. doi:10.1016/0168-9002(92)90011-r
26. Jospin LV, Buntine W, Boussaid F, Laga H, Bennamoun M (2020). Hands-on Bayesian Neural Networks–A Tutorial for Deep Learning Users. *arXiv preprint arXiv:2007.06823*
27. Klambauer G, Unterthiner T, Mayr A, Hochreiter S. Self-normalizing Neural Networks. In: Proceedings of the 31st international conference on neural information processing systems (2017). p. 972–81.
28. Levy S (2005). *Tau Identification at the Tevatron*. Tech. rep., Fermi National Accelerator Lab.(FNAL), Batavia, IL (United States). doi:10.1007/978-3-540-32841-4_20
29. Li Y, Bu R, Sun M, Wu W, Di X, Chen B. Pointcnn: Convolution on X-Transformed Points. *Adv Neural Inf Process Syst* (2018) 31. doi:10.48550/arXiv.1801.07791
30. Lindfeld L. Tau Leptons at Hera. *Nucl Phys B - Proc Supplements* (2005) 144: 315–22. doi:10.1016/j.nuclphysbps.2005.02.042
31. Neal RM, Mcmc Using Hamiltonian Dynamics. *Handbook of markov chain monte carlo* (2011) 2:2. doi:10.1201/b10905-6
32. Perl ML. The Tau Lepton. *Rep Prog Phys* (1992) 55:653–722. doi:10.1088/0034-4885/55/6/001
33. Proserpio E, Santoro R. *SimSiPM: A Library for SiPM Simulation*. Como, Italy (2021).
34. Qi CR, Su H, Mo K, Guibas LJ. Pointnet: Deep Learning on point Sets for 3d Classification and Segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition (2017). p. 652–60.
35. Qu H, Gouskos L. Jet Tagging via Particle Clouds. *Phys Rev D* (2020) 101: 056019. doi:10.1103/physrevd.101.056019
36. Safonov AN. Physics with Taus at Cdf. *Nucl Phys B - Proc Supplements* (2005) 144:323–32. doi:10.1016/j.nuclphysbps.2005.02.043
37. Sjöstrand T, Mrenna S, Skands P. A Brief Introduction to Pythia 8.1. *Comput Phys Commun* (2008) 178:852–67. doi:10.1016/j.cpc.2008.01.036
38. Wang Y, Sun Y, Liu Z, Sarma SE, Bronstein MM, Solomon JM. Dynamic Graph Cnn for Learning on point Clouds. *ACM Trans Graph* (2019) 38:1–12. doi:10.1145/3326362
39. Wigmans R. The DREAM Project-Results and Plans. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2007) 572: 215–7. doi:10.1016/j.nima.2006.10.211
40. Wigmans R, New Results from the Rd52 Project. *Nucl Instr Methods Phys Res Section A: Acc Spectrometers, Detectors Associated Equipment* (2016) 824: 721–5. doi:10.1016/j.nima.2015.09.069
41. Wu W, Qi Z, Fuxin L. Pointconv: Deep Convolutional Networks on 3d point Clouds. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2019). p. 9621–30. doi:10.1109/cvpr.2019.00985
42. Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS. A Comprehensive Survey on Graph Neural Networks. *IEEE Trans Neural Netw Learn Syst* (2020) 32:4–24. doi:10.1109/TNNLS.2020.2978386
43. Zhou J, Cui G, Hu S, Zhang Z, Yang C, Liu Z, et al. Graph Neural Networks: A Review of Methods and Applications. *AI Open* (2020) 1:57–81. doi:10.1016/j.aiopen.2021.01.001
44. Zhou Z-H. *Ensemble Methods: Foundations and Algorithms*. New York, NY: CRC Press (2012).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Giagu, Torresi and Di Filippo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

REVIEW
published: 09 August 2022
doi: 10.3389/fphy.2022.933479

RF Technologies for Future Colliders

Sergey Belomestnykh *

Fermi National Accelerator Laboratory, Batavia, IL, United States

Particle colliders remain indispensable scientific instruments to discover and study new elementary particles and fundamental forces of nature. Whether the collider is a factory (used to improve precision of measuring properties of already discovered particles or to enable studies of rare decay channels), an energy frontier machine (aimed at discovering new particles and forces), a heavy ion collider (allowing studies of what the universe looked like in the early moments after its creation), or an electron-hadron collider (where electrons are used for probing heavy ions or protons to study the fundamental force binding all visible matter), the radio frequency technologies play a key role in enabling the machine to reach its goals. This article considers challenges presented to the radio frequency technologies by the next generation of particle colliders and reviews R&D approaches and directions to address these challenges.

Keywords: acceleration, cavity, collider, klystron, linac, radio frequency, storage ring, superconductivity

OPEN ACCESS

Edited by:
Patrizia Azzi,
National Institute of Nuclear Physics of Padova, Italy

Reviewed by:
Serkant Ali Çetin,
University of Istinye, Turkey
Jie Gao,
Institute of High Energy Physics (CAS), China
Akira Yamamoto,
High Energy Accelerator Research Organization, Japan

***Correspondence:**
Sergey Belomestnykh
sbelomes@fnal.gov

Specialty section:
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

Received: *30 April 2022*
Accepted: *27 May 2022*
Published: *09 August 2022*

Citation:
Belomestnykh S (2022) RF Technologies for Future Colliders. Front. Phys. 10:933479. doi: 10.3389/fphy.2022.933479

1 INTRODUCTION

The quest for fundamental laws of nature requires studying interactions of elementary particles at smaller and smaller scales, shorter and shorter time intervals, and hence higher and higher energies. The colliders for high-energy physics and nuclear physics experiments are arguably the main scientific tools that allow such studies. However, each step to higher interaction energies demands building bigger and more technologically challenging accelerators. While over the last few decades other application of accelerators, for example, free electron lasers and storage-ring-based light sources began to present new challenges, the colliders remain the main driving force behind advancing accelerator technologies. Large accelerators are costly to build and operate. So, in addition to technological challenges, there are always efforts to develop cost-saving technologies for building colliders and making them more efficient to operate. For a comprehensive review of modern and future colliders, the readers are referred to a recent article [1].

Radio frequency (RF) is one of the key technologies for modern particle accelerators [2, 3]. The main function of radio frequency systems in colliders is to accelerate particles either to increase their energy or to compensate the energy loss, for example, due to synchrotron radiation and thus maintain the energy of the experiment constant. The acceleration is achieved *via* interaction of the particle beam with the time-varying longitudinal electric field in an accelerating structure. The particles traverse the accelerating structure along its axis, where the electric field of the fundamental mode is at maximum. Resonant frequencies of the collider's RF structures are in the range from tens of megahertz to tens of gigahertz. Some colliders use a "crab-crossing" collision scheme when instead of head-on, the counter-rotating bunches are colliding at a small crossing angle [4]. To avoid the geometrical loss of luminosity, special RF deflecting cavities, called crab cavities, are used to "chirp" the bunches before the collision so they are oriented head-on at the interaction point and then "de-chirp" them afterward.

In addition to accelerating and crabbing/deflecting structures, the RF systems include auxiliary or peripheral sub-systems and components. Among those are fundamental RF power couplers,

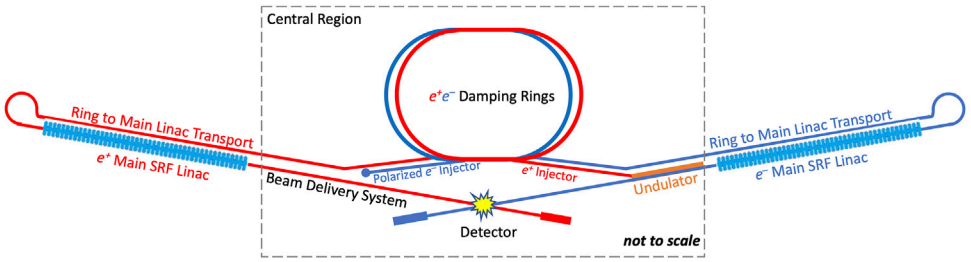


FIGURE 1 | Layout of a "generic" SRF linear collider.

higher-order mode couplers, resonance control sub-systems, RF power amplifiers and distribution sub-systems, low level RF controls, and some others.

In this article, first and foremost, the challenges faced by RF technologies of the next-generation particle colliders are addressed, those sufficiently developed to be built in the next few decades. The challenges of linear and circular colliders are distinctly different for the most part and will be considered separately. However, some RF technologies, such as crab cavities and high-efficiency, low-cost klystrons have enough commonality for both collider types, so they are considered as common RF technologies. Finally, specific challenges put forward by muon collider concepts are touched upon.

While the mainstream technologies are the focus of this article, some recent exotic (if not far-fetched) technological proposals as well as challenges of some collider concepts that require longer-term R&D efforts will be briefly mentioned.

2 RADIO FREQUENCY TECHNOLOGIES FOR FUTURE e^+e^- LINEAR COLLIDERS

In a linear collider configuration, two beams are accelerated in two linear accelerators (linacs) and then transported to a collision point inside a detector, as shown in **Figure 1**. The luminosity of a symmetric (two beams having equal energies) linear collider is given by [5].

$$L = \frac{n_b N_e^2 f_{rep}}{4\pi\sigma_x^*\sigma_y^*} \cdot H_D = \left(n_b N_e f_{rep}\right)\frac{N_e}{4\pi\sigma_x^*\sigma_y^*} \cdot H_D, \tag{2.1}$$

where n_b is the number of bunches per pulse, N_e is the number of electrons (positrons) per bunch, f_{rep} is the repetition frequency, $\sigma^*_{x(y)}$ is the horizontal (vertical) beam size at the interaction point, and H_D is the disruption enhancement factor, typically ≈ 2. The term in parenthesis is essentially an average beam current I_{beam}, which can be expressed as

$$n_b N_e f_{rep} = \frac{I_{beam}}{e} = \frac{P_{beam}}{E_{c.m.}}, \tag{2.2}$$

with e being the elementary charge, P_{beam}—the average beam power in two beams, and $E_{c.m.}$—the center-of-mass energy. Then, we can re-write the formula for luminosity as

$$L = \frac{P_{beam}}{E_{c.m.}} \cdot \frac{N_e}{4\pi\sigma_x^*\sigma_y^*} \cdot H_D. \tag{2.3}$$

As one can see, a high beam power and very small spot sizes at the collision point are required to obtain high luminosity. Thus, achieving a high efficiency of transferring RF to beam power and wall-plug to RF power is very important to colliders. The overall site wall-plug power is usually limited to a few hundreds of megawatts, while the wall-plug to beam power efficiency for linear colliders is of the order of 10%. This constraint necessitates running of linear colliders in a pulsed mode.

On the other hand, it is desirable to reach the collision energy with a linac of a reasonable length. Hence, the RF structures must provide as high rate of acceleration as possible while still maintaining reliable operation. The pulsed mode of operation is beneficial for reaching this goal.

Two types of linear electron–positron colliders under consideration are based on either superconducting or normal-conducting RF accelerating structures. Both have a long history of technology development: a superconducting RF (SRF) linear collider was first proposed in 1965 by Tigner [6]; colliders based on normal-conducting RF accelerating structures are under discussion since the 1970s [7]. The only linear collider realized so far was the Stanford Linear Collider (SLC) at SLAC [8], based on a normal-conducting S-band linac. Among the potential next-generation SRF colliders are the International Linear Collider (ILC) [9], recently proposed HELEN collider [10], and two concepts based on energy recovery linacs: ERLC [11] and ReLiC [12]. The normal-conducting RF linacs are utilized in the Compact Linear Collider (CLIC) [13] and Cool Copper Collider (C^3) [14] proposals. In this section, state of the art for the two linear collider technology options and ongoing R&D are considered.

2.1 Superconducting Radio Frequency Linear Colliders

Development of the SRF technology for a linear collider had begun in the 1980's with a world-wide effort which focused on the TESLA collider proposal [15]. By early 2000's, the TESLA technology was well-developed already [5, 16], and the SRF option was eventually selected for the ILC. However, the SRF technology continued to make advances to improve the cavity performance, and new e^+e^- linear collider proposals (including

FIGURE 2 | 1.3 GHz nine-cell SRF TESLA cavity made of bulk niobium. Courtesy of linearcollider.org.

ILC upgrades) have been put forward recently that would utilize the most recent achievements in the SRF technology.

2.1.1 Baseline Superconducting Radio Frequency Technology for the International Linear Collider

The ILC has been the prime candidate for the next e^+e^- collider, especially since the discovery of the Higgs boson in 2012. The machine was baselined in 2013 [17, 18] and is under consideration to be hosted by Japan [9]. At the first stage, the machine is proposed to operate as a Higgs factory with a center-of-mass energy of 250 GeV. Each of the two Main Linacs accelerate beams from 5 to 125 GeV. The linacs are based on the TESLA technology: 1.3 GHz nine-cell SRF cavities made of bulk niobium (**Figure 2**), operating at 2 K with an intrinsic cavity quality factor of $1{\cdot}10^{10}$ and providing accelerating gradients of 31.5 MV/m. Because the power dissipation in the cavity walls is extremely small, the accelerating field can be produced with long, ~0.7 ms, low peak power RF pulses, and a high RF-to-beam-power transfer efficiency, even considering the need to operate at 2 K.

The TESLA cavities are enclosed in helium vessels, which are filled with 2 K superfluid liquid helium during operation and equipped with fundamental RF power couplers, frequency tuning mechanisms, magnetic shield, and other peripherals. These "dressed" cavities are then assembled into cryomodules comprising either nine cavities or (every third module) eight cavities plus a quadrupole/corrector/beam position monitor unit and all necessary cryogenic supply lines, heat shields, multilayer insulation, etc. Nine cryomodules are powered by two 10-MW L-band multibeam klystrons, thus forming two RF units or one cryostring. RF sources meeting the ILC specifications are commercially available from two vendors, both of which provided klystrons for the European XFEL. The ILC specifications ask for a 65% RF efficiency (drive beam to output RF power). In total, 987 cryomodules will be installed for operation at 250 GeV.

The baseline ILC SRF technology is well-established and was used to build several SRF linacs, with free electron lasers (FELs) European XFEL at DESY, Hamburg, Germany [19], and LCLS-II at SLAC in the United States [20] being the biggest ones. More SRF linacs utilizing this technology are on the horizon, for example, LCLS-II-HE—the high-energy upgrade of LCLS-II at SLAC—and SHINE FEL in China. The European XFEL linac is the largest application of the ILC SRF technology to date. The performance of European XFEL cavities is close to the requirements of the ILC Technical Design Report [17, 18]. For example, 420 cavities from one vendor, which followed the ILC cavity treatment recipe, reached an average gradient of 33.0 ± 6.5 MV/m. More than 10% of cavities from this vendor exceeded 40 MV/m. It is extremely rare that demonstration of a core technology for a future machine exists at such a large scale. Furthermore, ILC-related R&D efforts produced additional proofs of technology readiness. The studies of the cavity performance yield resulted in a (94 ± 6)% yield for cavities with accelerating gradients greater than 28 MV/m and (75 ± 11)% for 35 MV/m. The ensemble of cavities had an average gradient of 37.1 MV/m. Two cryomodules, one at FNAL [21] and one at KEK [22], were tested with beams reaching and even exceeding the ILC specifications. Thus, the baseline ILC SRF technology has been demonstrated and industrialized on a mid-scale and is ready to be deployed.

2.1.2 Superconducting Radio Frequency Advances for International Linear Collider Upgrades, HELEN Collider, and Other Applications

Since the development of baseline ILC technology, the SRF field has continued to make progress in several areas. The key directions for SRF cavity R&D to support future needs of high-energy physics are outlined in the recently published article [23]. Some of the recent advances and results expected in the near future can be applied to the ILC luminosity and energy upgrades [9, 24, 25] or to the recently proposed HELEN collider [10]. Some of the developments discussed below are relevant to other future colliders that utilize SRF technology.

It is important to note that the improvements in gradients should be accompanied by improvements in cavity quality factors to avoid excessively high cavity wall losses and associated thermal effects. This follows from the expression for RF power dissipation in the cavity walls P_c

$$P_c = \frac{V_c^2}{R/Q \cdot Q_0} = \frac{(E_{acc}L_c)^2}{R/Q \cdot Q_0}, \tag{2.4}$$

where $V_c = E_{acc}L_c$ is the cavity voltage, E_{acc} is the accelerating gradient, L_c is the effective cavity length, R/Q is the specific shunt impedance (depends only on the cavity geometry, not dimensions), and Q_0 is the cavity intrinsic quality factor. However, improving the cavity quality factor is as important for large SRF systems operating at medium, 16–25 MV/m, gradients. Higher Q factors allow both capital (fewer cryoplants such as in the case of LCLS-II SRF linac [20]) and operational savings (reduced cryogenic load). It is customary to present the SRF cavity performance as a plot of its intrinsic quality factor versus accelerating gradient.

Superconducting properties of niobium—including its surface resistance and breakdown magnetic field—are determined by the state of material within first few tens of nanometers of the surface subjected to the RF field, as the magnetic field penetration depth for clean niobium is about 40 nm. Studies demonstrate that depending on the surface treatment applied to SRF cavities, their surface resistance (and hence quality factor) can vary by almost an order of magnitude. The surface treatment also affects the maximum accelerating gradient that the cavities can operate at.

There was a rapid progress in developing advanced surface treatments over the last decade [23]. Most notably, a nitrogen

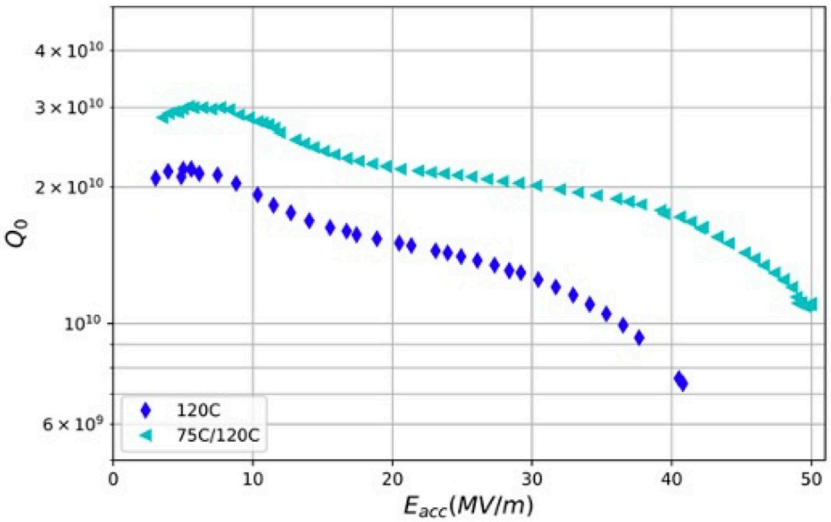


FIGURE 3 | Q vs. E curve of single-cell cavity reaching 49 MV/m from cold EP/optimized baking (75/120°C) compared to the curve of a cavity prepared by the standard ILC recipe.

doping technology allowed reaching unprecedented high-quality factors greater than $3 \cdot 10^{10}$ at 2 K and medium gradients for free-electron laser linacs LCLS-II and LCLS-II-HE [26] (1.3 GHz cavities) and proton SRF linac PIP-II (650 MHz cavities) [27]. Another promising surface treatment under development is mid-temperature (~300°C) baking of cavities in vacuum [28]. These methods could be adapted—after developing cavity-specific procedures—for collider SRF systems operating at medium gradients, such as CEPC (650 MHz), FCC-ee (800 MHz), ERLC (1.3 GHz), or ILC crab cavities. For high-gradient structures of ILC and HELEN, two low-temperature surface treatments demonstrated higher accelerating gradients and quality factors than those achieved with the standard ILC treatment. Thus, developing advanced surface treatments tailored to different applications remains an important direction in the SRF cavity R&D for future colliders.

In the next sections, a brief review of three possible pathways to achieve higher accelerating gradients in SRF cavities is given.

2.1.2.1 Higher Gradients in Bulk Niobium Standing Wave Accelerating Structures

As mentioned above, over the last decade, a remarkable progress was achieved in improving quality factors and accelerating gradients of bulk niobium cavities. In a recent development, a quench field near 50 MV/m for 1.3 GHz niobium TESLA-shaped SRF single-cell cavities has been achieved with a new 75/120°C two-step bake treatment [29], as shown in **Figure 3**. The statistics of over 50 cavity tests shows gradients in the range 40–50 MV/m with an average value of 45 MV/m [30].

The peak surface magnetic field H_{pk} presents the hard ultimate limit to the performance of SRF cavities when reaching the critical superheating field of Nb. As the ratio H_{pk}/E_{acc} is determined by the cavity geometry, there are several geometries (details of some of these geometries can be found in [31]) that lower this ratio by 10%–20% relative to the TESLA cavity by allowing E_{pk}/E_{acc} increase up to 20%. At the same time, the new geometries have $R/Q \cdot G$ values higher by about 30%–40%; hence, their cryogenic losses will be reduced proportionally. $R/Q \cdot G$ is the figure of merit that depends only on the cavity shape. It allows comparing the efficiency of different geometries by separating the geometry-only parameters from the surface resistance R_s of cavity material. $G = Q_0 R_s$ is called the geometry factor. For clarity, we can re-write (2.4) as

$$P_c = \frac{V_c^2 \cdot R_s}{G \cdot R/Q}. \tag{2.5}$$

Combining the two-step bake with one of the advanced cavity shapes, one can possibly improve the accelerating gradients up to ~60 MV/m. This combination of advanced cavity treatments and cavity shapes still must be tried and demonstrated, but if successful, it will provide a relatively straightforward path to operating SRF linear colliders at higher gradients with standing wave structures.

2.1.2.2 Traveling Wave Superconducting Radio Frequency Structures

Traveling wave (TW) structures [32] offer several main advantages compared to standing wave ones: substantially lower peak magnetic field, lower peak electric field, and substantially higher R/Q. To achieve the highest possible accelerating gradient, the main emphasis of design optimization is to lower H_{pk}/E_{acc} as much as possible. For example, an optimized TW structure with a phase advance of π/2 has this parameter almost a factor of 1.5 lower than for the TESLA cavity [33]. If results for the best single-cell TESLA shape cavities prepared today (E_{acc} = 49 MV/m, H_{pk} = 2090 Oe) can be repeated in this TW structure, one can optimistically expect an accelerating gradient $E_{acc} > 70$ MV/m.

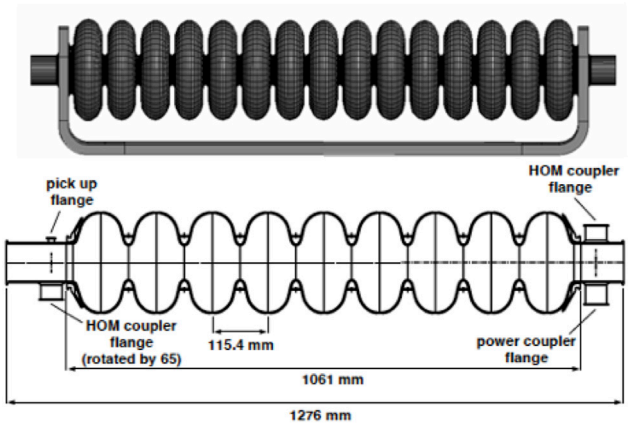


FIGURE 4 | Traveling wave structure with a 105° phase advance per cell (top) [34] compared to the one-meter standing wave TESLA structure [16].

A TW structure provides high stability of the field distribution along the structure with respect to geometrical perturbations. This allows for much longer accelerating structures than TESLA cavities. However, as we can see in **Figure 4**, an example TW structure with a phase advance of 105 degrees requires almost twice the number of cells per meter—as compared to the TESLA structure—and needs a feedback waveguide for redirecting power from the end to the front of the accelerating structure. The feedback requires careful tuning to compensate reflections along the TW ring and thus obtain a pure traveling wave regime. The challenges of fabricating and surface treatment remain to be addressed to demonstrate feasibility of traveling wave SRF structures, and efforts are underway for a proof-of-principle demonstration [35]. The potential payback of developing full-scale traveling wave SRF cavities is big. It would pave the way for the HELEN collider and/or ILC upgrades to center-of-mass energies beyond 1 TeV.

2.1.2.3 Advanced Superconducting Radio Frequency Materials

All significant SRF cavity performance results for linear colliders have been achieved with structures made of bulk niobium. There are two main reasons: 1) Nb has the highest critical temperature and critical magnetic field of all elemental superconductors, and 2) it is a ductile metal that can easily be formed into complex geometries of SRF cavities. With niobium cavity gradients approaching the fundamental limit (determined by the Nb superheating magnetic field $H_s \approx 2,200\,Oe$), R&D on alternative superconductors becomes more and more important [36]. There are several materials with higher critical temperature and critical field that could potentially surpass the niobium technology for SRF applications. These materials cannot be used in bulk form due to their poor mechanical (ceramic-like) and thermal properties (poor thermal conductivity). They must be deposited on a suitable substrate—often Nb or Cu—as either thick or thin films. As application of most alternative superconductors to the SRF is still far in the future, it is suitable for the readers to be referred to review articles [36, 37]. Here we only briefly address Nb_3Sn as an alternative superconductor closest to practical use [38].

In cavity studies, Nb_3Sn has a critical temperature of ~18 K, twice that of Nb, and has already shown high $Q \sim 10^{10}$ even at relatively high temperatures around 4 K. This represents a significant cryogenic infrastructure and operation cost savings for future machines. Single-cell cavities have reached 24 MV/m and 9-cell TESLA cavities have reached 15 MV/m [39]. There have already been proposals to use this technology, for example, for compact industrial accelerators [40]. However, the full potential of Nb_3Sn is significantly higher: based on its superconducting properties, the fundamental limit for Nb_3Sn should be approximately 100 MV/m. Experiments suggest that even high performing Nb_3Sn cavities are currently limited by surface defects. There are promising directions, including electropolishing, oxypolishing, and mechanical polishing, and new deposition methods that could lead to inherently smooth and defect-free films, and improved surfaces are expected to lead to improved gradients. Reaching even 60%–70% of the full potential of Nb_3Sn would be extremely beneficial for enabling advanced SRF linear colliders.

2.1.3 Other Superconducting Radio Frequency Linear Collider Concepts

Two SRF linear collider proposals based on the energy recovery linac (ERL) concept were published recently: ERLC [11] and ReLiC [12]. While both are still early concepts and require more detailed studies, their main challenges for RF technologies are briefly discussed. The main advantage of using an ERL is that most of the beam energy is recovered after interactions in the detector and re-used for acceleration of the fresh beam. As ERLs operate in either continuous wave (CW) or high-duty factor mode, they promise to deliver much higher luminosities. However, the ERL-based machines face significant challenges that must be addressed when the concepts are developed further. Here, a few of them are listed. First, to keep the cryogenic losses at a reasonable level during CW operation, the accelerating gradients in these two machines are chosen at ~20 MV/m, which makes the SRF linac significantly longer than the ILC baseline at 31.5 MV/m. Second, the CW beam will produce large higher-order mode (HOM) power. This must be dealt with using beam line HOM absorbers [2], which will further lengthen the linacs. Third, both concepts would need much more expensive SRF cavities compared to the ILC. The ERLC will utilize either twin 1.3 GHz SRF cavities in a single linac or two parallel linacs connected by RF couplers. In either case, the number of cavities increases twofold. ReLiC plans to use five-cell 500 MHz cavities with strong HOM damping, which are much larger than the nine-cell 1.3 GHz TESLA cavities of ILC. These and other challenges make the ERL-based concepts more expensive and power-hungry than ILC, HELEN, or normal conducting linear colliders CLIC and C^3. Nevertheless, due to their potential higher luminosities, these two concepts need further detailed scientific and technical validations to address the challenges.

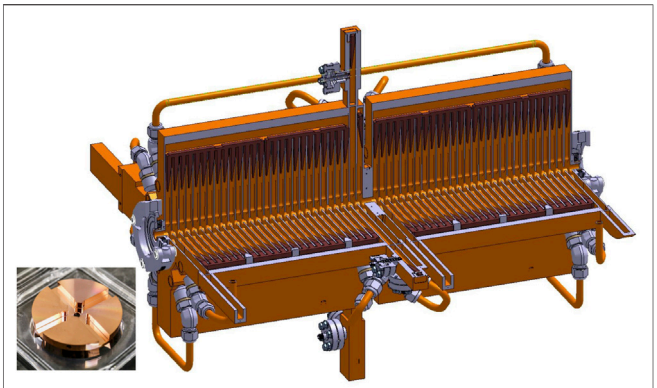


FIGURE 5 | Left: The basic assembly block of the CLIC structure with four waveguides providing HOM damping. Right: Assembly drawing of the double-structure acceleration unit (Image credit: CLIC) [13].

2.2 Normal Conducting Linear Colliders

Two normal conducting linear colliders are being actively developed at present: Compact Linear Collider (CLIC) [13] and Cool Copper Collider (C^3) [14]. To limit average power dissipation in the cavity walls at high accelerating gradients, normal conducting structures must operate in the pulsed mode with very short pulses. In addition, the RF power loss per unit length in normal conducting cavities scales as $P_c/L \sim 1/\sqrt{\omega}$, thus making higher frequencies preferable. In this section, RF technologies and challenges related to these two concepts are considered.

2.2.1 Radio Frequency for the Compact Linear Collider

The CLIC [13] has been optimized for three center-of-mass energies: 380 GeV, 1.5 TeV, and 3 TeV. It utilizes a novel two-beam acceleration technique. The main linacs accelerate particles from 9 GeV to the collision energy in normal-conducting X-band accelerating structures operating in the range of 70–100 MV/m. The RF power for each main linac is provided by a high-current, low-energy drive beam that runs parallel to the colliding beam through a sequence of power extraction and transfer structures (PETS). The drive beam generates RF power in the PETS that is then transferred to the accelerating structures using a waveguide network. The drive beam is generated centrally with a fundamental frequency of 1 GHz. This concept significantly reduces the cost and power consumption compared with powering the structures directly by klystrons.

CLIC uses 12 GHz traveling wave accelerating structures with a tapered inner aperture diameter ranging from 8.2 to 5.2 mm and are approximately 25 cm in length. The copper structures are assembled from disks machined with micron precision that are joined together using diffusion bonding. HOM suppression is provided by a combination of heavy damping *via* four terminated waveguides connected to each cell and detuning accomplished through the iris aperture tapering. **Figure 5** illustrates the micron-precision disk (the basic assembly block of the CLIC structure) and an assembly drawing of the acceleration unit.

For the collision energy of 380 GeV, the main linac structures accelerate a train of bunches with a gradient of 72 MV/m and must maintain a breakdown rate of less than $3{\cdot}10^{-7}$ m^{-1}. The pulse length is 233 ns, and the repetition rate is 50 Hz. The linacs operate at very high beam loading to enable high beam current and high RF-to-beam efficiency. The performance aspects of the structures have been validated in a series of dedicated tests. The tests included an experiment at CTF3 at CERN to determine the effect of heavy beam loading. Fully assembled two-beam modules have been tested with and without the beam. All elements for the waveguide system have been designed, fabricated, and operated to full specifications. All these efforts demonstrated that the CLIC RF design parameters are well-understood and can be reliably reproduced in tests. Further studies will put an emphasis on optimizing cost and energy efficiency of the RF system.

2.2.2 Optimized C-Band RF Structures for C^3

C^3 [14] is an e^+e^- linear collider based on a novel approach to normal conducting linacs that achieves both high gradient and high efficiency. C^3 would be a Higgs factory operating at 250 GeV with a straightforward upgrade path to 550 GeV, while staying on the same short facility footprint.

There are several distinct features that set the C^3 accelerating structures apart from traditional electron linacs. First, the cavity cells are optimized for high shunt impedance and low surface electric and magnetic fields. This necessarily led to a spherical reentrant geometry with "nose cones" that is typical, for example, for single-cell cavities of synchrotrons and storage rings [41]. However, such cells have beam apertures that are too small for sufficient cell-to-cell coupling of the fundamental RF mode, thus precluding on-axis coupling of RF power to the structure. Second, to circumvent this, a distributed coupling scheme (previously used at other laboratories in different configurations, see for example, [42, 43]) *via* parallel manifold RF waveguides provides side-coupling into each cell with the proper RF phase. **Figure 6** shows a 3D model of the proposed C-band (5.712 GHz) structure and magnitude of electric and magnetic field in the accelerating cell. While this structure is relatively complex, it was demonstrated that it could be machined in two

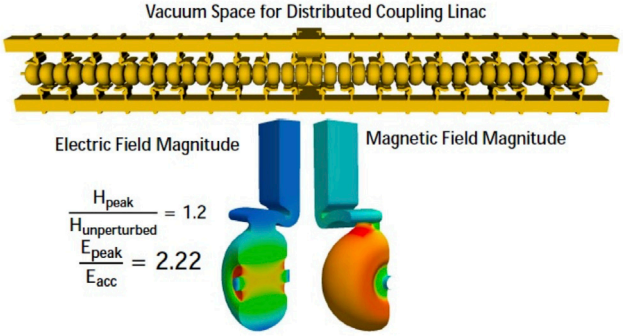


FIGURE 6 | Top: C^3 accelerating structure. Bottom: 3D model of a 1-m-long 40-cell C-band accelerating structure operating in the π mode. The magnitude of the electric and magnetic field in each cell. The peak surface electric field to accelerating gradient ratio is 2.22. The perturbation to the magnetic field from the RF coupler increases the peak magnetic field by 1.2 [14].

halves (or four quarters) by a CNC milling machine [44]. This fabrication process provides ultra-high-vacuum (UHV) quality surfaces that need no further finishing apart from a standard copper surface etch. Finally, the copper structures will operate in a liquid nitrogen bath at a cryogenic temperature of about 80 K. The advantage is two-fold: the increase of electrical conductivity and improvement of the material strength. The increased electrical conductivity results in reduction of RF losses in the cavity walls and the required RF power. The lower thermal stresses in the material and improved material strength reduce the probability of breakdown. Theoretical predictions of the copper surface resistance, confirmed by the experiment [45], show that almost all the improvement (factor of 2.55 compared to 300 K for RRR = 400 copper at 5.712 GHz) is achieved by cooling the material down to liquid nitrogen temperature.

The main linac will consist of one-meter-long structures operating at an accelerating gradient of 70 MV/m with pulse length of 700 ns and repetition rate of 120 Hz. For 550 GeV upgrade, the gradient will be raised to 120 MV/m with a shorter, 250 ns, pulse length. Prototype one-meter structures have been fabricated and tested at high gradient and at cryogenic temperatures. The next step is to develop an HOM damped and detuned design to mitigate the effect of the long-range wakefields. This will be accomplished differently from the CLIC design. Detuning of HOMs will be achieved by modifying the geometry of each cavity, while maintaining constant frequency of the fundamental mode. For damping, longitudinal damping slots in quadrature will be added to the structure design. While for the Higgs factory operation, C^3 can use commercially available 50 MW C-band klystrons and modulators, it is highly desirable to develop RF sources with better efficiency and higher power (especially for the energy upgrade).

3 FUTURE CIRCULAR COLLIDERS

First lepton circular colliders began operating in mid-1960s, see overview of collider development in [1]. Since that time, all colliders except the SLC at SLAC have been of this type. The energy and luminosity of every new generation of the particle colliders grew steadily in tandem with advances in key accelerator technologies and beam physics techniques. All contemporary colliders consist of two rings storing counter-rotating beams that intersect at one or more interaction points. The two rings may store particles of the same type or particles and their anti-particles or particles of two different types. The collisions may be symmetric (two beams having the same energy) or asymmetric. The current frontier colliders are the 13-TeV Large Hadron Collider (LHC) at CERN and asymmetric 4 by 7 GeV e^+e^- Super B factory (Super-KEKB) at KEK in Japan.

Several new colliders are under consideration to be built in the next couple of decades. For hadrons, the luminosity of the LHC will be upgraded first to the HL-LHC to improve the discovery reach of the machine at center-of-mass energies up to 14 TeV. The next-generation energy frontier hadron machines would be ~100-km circumference colliders FCC-hh [46] at CERN and SppC [47] in China, both of which have a target energy of ~100 TeV. However, to reach their design energies, these colliders need to develop superconducting magnet technology beyond state of the art that would take some time. Thus, before building hadron colliders, both teams consider building symmetric circular e^+e^- colliders in the same tunnels first. These are called FCC-ee [48] and CEPC [49], respectively. They will be able to run in a range of center-of-mass energies from 91 GeV (Z-pole) to ~360 GeV ($t\bar{t}$). At these high energies, the electron (positron) beams will generate huge amounts of synchrotron radiation, which is the main limitation to the luminosity. To keep the total electrical power consumption limited to about 300 MW, the synchrotron radiation power is constrained to 30 MW per beam in the CEPC and to 50 MW per beam in FCC-ee. The synchrotron radiation power for a high-energy electron beam can be calculated as

$$P_{SR}[\mathrm{MW}] = I_{beam} \cdot \Delta E_{SR} = I_{beam}[\mathrm{A}] \cdot 0.0885 \cdot \frac{E_{beam}^4[\mathrm{GeV}]}{\rho[m]}, \tag{3.1}$$

where ΔE_{SR} is the energy loss per turn due to synchrotron radiation, E_{beam} is the beam energy, and ρ is the average bending radius of the ring. Accordingly, the beam current will have to be lowered with the collision energy going up. Running at Z-pole, the RF systems will encounter heavy beam loading from ampere-class beams. At the $t\bar{t}$ energy, the e^+e^- colliders will operate in a "high energy machine" regime with relatively small beam currents.

Now, for flat beams ($\sigma_x \gg \sigma_y$), one can write the formula for luminosity as

$$L = K \cdot \frac{E_{beam} \cdot I_{beam} \cdot \xi_y}{\beta_y^*} \cdot R_{HG}, \tag{3.2}$$

where K is a constant, ξ_y is the beam–beam tune shift parameter, β_y^* is the beta function at the interaction point, and R_{HG} is the hour-glass factor, which should be kept reasonably large (e.g., it is ~ 0.9 for FCC-ee [50]). Combining (3.1) and (3.2), we obtain that the luminosity is proportional to the synchrotron radiation power and drops as cube of the beam energy:

$$L\left[\mathrm{cm}^{-2}s^{-1}\right] = 2.45 \cdot 10^{33} \cdot P_{SR}[\mathrm{MW}] \cdot \frac{\rho[\mathrm{m}] \cdot \xi_y}{E_{beam}^3[\mathrm{GeV}] \cdot \beta_y^*[\mathrm{m}]} \cdot R_{HG}. \tag{3.3}$$

The synchrotron radiation power must be compensated by the collider's RF system. Thus, achieving a high efficiency of transferring RF power to beam and wall-plug to RF power is very important to colliders, resembling the case of linear colliders.

In addition to FCC-ee and CEPC, an ERL-based e^+e^- collider CERC was proposed recently [51]. Finally, there is an interest in electron-hadron colliders. In particular, an electron-ion collider EIC is being designed at BNL in the United States [52]. EIC will be built in the existing RHIC tunnel by adding three new rings, strong hadron cooling, and upgrading the injector complex.

All circular colliders will need advanced radio frequency systems that must be able to deliver high CW RF power to

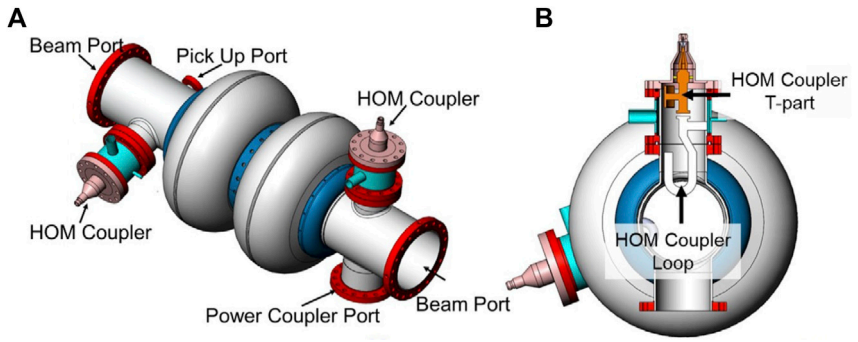


FIGURE 7 | CEPC cavity model with HOM couplers **(A)** and a cross-sectional view at the fundamental power port **(B)** [55].

ampere-class beams *via* strongly HOM-damped SRF cavities. In this section, the designs and challenges of several such cavities are reviewed.

3.1 Higher-Order Mode Damped Superconducting Radio Frequency Cavities for Circular Colliders

The first strongly HOM damped SRF cavities [53] were developed in the early 1990s for e^+e^- colliders CESR at Cornell in the United States and KEKB in Japan. These single-cell cavities are made of bulk niobium and operate at 500 and 509 MHz, respectively. Both cavities have very large beam pipes allowing HOMs to propagate toward water-cooled ferrite absorbers located outside the cryomodules. The 400-MHz LHC cavities use niobium-film-on-copper (Nb/Cu) technology, developed originally for LEP2 cavities. HOM damping is achieved with two types of coaxial HOM couplers attached to the large beam tubes. The first type is a narrow-band coupler (one per cavity) which damps the first two dipole modes, while two couplers of the second type are broad-band to damp all other HOMs [54]. These three HOM damped cavities have been state of the art for colliders and storage ring-based light sources since their development and implementation. The next-generation particle colliders require updated, custom designs.

CEPC will use two-cell 650 MHz cavities. Two coaxial HOM couplers with double-notch filters are installed on both sides of the cavity, as shown in **Figure 7**, with orientation optimized for best damping [55]. 240 such cavities will be installed for the Higgs operation. For the $t\bar{t}$ operation of the collider, additional 240 five-cell cavities will be installed to bring the total RF voltage to 10 GeV [49].

The conceptual design of FCC-ee envisages using cavities at two frequencies. Fifty-two single-cell, LHC-like, 400 MHz Nb/Cu cavities will be used at the Z-pole energy. At the WW threshold and ZH modes of operation, four-cell 400 MHz Nb/Cu cavities (52 and 136, respectively) will replace the single-cell cavities. Five-cell 800 MHz cavities made of bulk niobium will be added to boost the required RF voltage to 11.25 GV at the highest energy, where HOM damping requirements are relaxed, and smaller cavities would provide better efficiency and cost savings. The 400 MHz cavities would operate at 4.5 K and the 800 MHz cavities—at 2 K. A new design concept of a two-cell 600 MHz superconducting Slotted Waveguide Elliptical cavity (SWELL) was recently proposed as a potential option for FCC-ee [56]. The design (**Figure 8**) borrows the idea from normal conducting structures where a similar approach was realized, for example, in Slotted Irises Constant Aperture (SICA) 3 GHz accelerating structure used in CTF3 [57]. It is proposed that the SWELL cavity quadrant would be machined from copper and then coated with a thin film of niobium. While computer simulations show good HOM damping, the design looks quite exotic for a superconducting cavity. There are still many technological challenges to be addressed—fabrication tolerances, niobium coating, clean assembly, tuning, etc.—before one can discuss practicality of this proposal.

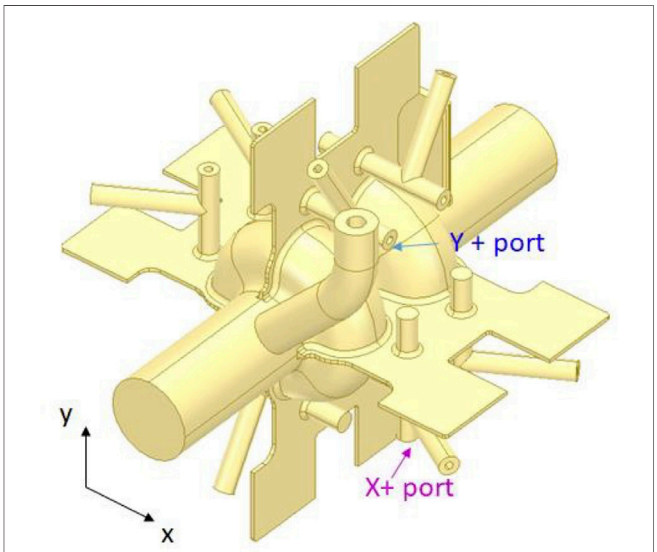


FIGURE 8 | 3D model showing two coaxial HOM ports located on the orthogonal branches of the SWELL cavity [56].

Several new SRF and normal conducting cavities of different types will be installed in the new EIC accelerators at BNL [52]. Consider, for example, the electron storage ring, which will

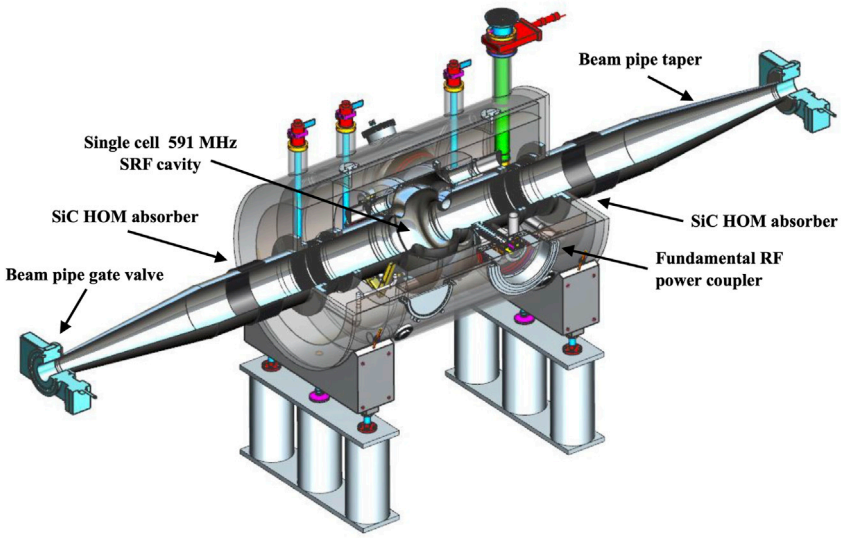


FIGURE 9 | Cut-out view of the EIC electron storage ring SRF cryomodule concept (adapted from [52]).

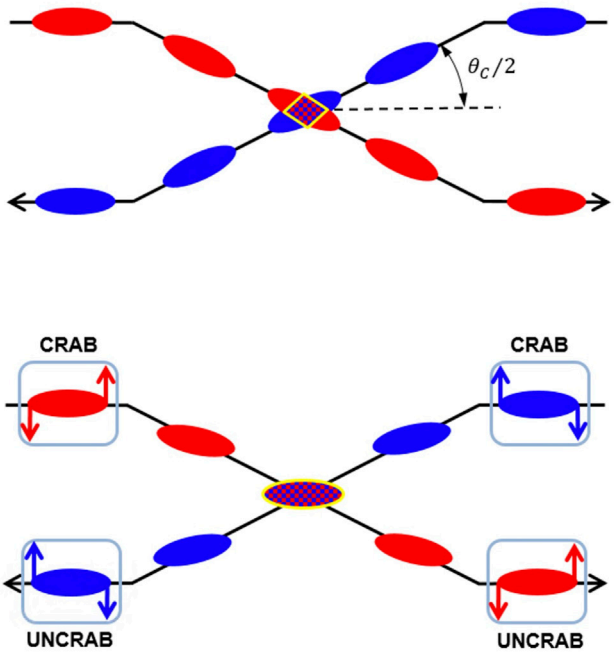


FIGURE 10 | Sketch illustrating crab crossing concept. Top: Beam collision scheme with crossing angle. Bottom: Crab crossing restores full bunch overlapping (Adapted from [58]).

operate in a beam energy range from 2.5 to 18 GeV, with beam current up to 2.5 A. Seventeen single-cell 591 MHz SRF cavities will have to deal with compensating 9 MW of beam power loss to synchrotron radiation and 1 MW of HOM losses. Each cavity has dual 400-kW fundamental power couplers. Strong HOM damping is provided by broad-band SiC HOM absorbers located outside cryomodules as shown in **Figure 9**.

4 COMMON RADIO FREQUENCY TECHNOLOGIES

There are RF technology areas that are common to linear and circular colliders. Here, two such areas: crab cavities and high-efficiency, low-cost klystrons are considered.

4.1 Crab Cavities

In future machines, the colliding beam trajectories will intersect at interaction points (IPs) at a small crossing angle. This is carried out to mitigate parasitic collisions, simplify removing of spent beams, etc. Such an intersection reduces the geometric overlap of the beams and the peak luminosity, as a result. In some cases, for example, with a relatively large crossing angle or with mismatch of the bunch lengths, the effect can be large. Using a crab crossing scheme, proposed by Palmer in 1988 [4], allows reestablishing full bunch overlapping during collision.

A deflecting (crab) cavity operates in such a way that the bunch center gets a zero kick, while its head and tail receive opposite transverse kicks with equal magnitude. The bunch is "chirped" and it moves along the trajectory in a crab-like manner. After the IP, another crab cavity can be installed to un-crab the bunch. **Figure 10** illustrates the crab crossing concept.

The crab crossing scheme was first implemented at the KEK B factory. The KEKB crab cavities were single-cell elliptical SRF structures operating at 4.5 K. The TM_{110}-like mode with a resonant frequency of 509 MHz was used for crabbing the beams. The cavity had a special coaxial coupler to damp the fundamental mode and large beam pipes for suppressing HOMs. The cell had a squashed shape to split two deflecting modes and select one polarization [59]. The cavities successfully crabbed the KEKB bunches to provide head-on collisions in the detector. To date, this has been the only operational collider with the crab crossing collision scheme.

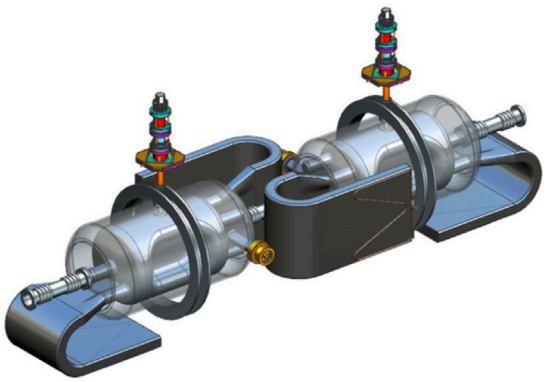


FIGURE 11 | 3D model of the RFD cavity concept showing two cavities with HOM waveguides [52].

HL-LHC will increase the LHC luminosity *via* several upgrades of the machine. As part of these upgrades, a crab crossing scheme will be implemented. The HL-LHC SRF crab cavities will operate at 400 MHz. As the real estate available for crab cavities in the collider is very tight, design studies were initiated to develop very compact deflecting cavities that would fit into the available space. Two novel compact designs were selected: RF Dipole (RFD) and Double Quarter Wave (DQW) resonators. For more details, see [58, 60] and the references therein.

ILC is currently following a similar path, going through a design study of various crab cavity proposals, including multi-cell elliptical cavity, RFD, DQW, and other design options with operating frequencies from 1.3 to 3.9 GHz.

Due to different bunch lengths of hadron and electron beams in the EIC, the crab crossing scheme will utilize cavities at two frequencies, 197 and 394 MHz. The hadron storage ring will use eight and four cavities at these two frequencies, respectively, while the electron ring will need only two cavities at 394 MHz [52]. The design studies are underway with the two main options being RFD and DQW. EIC imposes more stringent requirements on the HOM impedance of crab cavities than HL-LHC. As a result, the HL-LHC crab cavity HOM dampers cannot be adopted to EIC directly. The design updates are in progress. **Figure 11** illustrates the design of the baseline 197 MHz RFD crab cavity with two waveguide HOM couplers.

4.2 High-Efficiency, Low-Cost Klystrons

As it was mentioned already, achieving high efficiency is very important, especially for the RF systems of future colliders. Here, several recent developments in this area are mentioned. The High-Efficiency International Klystron Activity (HEIKA) collaboration [61, 62] investigates novel techniques for high-efficiency klystrons. Techniques such as the Bunching, Alignment, and Collecting (BAC) method [63] and the Core Oscillation Method (COM) [64] have been developed that promise increased efficiencies up to 90% [65]. One advantage of these methods is that it is possible to increase the efficiency of existing klystrons by equipping them with a new electron optics, as was demonstrated retrofitting an existing tube: the output power was increased by almost 50% and the klystron efficiency from 42% to 66% [66]. Incorporating periodic permanent magnet focusing can further reduce fabrication and operation costs. Another approach is to develop a simple modular system that would combine power of several low-voltage, highly efficient klystrons [44, 67].

5 RADIO FREQUENCY TECHNOLOGY CHALLENGES OF MUON COLLIDER

The muon collider can potentially extend the energy reach of lepton colliders to much higher energies, up to a 10 TeV center of mass [68]. Thus, such a machine can serve for both precision and discovery studies. It is expected that the collider's overall power consumption will be lower than that of the CLIC at 3 TeV. The machine would be relatively compact and could accommodate two experiments. While past efforts have demonstrated several key technologies, more R&D is needed before a fully integrated design could be developed. RF is one of the key technologies for the muon collider that must be part of a future R&D program.

Both normal conducting and superconducting RF will be utilized in different accelerators of the muon collider complex. A muon ionization cooling employs normal conducting 325 and 650 MHz RF cavities capable of providing high accelerating gradients in the presence of multi-tesla DC solenoid magnetic fields. Following that, a beam acceleration system includes SRF cavities operating at frequencies of 325, 650, and 1,300 MHz. These cavities must be robust to beam loading and capable of delivering significant gradients for rapid muon acceleration. No such cavities have been developed yet.

To yield good cooling performance in the cooling system, a compact lattice with large real-estate RF gradient is required. This results in RF cavities operating near to the breakdown limit while immersed in a strong solenoid field and poses specific challenges. The solenoid field guides electrons that are emitted at one location of the cavity surface to another location on the opposing wall and leads to localized heating that can result in breakdown and cavity damage. Operation of copper cavities in a 3 T magnetic field showed a maximum useable gradient of only 10 MV/m. It was proposed to use lower-Z materials such as beryllium to limit the energy loss density. Experiments with an 805 MHz beryllium cavity under vacuum yielded a gradient of 50 MV/m in an external magnetic field of 3 T [69]. Further efforts are needed to study alternative materials, for example, aluminum and fully develop the cavities at required frequencies.

6 SUMMARY

Radio frequency technologies will continue to be critical to the success of future colliders regardless of the collider configuration or particles being collided. In this article, we have tried to give the readers a broad, high-level overview of the field, highlighting recent advances, new ideas to pursue, and challenges that future colliders need to address. While most of the article was focused on accelerating structures, a couple of other technologies were mentioned: crab cavities and high-efficiency, high-power

klystrons. Performance of the RF systems strongly influences the luminosity and energy reach of the colliders either directly or indirectly *via* the capital and operational costs. Advances in RF technologies will continue to enable new proposals, such as the recently put forward linear colliders C^3 and HELEN.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

Work supported by the Fermi National Accelerator Laboratory, managed and operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the US Department of Energy. The US Government retains and the publisher, by accepting the article for publication, acknowledges that the US Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US Government purposes.

REFERENCES

1. Shiltsev V, Zimmermann F. Modern and Future Colliders. *Rev Mod Phys* (2021) 93(61):015006. doi:10.1103/revmodphys.93.015006
2. Belomestnykh S. Superconducting Radio-Frequency Systems for High-β Particle Accelerators. *Rev Accel Sci Tech* (2012) 05:147–84. doi:10.1142/s179362681230006x
3. Nassiri A, Chase B, Craievich P, Fabris A, Frischholz H, Jacob J, et al. History and Technology Developments of Radio Frequency (RF) Systems for Particle Accelerators. *IEEE Trans Nucl Sci* (2016) 63:707–50. doi:10.1109/tns.2015.2485164
4. Palmer R. B. *Energy Scaling, Cab Crossing and the Pair Problem*. C8806271: 613-619. SLAC-PUB-4707 (1988).
5. TESLA Technical Design Report. Part II: The Accelerator. R Brinkmann, K Flöttmann, J Roßbach, P Schmüser, N Walker, H Weise, editors (2001). Available from: https://flash.desy.de/sites2009/site_vuvfel/content/e1549/e1506/e1509/infoboxContent1511/partII.pdf (Accessed May 26, 2022).
6. Tigner M. A Possible Apparatus for Electron Clashing-Beam Experiments. *Nuovo Cim* (1965) 37:1228–31. doi:10.1007/bf02773204
7. Balakin VE, Skinsky AN. A Superhigh Energy Colliding Electron-Positron Beam Facility (VLEPP). In: Proc. 2nd ICFA Workshop on Possibilities and Limitations of Accelerators and Detectors (1979). p. 31.
8. Phinney N. SLC Final Performance and Lessons. arXiv:physics/0010008 [physics.acc-ph] (2000).
9. The ILC International Development Team and the ILC Community: Report to Snowmass 2021. arXiv:2203.07622 [physics.acc-ph] (2022).
10. Belomestnykh S, Bhat PC, Grassellion A, Checchin M, Denisov D, Geng RL, et al. Higgs-Energy LEptoN (HELEN) Collider Based on Advanced Superconducting Radio Frequency Technology. arXiv:2203.08211 [physics.acc-ph] (2022).
11. Telnov VI. A High-Luminosity Superconducting Twin e^+e^- Linear Collider with Energy Recovery. *J Inst* (2021) 16:P12025. doi:10.1088/1748-0221/16/12/p12025
12. Litvinenko VN, Bachhawat N, Chamizo-Llatas M, Jing Y, Méot F, Petrushina I, et al. The ReLiC: Recycling Linear e^+e^- collider. arXiv:2203.06476 [hep-ex] (2022).
13. Brunner O, Burrows PN, Calatroni S, Catalan Lasheras N, Corsini R, D'Auria G, et al. The CLIC Project. arXiv:2203.09186 [physics.acc-ph] (2022).
14. Bai M, Barklow T, Bartoldus R, Breidenbach M, Grenier P, Huang Z, et al. C^3: A "Cool" Route to the Higgs Boson and Beyond. arXiv:2110.15800 [hep-ex] (2021).
15. The First TESLA Collider Workshop. *Cornell Report CLNS 90-1029*. Ithaca, NY: Cornell University (1990).
16. Aune B, Bandelmann R, Bloess D, Bonin B, Bosotti A, Champion M, et al. Superconducting TESLA Cavities. *Phys Rev Accel Beams* (2000) 3:092001. doi:10.1103/PhysRevSTAB.3.092001
17. Adolphsen C, Barone M, Barish B, Buesser K, Burrows P, Carwardine J, et al. The International Linear Collider Technical Design Report – Volume 3.I: Accelerator R&D in the Technical Design Phase. arXiv:1306.6353 [physics.acc-ph] (2013).
18. Adolphsen C, Barone M, Barish B, Buesser K, Burrows P, Carwardine J, et al. The International Linear Collider Technical Design Report – Volume 3.II: Accelerator Baseline Design. arXiv:1306.6328 [physics.acc-ph] (2013).
19. The European X-Ray Free-Electron Laser: Technical Design Report. (2007). Available from: https://bib-pubdb1.desy.de/record/77248/files/european-xfel-tdr.pdf. (Accessed May 26, 2022).
20. Raubenheimer TO for the LCLS-II Collaboration. Technical Challenges of the LCLS-II CW X-Ray FEL. In: Proc. 6th Int. Particle Accel. Conf. IPAC2015. Richmond, VA, USA (2015). p. 2434–8.
21. Broemmelsiek D, Chase B, Edstrom D, Harms E, Leibfritz J, Nagaitsev S, et al. Record High-Gradient SRF Beam Acceleration at Fermilab. *New J Phys* (2018) 20:113018. doi:10.1088/1367-2630/aaec57
22. Yamamoto Y, Kako E, Umemori K, Sakai H, Saeki T, Dohmae T, et al. Stable Beam Operation at 33 MV/m in STF-2 Cryomodules at KEK. In: Proc. SRF2021. East Lansing, MI, USA (2021).
23. Belomestnykh S, Posen S, Bafia D, Balachandran S, Bertucci M, Burrill A, et al. Key Directions for Research and Development of Superconducting Radio Frequency Cavities. arXiv:2204.01178 [physics.acc-ph] (2022).
24. Padamsee H, Grassellino A, Belomestnykh S, Posen S. Impact of High Q on ILC250 Upgrade for Record Luminosities and Path Toward ILC380. arXiv:1910.01276 [physics.acc-ph] (2019).
25. Padamsee H. ILC Upgrades to 3 TeV. arXiv:2108.11904 [physics.acc-ph] (2021).
26. Gonnella D for the LCLS-II-HE Collaboration. LCLS-II-HE High Q_0 & Gradient R&D Program, First CM Test Results, and CM Plasma Processing Results. In: Presentation at the Virtual TESLA Technology Collaboration Meeting (2022).
27. Martinello M, Bice DJ, Boffo C, Chandrasekeran SK, Eremeev GV, Furuta F, et al. Q-Factor Optimization for High-Beta 650 MHz Cavities for PIP-II. *J Appl Phys* (2021) 130:174501. doi:10.1063/5.0068531
28. Posen S, Romanenko A, Grassellino A, Melnychuk OS, Sergatskov DA, et al. Ultralow Surface Resistance via Vacuum Heat Treatment of Superconducting Radio-Frequency Cavities. *Phys. Rev. Appl.* (2020) 13:014024. doi:10.1103/PhysRevApplied.13.014024
29. Grassellino A, Romanenko A, Bice D, Melnychuk O, Crawford AC, Chandrasekaran S, et al. Accelerating Fields Up to 49 MV/m in TESLA-Shape Superconducting RF Niobium Cavities via 75C Vacuum Bake. arXiv:1806.09824 [physics.acc-ph] (2018).
30. Bafia D, Grassellino A, Sung Z, Romanenko A, Melnychuk OS, Zasadzinski JF. Gradients of 50 MV/m in TESLA Shaped Cavities via Modified Low Temperature Bake. In: Proc. 19th Int. Conf. on RF Superconductivity SRF2019. Dresden, Germany (2019). p. 586–91.
31. Padamsee H. New Cavity Geometries. In: *RF Superconductivity: Science, Technology, and Applications*. WILEY-VCH Verlag GmbH & Co. KGaA (2009). p. 10–8.
32. Avrakhov P, Kanareykin A, Solyak N. *Traveling Wave Accelerating Structure for a Superconducting Accelerator*. Knoxville, TN, USA: Proc. PAC2005 (2005). p. 4296–8.
33. Shemelin V, Padamsee H, Yakovlev V. Optimization of a Traveling Wavesuperconducting Rf Cavity for Upgrading the International Linear Collider. *Phys Rev Accel Beams* (2022) 25:021001. doi:10.1103/physrevaccelbeams.25.021001
34. Yakovlev V, Avrakhov P, Kanareykin A, Kazakov S, Solyak N. *Progress Towards Development of a Superconducting Traveling Wave Accelerating Structure*. Albuquerque, NM, USA: Proc. PAC2007 (2007), p. 2182-4.

35. Kostin R, Avrakhov P, Kanareykin A, Yakovlev V, Solyak N. Progress Towards 3-Cell Superconducting Traveling Wave Cavity Cryogenic Test. *J Phys Conf Ser* (2017) 941:012100. doi:10.1088/1742-6596/941/1/012100
36. Valente-Feliciano A-M, Antoine C, Anlage S, Ciovati G, Delayen J, Gerigk F, et al. Next-Generation Superconducting RF Technology Based on Advanced Thin Film Technologies and Innovative Materials for Accelerator Enhanced Performance & Energy Reach. arXiv:2204.02536 [physics.acc-ph] (2022).
37. Valente-Feliciano A-M. Superconducting RF Materials Other Than Bulk Niobium: A Review. *Supercond Sci Technol* (2016) 29:113002. doi:10.1088/0953-2048/29/11/113002
38. Posen S, Liepe M, Eremeev G, Pudasaini U, Reece CE. Nb_3Sn Superconducting Radiofrequency Cavities: A Maturing Technology for Particle Accelerators and Detectors. arXiv:2203.06572 [physics.acc-ph] (2022).
39. Posen S, Lee J, Seidman DN, Romanenko A, Tennis B, Melnychuk OS, et al. Advances in Nb3Sn Superconducting Radiofrequency Cavities Towards First Practical Accelerator Applications. *Supercond Sci Technol* (2021) 34:025007. doi:10.1088/1361-6668/abc7f7
40. Dhuley RC, Gonin I, Kazakov S, Khabiboulline T, Sukhanov A, Yakovlev V, et al. Design of a 10 MeV, 1000 kW Average Power Electron-Beam Accelerator for Wastewater Treatment Applications. *Phys Rev Accel Beams* (2022) 25:041601. doi:10.1103/physrevaccelbeams.25.041601
41. Rimmer R. Single Cell Cavities. In: *Handbook of Accelerator Physics and Engineering*. 2nd ed. World Scientific (2013). p. 672–4.
42. Sundelin RM, Kirchgessner JL, Tigner M. Parallel Coupled Cavity Structure. *IEEE Trans Nucl Sci* (1977) 24(3):1686–8. doi:10.1109/tns.1977.4329052
43. Brezhnev ON, Logatchev PV, Pavlov VM, Pirogov OV, Shiyankov SV, Chernousov JD, et al. *Parallel-Coupled Accelerating Structures*. Gyeongju, Korea: Proc. LINAC2002 (2002). p. 213–5.
44. Nanni EA, Breidenbach M, Vernieri C, Belomestnykh S, Bhat PC, Nagaitsev S, et al. C^3 Demonstration Research and Development Plan. arXiv:2203.09076 [physics.acc-ph] (2022).
45. Cahill A, Fukasawa A, Rosenzweig J, Higashi Y, Bowden GB, Dolgashev VA, et al. *Measurements of Copper RF Surface Resistance at Cryogenic Temperatures for Applications to X-Band and S-Band Accelerators*. Busan, Korea: Proc. IPAC2016 (2016). p. 487–90.
46. Benedikt M, Chance A, Dalena B, Denisov D, Giovannozzi M, Gutleber J, et al. Future Circular Hadron Collider FCC-hh: Overview and Status. arXiv:2203.07804 [physics.acc-ph] (2022).
47. Tang J, Zhang Y, Xu Q, Gao J, Lou X, Wang Y. Study Overview for Super Proton-Proton Collider. arXiv:2203.07987 [hep-ex] (2022).
48. Agapov I, Benedikt M, Blondel A, Boscolo M, Brunner O, Chamizo Llatas M, et al. Future Circular Lepton Collider FCC-Ee: Overview and Status. arXiv:2203.08310 [physics.acc-ph] (2022).
49. CEPC Accelerator Study Group. Snowmass2021 White Paper AF3- CEPC. arXiv:2203.09451 [physics.acc-ph] (2022).
50. Abada A, Abbrescia M, AbdusSalam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. FCC-ee: The Lepton Collider. *Eur Phys J Spec Top* (2019) 228:261–623. doi:10.1140/epjst/e2019-900045-4
51. Litvinenko VN, Bachhawat N, Chamizo-Llatas M, Meot F, Roser T. CERC – Circular e+e- Collider Using Energy-Recovery Linac. arXiv:2203.07358 [physics.acc-ph] (2022).
52. Electron-Ion Collider at Brookhaven National Laborator. Electron Ion Collider Conceptual Design Report (2021). Available from: https://www.bnl.gov/ec/files/eic_cdr_final.pdf (Accessed May 26, 2022).
53. Belomestnykh S, Kirchgessner J. Superconducting Single Cell Cavities. In: *Handbook of Accelerator Physics and Engineering*. World Scientific (2013). p. 683–7.
54. Boussard D, Linnecar T. The LHC Superconducting RF System. *Adv Cryogen Eng A* (2000) 45:835.
55. Zheng H, Sha P, Zhai J, Pan W, Li Z, Mi Z, et al. Development and Vertical Tests of 650 MHz 2-Cell Superconducting Cavities with Higher Order Mode Couplers. *Nucl Instrum Methods Phys Res A* (2021) 995–165093. doi:10.1016/j.nima.2021.165093
56. Syratchev I, Peauger F, Karpov I, Brunner O. A Superconducting Slotted Waveguide Elliptical Cavity for FCC-Ee. CERN Report FCC-2105010000-OBR (2021).
57. Jensen E. *CTF3 Drive Beam Accelerating Structures*. Gyeongju, Korea: Proc. LINAC2002 (2002). p. 34–6.
58. Verdú-Andrés S, Belomestnykh S, Ben-Zvi I, Calaga R, Wu Q, Xiao B. Crab Cavities for Colliders: Past, Present and Future. *Nucl Part Phys Proc* (2016) 273-275:193–7. doi:10.1016/j.nuclphysbps.2015.09.025
59. Abe T, Akai K, Akemoto M, Akiyama A, Arinaga M, Ebihara K, et al. *Compensation of the Crossing Angle with Crab Cavities at KEKB*. Albuquerque, NM, USA: Proc. PAC2007 (2007). p. 27–31.
60. Verdú-Andrés S, Ben-Zvi I, Wu Q, Calaga R. *Crab Cavity Systems for Future Colliders*. Copenhagen, Denmark: Proc. IPAC2017 (2017). p. 2474–7.
61. Syratchev I. *Introduction to the High Efficiency International Klystron Activity HEIKA*. Geneva, Switzerland: Presentation at CLIC Workshop 2015.
62. Gerigk F. Status and Future Strategy for Advanced High Power Microwave Sources for Accelerators. In: Proc. 9th International Particle Accelerator Conference (IPAC 2018). Vancouver, BC: Canada (2018). p. 12–7.
63. Guzilov IA. BAC Method of Increasing the Efficiency in Klystrons. In: Proc. 10th International Vacuum Electron Sources Conference (IVESC) (2014). p. 1–2.
64. Constable D, Baikov A, Burt G, Guzilov I, Hill V, Jensen A, et al. High Efficiency Klystron Development for Particle Accelerators. In: Proc. 58th ICFA Advanced Beam Dynamics Workshop on High Luminosity Circular e+e- Colliders (eeFACT2016). Daresbury, UK (2016). p. 185–7.
65. Baikov AY, Marrelli C, Syratchev I. Toward High-Power Klystrons with RF Power Conversion Efficiency on the Order of 90%. *IEEE Trans Electron Dev* (2015) 62:3406.
66. Jensen E. Recent Developments Towards Very High Efficiency Klystrons. In: Presentation at 9th CW and High average RF power workshop. Grenoble, France (2016).
67. Weatherford B, Nanni E, Tantawi S. Advanced RF Sources R&D for Economical Future Colliders. In: arXiv:2203.15984 [physics.acc-ph] (2022).
68. Stratakis D, Mokhov N, Palmer M, Pastrone N, Raubenheimer T, Rogers C, et al. A Muon Collider Facility for Physics Discovery. arXiv:2203.08033 [physics.acc-ph] (2022).
69. Bowring D, Bross A, Lane P, Leonova M, Moretti A, Neuffer D, et al. Operation of Normal-Conducting RF Cavities in Multi-Tesla Magnetic Fields for Muon Ionization Cooling: A Feasibility Demonstration. *Phys Rev Accel Beams* (2020) 23:072001. doi:10.1103/physrevaccelbeams.23.072001

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Belomestnykh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 28 September 2022
DOI 10.3389/fphy.2022.967881

Check for updates

OPEN ACCESS

EDITED BY
Carmen Palomares,
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Spain

REVIEWED BY
Uzma Tabassam,
COMSATS University, Islamabad Campus, Pakistan
Sally Seidel,
University of New Mexico, United States
Hirsohi Okada,
Asia Pacific Center for Theoretical Physics (APCTP), South Korea

*CORRESPONDENCE
J. Alimena,
juliette.alimena@cern.ch

SPECIALTY SECTION
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

RECEIVED 13 June 2022
ACCEPTED 08 August 2022
PUBLISHED 28 September 2022

CITATION
Verhaaren CB, Alimena J, Bauer M, Azzi P, Ruiz R, Neubert M, Mikulenko O, Ovchynnikov M, Drewes M, Klaric J, Blondel A, Rizzi C, Sfyrla A, Sharma T, Kulkarni S, Thamm A, Blondel A, Suarez RG and Rygaard L (2022), Searches for long-lived particles at the future FCC-ee.
Front. Phys. 10:967881.
doi: 10.3389/fphy.2022.967881

COPYRIGHT
© 2022 Verhaaren, Alimena, Bauer, Azzi, Ruiz, Neubert, Mikulenko, Ovchynnikov, Drewes, Klaric, Blondel, Rizzi, Sfyrla, Sharma, Kulkarni, Thamm, Blondel, Suarez and Rygaard. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Searches for long-lived particles at the future FCC-ee

C. B. Verhaaren[1], J. Alimena[2]*, M. Bauer[3], P. Azzi[4], R. Ruiz[5], M. Neubert[6,7], O. Mikulenko[8], M. Ovchynnikov[8], M. Drewes[9], J. Klaric[9], A. Blondel[10], C. Rizzi[10], A. Sfyrla[10], T. Sharma[10], S. Kulkarni[11], A. Thamm[12], A. Blondel[13], R. Gonzalez Suarez[14] and L. Rygaard[14]

[1]Department of Physics and Astronomy, Brigham Young University, Provo, UT, United States, [2]Experimental Physics Department, CERN, Geneva, Switzerland, [3]Department of Physics, Durham University, Durham, United Kingdom, [4]INFN, Section of Padova, Padova, Italy, [5]Institute of Nuclear Physics, Polish Academy of Sciences, Kracow, Poland, [6]Johannes Gutenberg University, Mainz, Germany, [7]Cornell University, Ithaca, NY, United States, [8]Leiden University, Leiden, Netherlands, [9]Université Catholique de Louvain, Louvain-la-Neuve, Belgium, [10]University of Geneva, Geneva, Switzerland, [11]University of Graz, Graz, Austria, [12]The University of Melbourne, Parkville, VIC, Australia, [13]LPNHE, Université Paris-Sorbonne, Paris, France, [14]Uppsala University, Uppsala, Sweden

The electron-positron stage of the Future Circular Collider, FCC-ee, is a frontier factory for Higgs, top, electroweak, and flavour physics. It is designed to operate in a 100 km circular tunnel built at CERN, and will serve as the first step towards ≥100 TeV proton-proton collisions. In addition to an essential and unique Higgs program, it offers powerful opportunities to discover direct or indirect evidence of physics beyond the Standard Model. Direct searches for long-lived particles at FCC-ee could be particularly fertile in the high-luminosity Z run, where 5×10^{12} Z bosons are anticipated to be produced for the configuration with two interaction points. The high statistics of Higgs bosons, W bosons and top quarks in very clean experimental conditions could offer additional opportunities at other collision energies. Three physics cases producing long-lived signatures at FCC-ee are highlighted and studied in this paper: heavy neutral leptons (HNLs), axion-like particles (ALPs), and exotic decays of the Higgs boson. These searches motivate out-of-the-box optimization of experimental conditions and analysis techniques, which could lead to improvements in other physics searches.

KEYWORDS

future circular collider, particle physics, axion like particles, heavy neutral lepton, Higgs

1 Introduction

The Standard Model (SM) of particle physics is a mature and consistent theory that, after the observation of the Higgs boson, still fails to explain important experimental observations such as dark matter (DM), neutrino masses, or the baryon asymmetry of the Universe (BAU), among others. Theoretical aspects of the SM, including the origins of the electroweak scale, the spectrum of fermions masses, or flavor patterns also await

explanation. These questions may be answered by the direct observation of new particles and phenomena, or by measuring deviations from SM predictions. This is a chief motivation for new colliders that can push both the energy and intensity frontiers.

Long-lived particles (LLPs) are new, beyond the SM (BSM) states that travel a substantial distance between creation and decay in collider systems, presenting distinct experimental signatures [1]. LLPs feature in many BSM models and could provide answers to central questions in particle physics and beyond. The lifetime of a particle depends mostly on its mass and couplings, and so feebly-interacting particles (FIPs), with couplings to the SM particles several orders of magnitude smaller than the SM couplings, are often glaring examples of LLPs.

The experimental signatures of LLPs are particularly interesting. In contrast to promptly decaying particles, LLPs can decay after flying some distance from the primary interaction point. This produces a displaced vertex, with decay products including charged and neutral SM particles (e.g., charged leptons, light neutrinos, and pions). This kind of displaced signature is most commonly associated with LLPs. Other models predict disappearing LLPs giving rise to "short" or "broken" tracks; some are "stopped" or delayed; or produce unusual jets, such as "dark showers." Such a variety of experimental signatures is very different from the usual SM processes studied at colliders, and any of these signatures would, if observed, constitute a striking "smoking gun" of new physics. In hadron collider environments, standard trigger and reconstruction techniques are often unable to recognize LLP signatures and their study requires dedicated techniques and experiments.

The Future Circular Collider (FCC) program is a design study for a post-LHC particle accelerator at CERN following the priorities set by the 2020 Update of the European Strategy for Particle Physics [2]. The first stage of the FCC design study (FCC-ee) is a high-luminosity, high-precision lepton collider with the goal of better understanding the electroweak (EW) sector, especially the Higgs boson. In addition to a robust program in its own right, FCC-ee will also act as a possible precursor to a high-energy hadron collider (FCC-hh), located in the same tunnel and complementary to it [3].

Though FCC-ee will be a high precision exploration tool, it also opens the possibility of directly discovering new physics [4]. In particular, a future FCC-ee program has an exciting potential for exploring LLPs. The large integrated luminosity of the FCC-ee run around the Z pole, producing 5×10^{12} Z bosons (Tera-Z run), will facilitate direct searches for LLPs that could be closely linked to neutrino masses, explain the BAU, be sound DM candidates, or all at the same time. In the following, three central physics cases are discussed: heavy neutral leptons (HNLs) [5] in the context of the Phenomenological Type I Seesaw model, axion-like particles (ALPs), and exotic Higgs boson decays. In Section 2, the theoretical landscapes of these three physics cases are outlined. Section 3 discusses the experimental outlook. In particular, the common simulation details (Section 3.1), the experimental aspects of HNLs (Section 3.2), ALPs (Section 3.3), and exotic Higgs boson decays (Section 3.4) at FCC-ee, and considerations for additional detectors for LLPs at FCC-ee (Section 3.5) are covered. Finally, the summary and conclusions are presented in Section 4.

2 Theoretical landscape

In this section, the theoretical frameworks considered are briefly summarized. These representative scenarios are: the Phenomenological Type I Seesaw model (Section 2.1), axion-like particles (Section 2.2), and scalar singlet extensions of the SM (Section 2.3).

2.1 Heavy neutral leptons

The oscillations between neutrino flavor eigenstates in long-baseline experiments [6, 7] is one of the most pressing theoretical puzzles in particle physics today. The neutrino masses that produce these oscillations are also interesting because they imply either the existence of new particles and interactions or substantial changes to the SM paradigm [8, 9].

It is simply not enough to write effective neutrino masses, given the SM's limited particle content and the desire to understand the mechanism (or mechanisms) that render neutrinos so much lighter than charged leptons and quarks. It is also desirable to understand the flavor/mixing pattern among neutrinos and the possible connections to lepton and quark flavors themselves.

Among the most popular solutions to these mysteries are the Seesaw models. These tie the smallness of neutrino masses (m_{ν_k} with $k = 1, 2, 3$) to the scale of new physics (Λ). Depending on the complexity, the scale (or scales) introduced by these models can range from well below the EW scale to the Planck scale. In the most minimal scenarios [9], general arguments only require Λ to be below 10^{14} GeV. In high-scale Seesaws, light neutrino masses scale inversely with this new physics scale, $m_{\nu_k} \propto 1/\Lambda$. In low-scale Seesaws, the behavior can be more complicated, and in some cases light neutrino masses scale proportionally with this new physics scale, $m_{\nu_k} \propto \Lambda$. The manner in which either is implemented can vary widely, cf. Section 5 in [10], and the most minimal, tree-level constructions are known popularly as the Types I [11–17], II [16, 18–22], and III [23] Seesaw models. Notably, these minimal scenarios are often stepping stones to fuller, more ultraviolet-complete models, including extended gauge theories and grand unified theories. Importantly, neutrino mass models predict a plethora of phenomenology that are testable at a variety of low-

energy and high-energy experiments [10, 24–27], including particle colliders.

A common feature of several popular solutions to the origin of neutrino masses is the hypothetical existence of heavy, sterile neutrinos N_i, or heavy neutral leptons (HNLs) as they are sometimes called. Depending on the precise scenario, they can be Dirac or Majorana fermions, and mediate processes that violate lepton flavor symmetries. In practice, fermions may be arranged in a way that they form a Majorana state with Dirac-like properties [28, 29]. This state is known as a pseudo-Dirac fermion and results from underlying symmetries that explain the smallness of neutrino masses [30–33]. This leads to a phenomenology that practically interpolates between these limiting cases in the sense that the ratio between the rates of the lepton number violating and conserving decays (R_{ll}) can smoothly interpolate between $R_{ll} = 0$ and $R_{ll} = 1$ [34]. Searches for heavy Dirac and Majorana neutrinos at e^+e^- facilities have a long history [35–38], and if they are discovered at the LHC, FCC-ee would be a natural program to study their properties [26, 39–51].

If HNLs mix with the SM neutrinos, they can participate in the SM weak interaction *via* the couplings.

$$\mathcal{L}^{\text{Int}}_{\text{Type I}} = \mathcal{L}^W + \mathcal{L}^Z + \mathcal{L}^H, \quad \text{where} \tag{1a}$$

$$\mathcal{L}^W = -\frac{g_W}{\sqrt{2}} \sum_{\ell=e}^{\tau} \sum_{i=1}^{n_s} \overline{N_i} V^*_{\ell i} W^+_\mu \gamma^\mu P_L \ell^- + \text{H.c.}, \tag{1b}$$

$$\mathcal{L}^Z = -\frac{g_W}{2\cos\theta_W} \sum_{\ell=e}^{\tau} \sum_{i=1}^{n_s} \overline{N_i} V^*_{\ell i} Z_\mu \gamma^\mu P_L \nu_\ell + \text{H.c.}, \tag{1c}$$

$$\mathcal{L}^H = -\frac{g_W}{2M_W} h \sum_{\ell=e}^{\tau} \sum_{i=1}^{n_s} \overline{N_i} V^*_{\ell i} m_{N_i} P_L \nu_\ell + \text{H.c.} \tag{1d}$$

Here, $N_1, \ldots N_{n_s}$ are the heavy mass eigenstates of the theory. This model (1) is a common HNL benchmark for the pure type I Seesaw we use in this study. In extended models, the HNLs may have extra interactions, such as new gauge interactions [52–60]. The number of right-handed neutrino chiral eigenstates n_s is not constrained by gauge anomaly considerations in the model (1) because the chiral states are gauge singlets. Here, the $V_{\ell N_i}$ are the complex-valued, active-sterile mixing matrix elements and describe the coupling between the heavy mass eigenstate i and lepton flavor state ℓ.

The Lagrangian (1) approximates interactions to first order in the parameter $|V_{\ell i}|$. In this phenomenological framework, the masses of N_i (m_{N_i}) and $V_{\ell i}$ are taken to be independent. Hypothesizing connections to other physics, e.g., the relic abundance of DM or the matter-antimatter asymmetry of the observable Universe, can greatly constrain masses and mixing, cf. Section 2.1.2. For simplicity, the analysis of Section 3.2 considers only the lightest heavy mass eigenstate N_1, denoted by N, with mass and mixing m_N and $V_{\ell N}$. It is important to stress that considering only one HNL is for bench-marking and discovery purposes; realistic scenarios usually contain multiple mass eigenstates.

In analogy to the SM effective field theory (SMEFT) framework [61–63], the above Lagrangian can be systematically extended by higher dimensional operators in a framework known as νSMEFT [64–66], which can parameterize ultraviolet completions.

In the minimal Type I Seesaw model, where $n_s = 2$, the requirement to reproduce the observed pattern of light neutrino masses and mixing imposes testable constraints on the relative size of the HNL couplings $|V_{\ell i}|^2$ to individual SM flavors [67–70]. These will improve in the future with the Deep Underground Neutrino Experiment (DUNE) [71], cf. Figure 1, leading to a prediction that can be tested with FCC-ee. The position in the triangle in Figure 1 is entirely determined by the low energy phases in the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. The number of events observed in displaced decays of HNLs produced during the Z pole run permits the determination of the relative mixing $|V_{\ell i}|^2/(\sum_\ell |V_{\ell i}|^2)$ at the percent level [44], allowing for the Majorana phase in the light neutrino mixing matrix to be indirectly constrained [68, 72]. For $n_s = 3$, the model is less constrained, and making a testable prediction would require an independent determination of the lightest neutrino mass in the SM, cf. Figure 11 in [70]. Beyond minimal models, measuring the $|V_{\ell i}|^2/(\sum_\ell |V_{\ell i}|^2)$ can give insight into flavor at charge-parity (CP) symmetries of the neutrino mixing matrix [73–75], providing a hint towards possible ultraviolet completions.

2.1.1 Phenomenology of Dirac and Majorana heavy neutral leptons

In the kinematically accessible regime, FCC-ee is an excellent machine to discover HNLs [39] and study their properties. An analysis was completed for prompt HNL signals [79] and reproduced in Ref. [3]. A comparison for various machines and setups was compiled for the European Strategy for Particle Physics Briefing Book [80] [1]. The sensitivity to active-sterile mixing, labeled here by Θ, is shown in the summary figure, Figure 2. Figure 3 shows an updated estimation of different sensitivities for current and proposed detectors including an FCC-ee displaced vertex analysis.

Figure 4 shows the four and one event contours. The four event contour corresponds to the 95% confidence level (CL) exclusion in the absence of signal events. The one event contour shows that for the analysis of the LLP signatures, there still is a 63% probability to observe one event, which, in the absence of background events, could be sufficient for discovery, all the way down to the see-saw limit around 20–40 GeV.

If N is a Majorana fermion, then it can mediate processes that conserve lepton number as well as those that violate lepton

1 Figure 8.19.

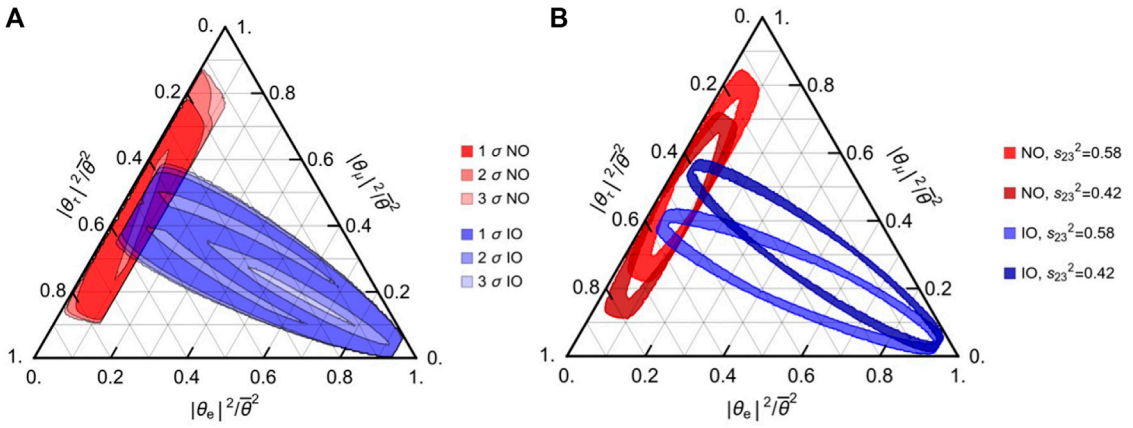


FIGURE 1
Allowed range for the relative magnitude of the HNL couplings to individual SM flavors in the model (27) with n_s =2, plot taken from [5]. **(A)**: The range of relative flavor mixings $(\sum_i|V_{ei}|^2)/(\sum_{i,\ell}|V_{\ell i}|^2)$ consistent with the current neutrino oscillation data, cf. e.g., [68, 69, 76]. The contours correspond to the allowed $\Delta\chi^2$ range taken from [77] for the case of normal (red) and inverted (blue) light neutrino mass ordering. **(B)**: The projected 90% CL contours for the relative mixings after 14 years of data taking at DUNE [78], assuming maximal CP violation $\delta = -\pi/2$ and two benchmark values of the PMNS angle θ_{23}, taken from the DUNE TDR [71], as indicated in the legend. FCC-ee can measure these ratios to the percent level in displaced HNL decays [44].

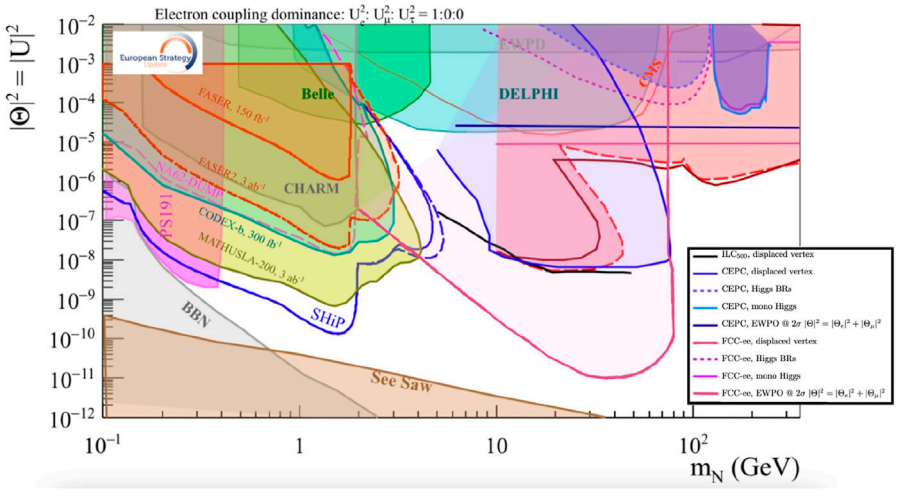


FIGURE 2
90% CL exclusion limits for a Heavy Neutral Lepton mixed with the electron neutrino, as presented in the European Strategy for Particle Physics Briefing Book [80]. The FCC-ee curves are in (overlined) dark purple—for FCC-ee, this is equivalent to a plot as function of the sum of matrix elements squared $|U_N|^2$. The curve below the Z boson mass corresponds to the combined LLP and prompt analysis performed with $10^{12}Z$ in Ref. [79]. The horizontal limit at high masses results from the effect of light-heavy neutrino mixing on the EW precision observables and remains valid up to O (1000 TeV).

number. Likewise, if lepton number is violated, then neutrinos must in principle possess Majorana properties [120, 121], though the amount of lepton number violation (LNV) in practice depends on the underlying model [121, 122]. Dirac neutrinos can only participate in processes that conserve lepton number. Therefore, differences between Dirac and Majorana N are closely related to differences between lepton number conservation (LNC) and LNV. The availability of LNV decay modes, for instance, leads to a Majorana N having a width (Γ_N) that is twice as large as a Dirac N. This implies that a Majorana N has a mean lifetime (τ_N), or mean displacement (d_N), that is half as long as that of a Dirac N.

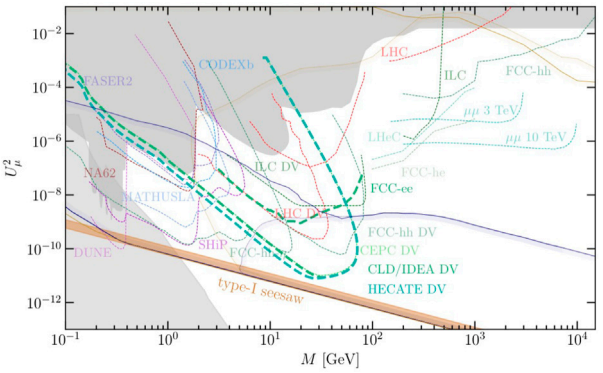


FIGURE 3
Bold green line: Sensitivity of displaced vertex searches at FCC-ee. The parameter region inside the curves corresponds to more than four observed HNL decays with $|V_{\ell N}|^2 = \delta_{\ell\mu} U_\mu^2$ from 5×10^{12} Z bosons, assuming no background events and 95% reconstructed HNL decays (i.e., all decays except the invisible decay) inside the main detectors based on the IDEA or CLD design with a displacement of over 400 μm. Based on Tables 7.2 and 7.3 in [3] with 1 m of instrumentation required for detection, we assume a cylinder of length l = 8.6 m and radius r = 5 m (CLD) or l = 11 m and r = 4.5 m (IDEA) as fiducial volumes. The resulting curves for the CLD and IDEA detectors are visually indistinguishable. For comparison, we show what CEPC can achieve with 4.2×10^{12} Z bosons [81] for an IDEA-type detector [82]. *Bold turquoise line*: Gain in sensitivity if the maximal observable displacement is increased with HECATE-like detectors [83] with l = 60 m, r = 15 m at two IPs. Medium gray: Constraints on the mixing of HNLs from past experiments [84–94]. Colorful lines: Estimated sensitivities of the main HL-LHC detectors [95–97] and NA62 [69], compared to the sensitivities of selected planned or proposed experiments (DUNE [98], FASER2 [99], SHiP [100, 101], MATHUSLA [102], CODEX-b [103], cf [10]. for a more complete list), prompt searches at FCC-ee or CEPC [50, 104], and searches at selected other proposed future colliders (FCC-hh [79, 97, 105, 106], ILC [43, 107] LHeC and FCC-he [108], and muon colliders [109], with DV indicating displaced vertex searches). The curves from [50, 95, 104] were re-scaled for a consistent integrated luminosity with [96, 97]. The sensitivity of FCC-ee and other future colliders can be further improved with dedicated long-lived particle detectors [83, 106, 110, 111]. *Brown band:* Indicative lower bound on the total HNL mixing $U_e^2 + U_\mu^2 + U_\tau^2$ from the requirement to explain the light neutrino oscillation data [77]. The band width corresponds to varying the light neutrino mass ordering and the lightest neutrino mass. The matter-antimatter asymmetry of the universe [112] can be explained by low scale leptogenesis [113–115] together with the light neutrino properties inside the mustard (violet) hashed contours with three [116] (two [117]) HNL flavours; solid and dashed contours indicate vanishing and thermal initial conditions in the early universe, respectively. *Light gray*: Lower bound on U_μ^2 from BBN [118, 119]. Plot adapted from [5].

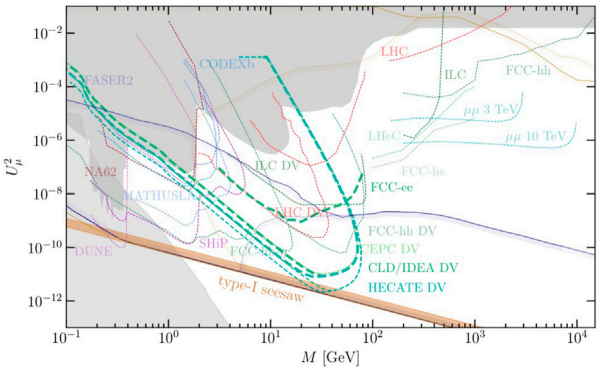


FIGURE 4
Comparison of the parameter regions in which four events (bold lines) and one event (non-bold lines) are expected in the IDEA/CLD detector or HECATE, with the same conventions and assumptions as in Figure 3.

Section 3.2.3 assumes a simple phenomenological model (1) with n_s = 1, i.e., only one mass eigenstate N that is either a Dirac or a Majorana fermion. Though this phenomenological model cannot completely reproduce the light neutrino oscillation data, it is sufficient to capture the collider phenomenology of the pure Dirac and Majorana HNLs benchmarks experimentally. Nature may be somewhere in the middle of these two cases. Therefore, observables that quantify differences between LNV and LNC processes, such as R_{ll} as defined in [34] or $\mathcal{A}$ as defined in [123, 124], can be used to interpolate between the benchmarks.

There is a rich phenomenology that connects R_{ll}, $\mathcal{A}$, and the decay rates of HNLs into different SM flavors to the mechanism that generates light neutrino masses. This connection can be accurately probed by studying the HNL properties at FCC [37, 48, 125] or other colliders [34, 123, 124, 126].

2.1.2 Probing cosmological questions

In addition to generating light neutrino masses, HNLs can, depending on their mass, affect the history of the Universe in different ways [114, 127, 128]. From a cosmological viewpoint, the most important motivation for HNL searches may be their potential connection to the origins of baryonic matter and DM.

Leptogenesis: Leptogenesis [129] provides an explanation for the matter-antimatter asymmetry in the observable Universe [112] and relates it to the properties of neutrinos. In the most popular scenario (based on the Type I Seesaw), HNLs generate the BAU *via* their CP-violating interactions with the thermal plasma. While it was originally believed that this mechanism only operates for HNL masses far above the EW scale [130], it is now established that sub-TeV HNLs can produce the observed BAU during their production [113, 114, 127] or freeze-out and decay [115, 131]. This implies that direct experimental searches have the potential to probe the origin of matter [132]. If any HNLs with masses at or below the EW scale are discovered in the near future, FCC-ee would provide a powerful tool to study their properties and test their connection to the BAU. A simple construction that supports low-scale leptogenesis is the Neutrino Minimal Standard Model (νMSM) [127]. Here, two HNLs simultaneously explain the neutrino masses and the BAU [114] for a wide range of experimentally accessible masses, cf.

Figure 3. Due to its minimality, the model is highly testable [67, 68]. In particular, leptogenesis constrains the flavor mixing pattern beyond the experimental fits shown in Figure 1, which can be tested by comparing flavored branching ratios in displaced decays. Finally, if accessible, HNL oscillations in the detectors are sensitive to the HNL mass splitting [44], which is a crucial parameter for leptogenesis.

Dark matter: HNLs with sufficiently small masses and mixing angles could be viable DM candidates [133]. Constraints on the HNL lifetime and from indirect searches restrict the range of masses and mixings to values that are inaccessible to direct searches at colliders, cf [134, 135]. for reviews. However, FCC-ee can indirectly probe sterile neutrino SM scenarios by searching for signatures of other particles that were involved in the DM production.

HNLs can be resonantly produced in the early Universe through their mixing-suppressed weak interactions if the lepton asymmetry at temperatures around the quantum chromodynamics (QCD) crossover greatly exceeded the BAU [136–139]. In the νMSM, this large lepton asymmetry can be generated by heavier HNLs that are also responsible for the BAU and neutrino masses [140]. The first parameter space studies [141–143] suggest that this is possible only for comparably small mixing angles, possibly making FCC-ee or a similar machine the only facility at which these HNLs could be discovered. If the HNLs have additional gauge interactions (cf. e.g., [144–149]), the extended gauge sector can be probed directly or indirectly at FCC-ee. If the DM is produced via the decay of a singlet [150–152] or charged [153, 154] scalar during freeze-out or freeze-in [155, 156], precision studies of the SM Higgs and of the portal can shed light on the mechanism. Most of these possibilities have not been studied in detail to date.

2.2 Axion-like particles

Many models that address open, fundamental problems of the SM are governed by global symmetries. If an approximate global symmetry is spontaneously broken, a pseudo Nambu-Goldstone boson appears in the theory that is light compared to the symmetry breaking scale. If this pseudo Nambu-Goldstone boson is a pseudoscalar, it is often referred to as an axion-like particle or ALP. The ALP's lightness singles it out as a uniquely promising experimental target that could open a first window onto high-scale new physics beyond the SM.

ALPs appear in many models that address open, fundamental problems in the SM. The most prominent example is the QCD axion, which was introduced in the 1980s to solve the strong CP problem [157–160] and found to simultaneously account for the observed DM relic abundance [161, 162]. QCD axions are typically very light, and these models are plagued by the "axion quality" problem, in which quantum gravity corrections destabilize the minimum of the axion potential, thereby reintroducing the strong CP problem [163–166]. Heavy-axion solutions to the strong CP problem circumvent this issue and so motivate ALPs with MeV-to-TeV scale masses [167–175]. ALPs in this mass range could also result from the breaking of global symmetries in low scale supersymmetric [176–178] or composite Higgs models [179–182]. Phenomenologically, they can also lead to successful EW baryogenesis [183].

An ALP dominantly couples to SM particles *via* dimension-5 operators,

$$\begin{aligned}\mathcal{L}_{\text{eff}} = &\frac{1}{2}\left(\partial_\mu a\right)\left(\partial^\mu a\right) - \frac{m_{a,0}^2}{2}a^2 + \sum_\psi \frac{c_{ff}}{2}\frac{\partial^\mu a}{f}\,\bar{\psi}\gamma_\mu\gamma_5\psi \\ &+c_{GG}\,\frac{\alpha_s}{4\pi}\,\frac{a}{f}\,G^a_{\mu\nu}\tilde{G}^{\mu\nu,a} + c_{\gamma\gamma}\,\frac{\alpha}{4\pi}\,\frac{a}{f}\,F_{\mu\nu}\tilde{F}^{\mu\nu} \\ &+c_{\gamma Z}\,\frac{\alpha}{2\pi s_w\, c_w}\,\frac{a}{f}\,F_{\mu\nu}\tilde{Z}^{\mu\nu} + c_{ZZ}\,\frac{\alpha}{4\pi s_w^2\, c_w^2}\,\frac{a}{f}\,Z_{\mu\nu}\tilde{Z}^{\mu\nu} \\ &+c_{WW}\,\frac{\alpha}{2\pi s_w^2}\,\frac{a}{f}\,W^+_{\mu\nu}\tilde{W}^{-\mu\nu},\end{aligned} \tag{2}$$

where $G^a_{\mu\nu}$ is the field-strength tensor of $SU(3)_c$, while $F_{\mu\nu}$, $Z_{\mu\nu}$ and $W^+_{\mu\nu}$ describe the photon, Z, and W boson in the broken phase of EW symmetry. The dual field-strength tensors are denoted by $\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta}$, etc., (with $\epsilon^{0123} = 1$); α_s and α are the QCD coupling and fine-structure constants, respectively; s_w and c_w denote the sine and cosine of the weak mixing angle; and the sum runs over all fermion mass eigenstates ψ. The suppression scale f is related to the new physics scale Λ via $\Lambda = 4\pi f$, and to the axion decay constant f_a by $f_a = -f/(2c_{GG})$. The ALP dominantly interacts with the Higgs boson *via* dimension-6 and -7 operators,

$$\mathcal{L}^H_{\text{eff}} = \frac{c_{ah}}{f^2}\left(\partial_\mu a\right)\left(\partial^\mu a\right)H^\dagger H + \frac{c_{Zh}}{f^3}\left(\partial^\mu a\right)\left(H^\dagger\, iD_\mu\, H + \text{h.c.}\right)H^\dagger H. \tag{3}$$

At FCC-ee, ALPs are predominantly produced in association with a photon, Z boson, or Higgs boson, as shown in the Feynman diagrams in Figure 5, or *via* exotic Z and Higgs decays. Resonant production of an ALP, e.g., $e^+e^- \to a$, is possible but suppressed by $m_e^2/(4\pi f)^2$. ALP production in vector boson fusion has been considered in Ref. [184] and detection prospects in light-by-light scattering in Refs. [185, 186].

The differential cross sections for associated γa/Za/ha production are given by [187, 188].

$$\begin{aligned}\frac{d\sigma\left(e^+e^- \to \gamma a\right)}{d\Omega} = &\frac{\alpha\alpha^2(s)}{128\pi^3}\,\frac{s^2}{f^2}\left(1 - \frac{m_a^2}{s}\right)^3\left(1 + \cos^2\theta\right) \\ &\times\left(|V_\gamma(s)|^2 + |A_\gamma(s)|^2\right),\end{aligned} \tag{4}$$

$$\begin{aligned}\frac{d\sigma\left(e^+e^- \to Za\right)}{d\Omega} = &\frac{\alpha\alpha^2(s)}{128\pi^3}\,\frac{s^2}{f^2}\,\lambda^{\frac{3}{2}}\left(x_a, x_Z\right)\left(1 + \cos^2\theta\right) \\ &\times\left(|V_Z(s)|^2 + |A_Z(s)|^2\right),\end{aligned} \tag{5}$$

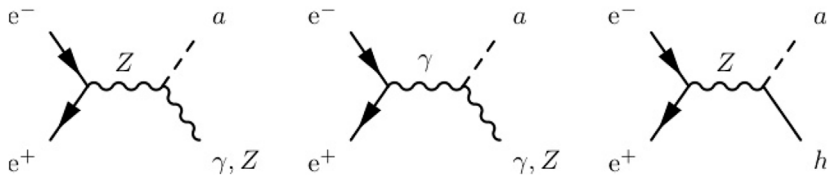


FIGURE 5
ALP production processes in electron-positron collisions.

$$\frac{d\sigma(e^+e^- \to ha)}{d\Omega} = \frac{2\pi^3\alpha}{c_w^2 s_w^2}\frac{|c_{Zh}|^2}{f^2}\frac{s m_Z^2}{(s-m_Z^2)^2}\lambda^{\frac{3}{2}}(x_a, x_h)\sin^2\theta\,(g_V^2+g_A^2), \tag{6}$$

where $\lambda(x,y) = (1-x-y)^2 - 4xy$, $x_i = (m_i^2/s)$, $\sqrt{s}$ is the center-of-mass energy, and θ describes the scattering angle of the photon, Z, or Higgs boson relative to the beam axis. The vector and axial-vector form factors are given by.

$$\begin{aligned} V_\gamma(s) &= \frac{c_{\gamma\gamma}}{s} + \frac{g_V}{2c_w^2 s_w^2}\frac{c_{\gamma Z}}{s-m_Z^2+im_Z\Gamma_Z}, \\ A_\gamma(s) &= \frac{g_A}{2c_w^2 s_w^2}\frac{c_{\gamma Z}}{s-m_Z^2+im_Z\Gamma_Z}, \end{aligned} \tag{7}$$

$$\begin{aligned} V_Z(s) &= \frac{1}{c_w s_w}\frac{c_{\gamma Z}}{s} + \frac{g_V}{2c_w^3 s_w^3}\frac{c_{ZZ}}{s-m_Z^2+im_Z\Gamma_Z}, \\ A_Z(s) &= \frac{g_A}{2c_w^3 s_w^3}\frac{c_{ZZ}}{s-m_Z^2+im_Z\Gamma_Z}, \end{aligned} \tag{8}$$

with $g_V = 2s_w^2 - 1/2$, $g_A = -1/2$, and Γ_Z is the total width of the Z boson. The process where an ALP is radiated off an initial-state electron exhibits an additional suppression of (m_e^2/s).

The integrated cross section of $e^+e^- \to \gamma a$ below the Z pole is dominated by the photon contribution, which is proportional to $c_{\gamma\gamma}$, while above the Z pole the process proportional to $c_{\gamma Z}$ also contributes. Combining these measurements at low and high energies therefore enables us to access these couplings separately. At the Z pole, the cross section becomes

$$\sigma(e^+e^- \to \gamma a) \approx \frac{\alpha}{24\pi^2}\alpha^2(m_Z^2)\left(1-\frac{m_a^2}{m_Z^2}\right)^3\left[\frac{|c_{\gamma\gamma}|^2}{f^2} + \frac{m_Z^2}{\Gamma_Z^2}\frac{|c_{\gamma Z}|^2}{16 s_w^4 c_w^4 f^2}\right]. \tag{9}$$

The contribution from the Z boson propagator is enhanced by $(m_Z^2/\Gamma_Z^2) \sim 1336$, which allows one to directly access the coupling $c_{\gamma Z}$ (as long as $c_{\gamma\gamma}$ is not much bigger than $c_{\gamma Z}$). ALPs can also be produced in exotic decays of Z and Higgs bosons [187–189]. The exotic decay rates are given by.

$$\Gamma(Z \to \gamma a) = \frac{\alpha\,\alpha(m_Z)\,m_Z^3}{96\pi^3 s_w^2 c_w^2 f^2}|c_{\gamma Z}|^2\left(1-\frac{m_a^2}{m_Z^2}\right)^3, \tag{10}$$

$$\Gamma(h \to Za) = \frac{m_h^3 v^2}{64\pi f^6}|c_{Zh}|^2\lambda^{3/2}\left(\frac{m_Z^2}{m_h^2}, \frac{m_a^2}{m_h^2}\right), \tag{11}$$

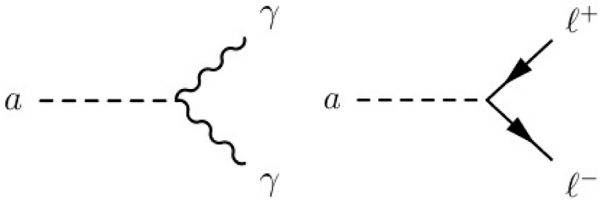


FIGURE 6
ALP decay processes at FCC-ee.

$$\Gamma(h \to aa) = \frac{m_h^3 v^2}{32\pi f^4}|c_{ah}|^2\left(1-\frac{2m_a^2}{m_h^2}\right)^2\sqrt{1-\frac{4m_a^2}{m_h^2}}. \tag{12}$$

Once produced, ALPs lead to a variety of signatures inside the detector. Very long-lived ALPs, for example, escape the detector and lead to a signature with missing momentum. ALPs with somewhat shorter lifetimes may decay into gauge bosons, leptons, and quarks inside the detector. The photon and lepton decay channels are shown in Figure 6. Their corresponding decay widths are given by.

$$\Gamma(a \to \gamma\gamma) = \frac{\alpha^2\, m_a^3}{64\pi^3 f^2}c_{\gamma\gamma}^2, \tag{13}$$

$$\Gamma(a \to \ell^+\ell^-) = \frac{m_a\, m_\ell^2}{8\pi f^2}\,c_{\ell\ell}^2\sqrt{1-\frac{4m_\ell^2}{m_a^2}}. \tag{14}$$

An ALP decay into hadrons can be computed perturbatively for relatively large ALP masses, i.e., $m_a \gg \Lambda_{\rm QCD}$. The decay width into bottom quarks specifically is given by

$$\Gamma(a \to b\bar{b}) = \frac{3\,m_a\, m_b^2(m_a)}{8\pi f^2}\,c_{bb}^2\sqrt{1-\frac{4m_b^2}{m_a^2}}, \tag{15}$$

and similarly for $\Gamma(a \to c\bar{c})$. The decay rate into light quarks (u, d, s) can be computed using quark-hadron duality and is given by [187, 190].

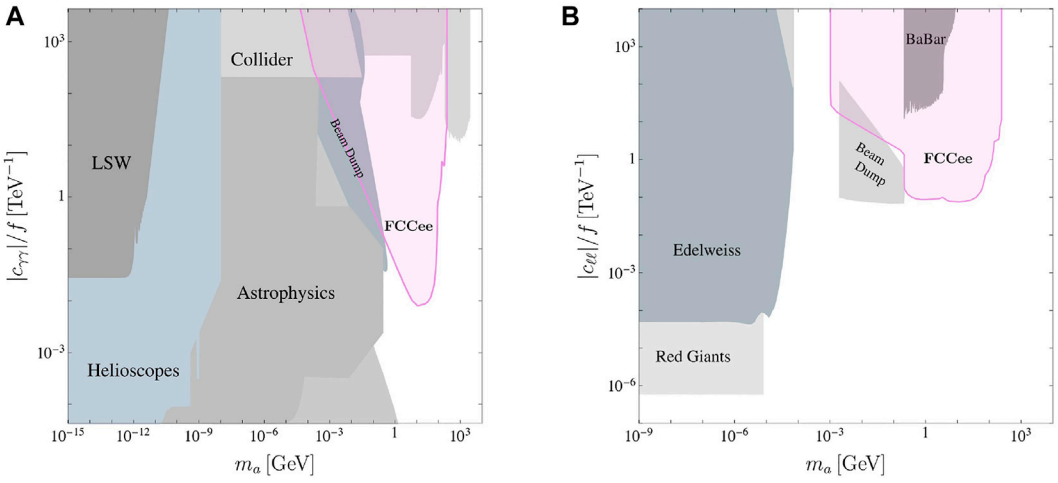


FIGURE 7
Projected sensitivity of FCC-ee in **(A)** $e^+e^- \to \gamma a \to 3\gamma$ and **(B)** $e^+e^- \to \gamma a \to \gamma \ell^+\ell^-$ in purple. Figure adapted from Figure [188].

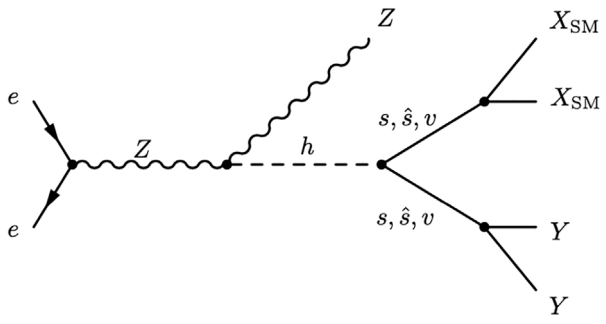


FIGURE 8
Example production of LLPs through exotic decays of the Higgs boson h. The Higgs decays to a pair of scalars s, pseudoscalars $\hat{s}$, or vectors v. At least one of these decays within the detector volume to SM particles. The other may or may not decay within the detector and may decay to visible or invisible states.

$$\Gamma(a \to \text{light hadrons}) = \frac{\alpha_s^2(m_a)\, m_a^3}{8\pi^3 f^2}\left[1 + \frac{83}{4}\frac{\alpha_s(m_a)}{\pi}\right]\left|C_{GG}^{\text{eff}}(m_a)\right|^2, \tag{16}$$

where the ALP couplings to both gluons and quarks contribute *via*

$$C_{GG}^{\text{eff}}(m_a) = c_{GG} + \frac{1}{2}\sum_{q\neq t} c_{qq}\, B_1\left(\frac{4m_q^2}{m_a^2}\right). \tag{17}$$

The function B_1 behaves as $B_1(4m_q^2/m_a^2) \approx 1$ for $m_q \ll m_a$ and $B_1(4m_q^2/m_a^2) \approx -m_a^2/(12m_q^2)$ for $m_q \gg m_a$. The explicit form of B_1 is given in e.g., [187]. For light ALPs, $m_a \ll \Lambda_{\text{QCD}}$, the decay into three pions may be kinematically accessible, with a decay rate which is given in [187, 191]. However, it is worth noting that the FCC-ee program as currently envisioned will not be able to produce significant numbers of ALPs that are heavy enough to decay in two top quarks, due to the high center of mass energy that would be required. Depending on their lifetime, ALPs can decay promptly at the interaction point or they may decay after having travelled a certain distance inside the detector.

At FCC-ee, all combinations of ALP production modes with visible and invisible decay modes can be investigated [188, 192]. While many processes, in particular exotic Higgs decays, depend on two independent couplings, under certain assumptions a few processes only depend on a single coupling parameter. For example, the $e^+e^- \to \gamma a \to 3\gamma$ and $e^+e^- \to Za \to Z\gamma\gamma$ processes only depend on the ALP-photon coupling $c_{\gamma\gamma}$ when it is assumed that both the ALP-photon and the ALP-photon-Z couplings originate from the ALP coupling to either $SU(2)_L$ or $U(1)$ gauge bosons before EW symmetry breaking. If the ALP only couples to $U(1)$ gauge bosons, then $c_{\gamma Z} = -s_w^2 c_{\gamma\gamma}$. In this case, Figure 7A shows the projected sensitivity of FCC-ee to $c_{\gamma\gamma}$ using the $e^+e^- \to \gamma a \to 3\gamma$ channel [188]. This analysis assumes at least four signal events and combines the Z-pole run with runs at $\sqrt{s} = 2m_W$ and $\sqrt{s} = 250$ GeV. Further details are provided in [188]. Another process that depends only on a single coupling is $e^+e^- \to \gamma a \to \gamma\ell^+\ell^-$ when the ALP-photon and the ALP-photon-Z couplings are induced via a loop of leptons. In this case,

$$c_{\gamma\gamma} = \sum_{\ell=e,\mu,\tau} c_{\ell\ell} B_1(4m_\ell^2/m_a^2) \quad \text{and} \quad c_{\gamma Z} = (s_w^2 - 1/4)c_{\gamma\gamma}. \tag{18}$$

Figure 7B shows the projected sensitivity of FCC-ee to $c_{\ell\ell}$ using the process $e^+e^- \to \gamma a \to \gamma\ell^+\ell^-$ [188].

The $e^+e^- \to \gamma a \to 3\gamma$ and $e^+e^- \to \gamma a \to \gamma \ell^+\ell^-$ searches are sensitive to ALP decay lengths of up to 1.5 m and 2 cm, respectively. The search for long-lived ALPs may be significantly improved with the installation of a dedicated far detector that could probe decay lengths of up to 100 m [193, 194]. For FCC-ee reach on the relaxion, see Ref. [195]. In addition to direct measurements, FCC-ee will be able to significantly constrain the ALP contribution to the oblique parameters [187, 188], whose determination is expected to improve by an order of magnitude [196].

2.3 Exotic Higgs boson decays

The Higgs boson has a unique role within the SM. It is the only apparently elementary scalar particle that has been discovered. In particular, whatever new physics is responsible for the cosmological DM, the apparent asymmetry between matter and antimatter, or small neutrino masses may well have some coupling to the Higgs boson. In short, the Higgs boson is a likely gateway to what lies beyond the SM.

The two-body decays of the Higgs boson to SM particles are controlled by small Yukawa couplings or loop suppression making its decay width much smaller than its mass. Consequently, current bounds on the Higgs width leave plenty of room for "exotic" decays, that is, decays not predicted by the SM. However, future colliders, like FCC-ee, will be able to measure the Higgs width much more precisely than at the LHC [197]. The products of these exotic Higgs boson decays can decay promptly themselves or be completely stable, each of which present their own experimental challenges and advantages. However, searches for particles whose lifetimes are more intermediate, i.e., that decay within the experimental detector but at a measurable distance from the interaction point, can have very low backgrounds in comparison to prompt searches. This gives Higgs boson decays to LLPs remarkable power to probe particles and sectors whose couplings to the Higgs are small but nonzero. The e^+e^- to Zh processes shown in Figure 8 illustrate the utility of the FCC-ee collider. Because the initial state and the Z decays are well understood, invisible, partially invisible, and displaced decays of the Higgs boson can be probed with confidence. For a review, see Ref. [102] and the recent work in Ref. [198].

The characteristics of LLPs vary considerably. Exotic Higgs decays to spin-zero particles are considered first. Such decays at future lepton colliders were considered in Ref. [199]. Long-lived scalars may result from simple constructions, such as adding a single scalar field to the SM:

$$V_{\text{scalar}} = V_H + V_S + c_1 S|H|^2 + c_2 S^2|H|^2 \ . \tag{19}$$

They may also arise in rich, hidden sectors such as Hidden Valley models [200–202]. Of particular interest are hidden sectors motivated by Neutral Naturalness [203–206]. These models address the little hierarchy problem through new symmetries, but the symmetry partners of the SM quarks do not carry SM color. Instead they are charged under a hidden, QCD-like confining force.

In many models with the long-lived scalar s or pseudoscalar $\hat{s}$, the Higgs boson decay products inherit much of the Higgs' coupling structure. While the actual size of the couplings are reduced by a common small mixing angle θ, the branching fractions are those of a SM Higgs boson with the mass of the LLP. In the scalar case, one often finds

$$\Gamma(s \to X_{\text{SM}} X_{\text{SM}}) \ = \ \sin^2\theta \ \Gamma(h(m_s) \to X_{\text{SM}} X_{\text{SM}}) \ . \tag{20}$$

The pseudoscalar case is slightly modified [207–209] and can also include the $h \to \hat{s}\, Z$ decay channel, see for instance the ALP results given in Eqs 11, 12. Since the masses of the LLPs must be less than half the Higgs boson mass, the dominant decays modes are into the heaviest kinetically accessible SM quarks. Thus, for Higgs boson decays into spin-zero LLPs, hadronic final states, and b-jets especially, are particularly motivated.

Rather than scalars, the LLPs may be spin-half fermions. These can be related to the BAU [210] or to Seesaw explanations of the neutrino masses [11–17]. The heavy neutrinos N in these models have been shown to have a wide range of possible decay lengths, including within the volume of an FCC-ee detector [211–213]. The N mainly decay into a SM lepton and an off-shell weak gauge boson. This leads to three-body final states which may be composed of both quarks (jets) and leptons.

The Higgs boson can also decay to long-lived vectors v. A simple framework is the Hidden Abelian Higgs model [214]. In this case, a new $U(1)'$ gauge symmetry is broken by a hidden Higgs field h_D that generates a vacuum expectation value (VEV). The hidden photon A'_μ of the new gauge symmetry gets a mass proportional to the hidden Higgs VEV and can also have kinetic mixing with the SM through

$$-\frac{\epsilon}{2\cos\theta_W} F'_{\mu\nu} F_Y^{\mu\nu} \ , \tag{21}$$

where θ_W is the weak mixing angle and $F_Y^{\mu\nu}$ is the field strength for SM hypercharge. The parameter ϵ can vary over a huge range, and controls the degree to which SM fermions couple to A'_μ. For sufficiently small ϵ, the massive hidden photon is a LLP.

The hidden photon's coupling to SM fields is proportional to their hypercharge. This means that, when and if they are kinematically accessible, quark final states make up most of its branching fraction, though decay rates to leptons are non-negligible. A small ϵ also means that the direct coupling of the hidden photon to the SM Higgs boson is small. However, the mixing between SM Higgs and the hidden Higgs can be larger than ϵ. This allows the Higgs boson to decay to two hidden photons at a larger rate.

In summary, the Higgs boson may have appreciable decay widths into LLPs of various spin. The decay modes of the LLPs can vary, but it has been shown that hadronic final states play a significant role in all the decay types outlined above. Decays to

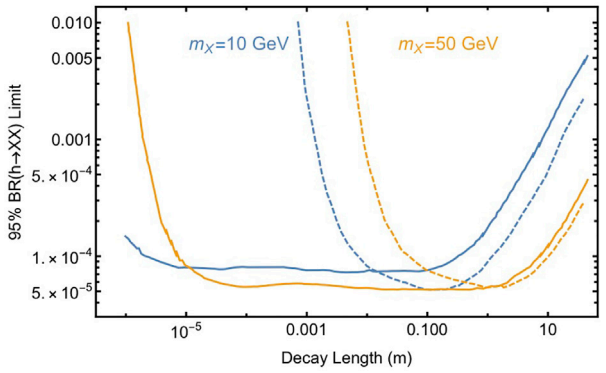


FIGURE 9
Plot of data recorded in [199] to illustrate the potential sensitivity of FCC-ee to exotic Higgs boson decays to LLPs, denoted *X*. Two LLP mass benchmarks are shown: 10 GeV (blue) and 50 GeV (tan). For each benchmark two search strategies are presented. The solid line employs an invariant mass cut to improve sensitivity at shorter decay lengths, the dashed line relies on longer decay lengths to reduce SM backgrounds.

long-lived fermions stand out as different, in that their leading decays are three-body. Pseudoscalars may also lead to $h \to \hat{s}\,Z$ decays, but in general, the $h \to XX$ process captures most of the interesting possibilities. Assuming the X particle has significant branching into SM quarks (and possibly into b quarks in particular) appears to be the most motivated benchmark. Of course, the variety of other decays can be leveraged in more model-specific analyses.

Figure 9 displays an illustration, taken from Ref. [199], of how sensitive FCC-ee can be to Higgs boson decays to long-lived X particles. The 95% limit on the exotic branching fraction to these particles is plotted as a function of the X's decay length. Two mass benchmarks, m_X = 10 (blue) and 50 (tan) GeV, are shown (additional benchmarks are considered in Ref. [199]) for two search strategies. The solid line corresponds to using an invariant mass cut to retain sensitivity to shorter decay lengths. In contrast, the dashed line depends on longer decay lengths to reduce SM backgrounds.

3 Experimental outlook

This section presents new studies produced for this paper that follow up on the theoretical landscape presented in Section 2.

3.1 Simulation details

For all signal and background processes, the event generator `MadGraph5_aMC@NLO` v3.2.0 [215, 216] is used to simulate at leading order unpolarized, parton-level e^+e^- collisions at $\sqrt{s}$ = 91 GeV. For all processes, parton-level events are passed to `Pythia` [217] v8.303 to simulate parton showering and hadronization. For each signal benchmark point, 50 × 10^3 unscaled events were generated, and for each background process, 10^7–10^9 unscaled events were generated, depending on the process.

The detector response is simulated with `Delphes` v3.4.2 [218], using the latest Innovative Detector for Electron–positron Accelerators (IDEA) FCC-ee detector concept [219] card. The IDEA detector comprises of a silicon pixel vertex detector; a large-volume, light short-drift wire chamber surrounded by a layer of silicon micro-strip detectors; a thin, low-mass superconducting solenoid coil; a pre-shower detector; a dual-readout calorimeter; and muon chambers within the magnet return yoke.

The k4SimDelphes project [220] converts Delphes objects to the EDM4HEP format [221], which is the common data format used for the simulation of future colliders. A sophisticated analysis framework has been developed for all FCC analyses using the EDM4hep format. It is based on RDataFrames [222], where `C++` code is compiled in a `ROOT` [223] dictionary as "analysers." These are subsequently called in `Python`. Several external packages, such as `ACTS` [224], `FastJet` [225], and `awkward` [226], are included.

3.1.1 Heavy neutral leptons

To study Dirac and Majorana HNLs at FCC-ee, the processes.

$$\textbf{Majorana}\ N\text{: } e^+e^- \to Z \to N\nu_e + N\overline{\nu_e}, \quad \text{with} \quad N \to e^+e^-\nu_e + e^+e^-\overline{\nu_e}, \tag{22a}$$

$$\textbf{Dirac}\ N\text{: } e^+e^- \to Z \to N\overline{\nu_e} + \bar{N}\nu_e, \ \text{with}\ N\ (\bar{N}) \to e^+e^-\nu_e\ (\overline{\nu_e}), \tag{22b}$$

are simulated using the `HeavyN` [227, 228] and `HeavyN_Dirac` [97, 228] universal `FeynRules` Object [229–231] libraries in conjunction with `MadGraph5_aMC@NLO`. These libraries implement the interaction Lagrangian described in Section 2.1 for Majorana and Dirac N, respectively. A representative subset of Feynman diagrams common to both the Dirac and Majorana case is shown in Figure 11. For the Majorana case, both LNC and LNV channels are included. The Dirac case only permits LNC channels. The preservation of spin correlation in the production and decay of N with this setup was checked in Ref. [124]. When unspecified, the results consider the Majorana case. As a further benchmark, the assumption that N couples only to the electron-flavor sector is made, i.e., $|V_{eN}|$ is kept nonzero and set $|V_{\mu N}|$, $|V_{\tau N}| = 0$. Only one heavy neutrino

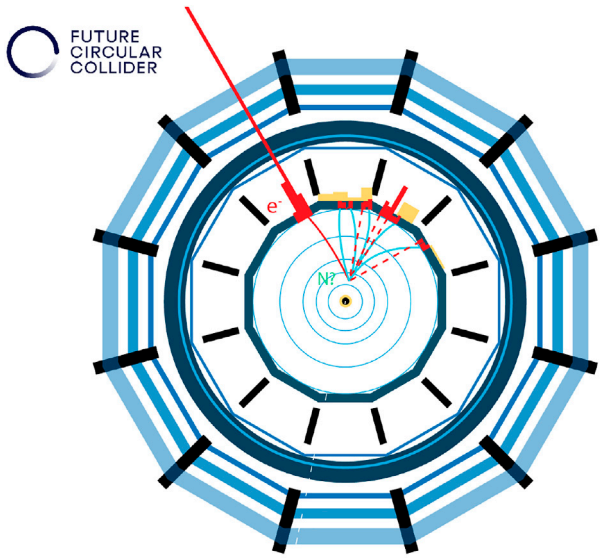


FIGURE 10
Representation of an event display at an FCC-ee detector of a HNL decay into an electron and a virtual W decaying hadronically. Courtesy of the FCC collaboration.

mass eigenstate is considered. SM inputs are fixed according to the values in Ref. [228].

3.1.2 Axion-like particles

To study the production ALPs a from Z decays at FCC-ee, the process

$$\textbf{ALP}: e^+e^- \to Z \to a\gamma, \quad \text{with} \quad a \to \gamma\gamma, \tag{23}$$

is simulated using the model libraries of Ref. [188] in conjunction with `MadGraph5_aMC@NLO`. These libraries implement the Lagrangian described in Section 2.2.

3.1.3 Exotic Higgs boson decays

A simulation study of exotic Higgs decays into LLPs is left for a future paper, as well as additional detector concepts, namely the CLIC-like detector (CLD) design [232].

3.2 Heavy neutral leptons

Although the most promising Seesaw models feature two or three HNL states in the same mass range, and possibly almost degenerate, a reasonable experimental approach is to begin by considering the production and decay of a single HNL particle.

The branching fraction of a Z boson decay into any light neutrino or antineutrino and a heavy neutrino N, which mixes with the three families of neutrinos is given by [97, 233]:

TABLE 1 The cross section and expected number of events at 150 ab^{-1}, for an HNL with a mass of 50 GeV and for several choices of $|V_{eN}|$.

Active-sterile mixing $\lvert V_{eN}\rvert$	Cross Section [pb]	Expected events at 150 ab^{-1}
1×10^{-1}	2.29	343,200,000
1×10^{-2}	2.29×10^{-2}	3,432,000
1×10^{-3}	2.29×10^{-4}	34,320
1×10^{-4}	2.29×10^{-6}	343
1×10^{-5}	2.29×10^{-8}	3
1×10^{-6}	2.29×10^{-10}	0

No event selection is applied.

$$\text{BR}(Z \to \nu N) = \frac{2}{3}|U_N|^2 \,\text{BR}(Z \to \text{invisible}) \left(1 + \frac{{m_N}^2}{2{m_Z}^2}\right) \times \left(1 - \frac{{m_N}^2}{{m_Z}^2}\right), \tag{24}$$

where $|U_N|^2 \equiv \sum_{\ell=e,\mu,\tau}|U_{\ell N}|^2$ is the sum of the mixing matrix elements of the HNL N with the three active neutrinos ν_ℓ. As the HNL masses considered here are much heavier than the tau lepton, the total charged current decay rate of the HNL $N \to \ell_\lambda W*$ is also proportional to the same combination of mixing angles.

$$\Gamma_N = \frac{1}{c\tau_N} \simeq C_0 C_{MD}|U_N|^2 \left(\frac{m_N}{50\text{GeV}}\right)^5 \times \left(\frac{3.10^9}{1\text{ cm}}\right) \tag{25}$$

Here, C_0 is a numerical coefficient of $\mathcal{O}(1)$ that takes into account the open charged- and neutral-current decays of the heavy neutrino, and C_{MD} is a coefficient that depends on the Dirac ($C_{MD} = 1$) or Majorana ($C_{MD} = 2$) nature of the particle, since twice as many decay channels are open for the Majorana particle decay. Potentially, with sufficient statistics, the direct comparison of the event rate with the lifetime for an HNL of a known mass would allow a discrimination between a Dirac and a Majorana particle.

The corresponding decay length is then of the order of a meter for a 50 GeV HNL. In those conditions, a HNL would decay in the volume of an FCC-ee detector, leading to the observable signature of a displaced vertex, with a significant time delay (several nanoseconds) with respect to ultra relativistic particles. This leads to a particularly clean signature, for which a first analysis [39] argued that it could be a background-free search, at least for the dominant charged current decay $N \to \ell W^* \to \ell+$ hadrons. Figure 10 shows what such a possible decay of the N at a future FCC-ee experiment would look like, in this case for a semileptonic final state.

Furthermore, for $Z \to N\nu_\ell$ decays, the two-body Z decay kinematics results in a mono-chromatic HNL.

Therefore, even in cases where a full, final-state reconstruction is not possible, a simultaneous measurement of

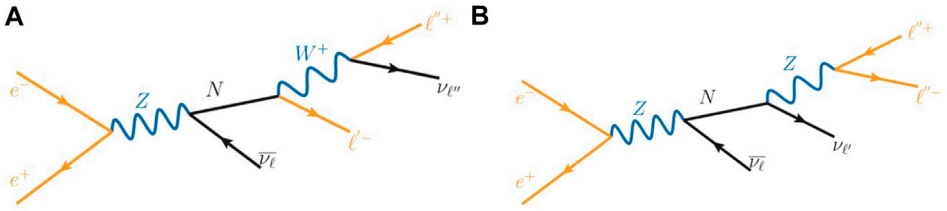


FIGURE 11
Representative diagrams depicting the $e^+e^- \rightarrow Z \rightarrow N\nu_\ell$ process at leading order, with N decaying *via* **(A)** charged current and **(B)** neutral current channels to the two-neutrino, two-charged lepton final state.

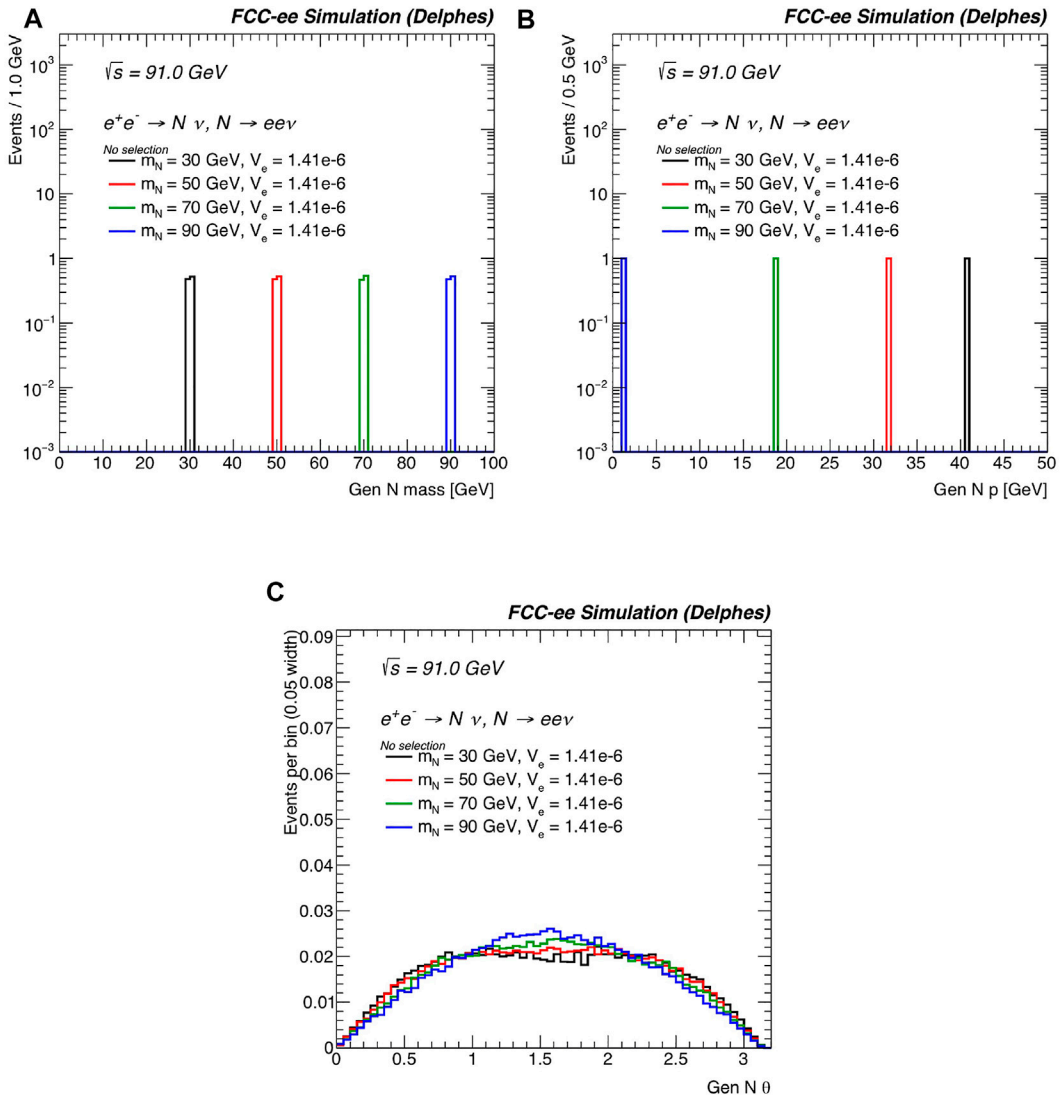


FIGURE 12
For the processes $e^+e^- \rightarrow N\nu_e + N\overline{\nu_e}$ with $N \rightarrow e^+e^-\nu_e + e^+e^-\overline{\nu_e}$ at $\sqrt{s}$ = 91 GeV, the generator-level distributions of **(A)** the invariant mass of N, **(B)** the magnitude of N's three-momentum in the lab frame, and **(C)** the polar angle of N with respect to the beam axis in the lab frame are shown, for representative HNL masses and representative active-sterile mixing $|V_{eN}| = 1.41 \times 10^{-6}$. The distributions are normalized to unit area.

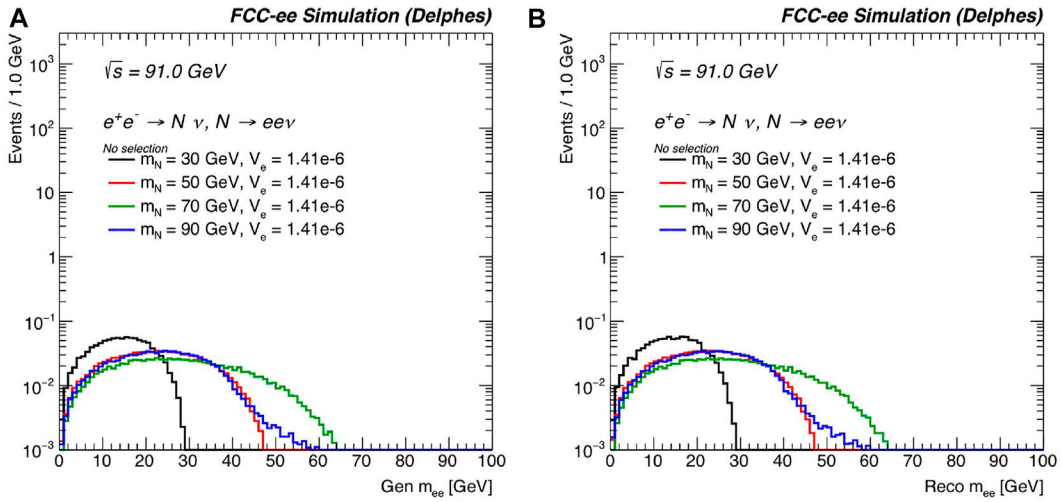


FIGURE 13
For the same processes and benchmark mass and $|V_{eN}|$ choices as in Figure 12, the differential distributions with respect to the invariant mass of the (e^+e^-) system m_{ee} at **(A)** the generator level and **(B)** after reconstruction. No selection criteria have been applied. The distributions are normalized to unit area.

the decay path and of the time-of-flight provides a determination of both the mass and proper decay time on an event-by-event basis. A detailed simulation of the process is thus of great interest to understanding how much statistics are required, first to establish the existence of the new particle, and then to establish the possible existence of a lepton number violating process (Majorana vs. Dirac nature). This also leads to the identification of specific detector requirements to optimize the discovery potential.

3.2.1 Production and kinematics of electroweak-scale HNLs

As a first step to exploring the sensitivity of FCC-ee to EW-scale HNLs, Table 1 shows the cross section (center column) and the expected number of events (right column) for an HNL with a mass of $m_N = 50$ GeV when produced and decayed through the process described in Eq. 22 and shown in Figure 11.

Results are shown for several choices of active-sterile mixing $|V_{eN}|$, and assume that an integrated luminosity of 150 ab^{-1} is collected during the Tera-Z run of FCC-ee [3]. No event selection is applied at this stage.

The kinematics of HNLs in the $m_N = 20$–90 GeV mass range at FCC-ee can also be studied. Figure 12 shows the baseline kinematics distributions of N when no event selection is applied at this stage. Here and below, active-sterile mixing of $|V_{eN}| = 1.41 \times 10^{-6}$ for representative masses of $m_N = 30$ (50) [70] {90} GeV is assumed.

Figure 12A shows the generator-level invariant mass of the HNL, which aligns with the pole mass of N. In Figure 12B, the magnitude of the normalized, generator-level three-momentum $|\vec{p}_N|$ in the lab frame is presented. From elementary kinematics, $|\vec{p}_N|$ is given analytically for a massless electron by the formula

$$|\vec{p}_N| = \frac{M_Z}{2}\left(1 - \frac{m_N^2}{M_Z^2}\right). \tag{26}$$

This corresponds to $|\vec{p}_N| \approx 40.7$ (31.9) [18.7] {1.2} GeV for the representative m_N under consideration and is in good agreement with the values $|\vec{p}_N| \approx 41$ (32) [19] {1.2} GeV shown in Figure 12B. Finally, the generator-level polar angle θ of N with respect to the beam axis in the lab frame is presented in Figure 12C. The distribution shows that a bulk of events feature central ($0.5 < \theta < 2.5$) HNLs, as one would expect from a high-p_T process.

To explore the potential impact of finite detector resolution, limited geometric coverage, and detector mismeasurements, Figure 13 shows the distributions with respect to the invariant mass of the (e^+e^-) system, which is given for massless electrons by the formula

$$m_{ee} = \sqrt{\left(p_{e^+} + p_{e^-}\right)^2} \approx \sqrt{2p_{e^+} \cdot p_{e^-}} = \sqrt{2E_{e^+}E_{e^-}\left(1 - \cos\theta_{ee}\right)}, \tag{27}$$

at (a) the generator level (Gen) and (b) the reconstruction level (Reco). In both cases, no selection criteria have been applied and the same representative inputs as above are assumed.

Consider first the generator-level case in Figure 13A. As both charged leptons in the final state originate from the $N \rightarrow$

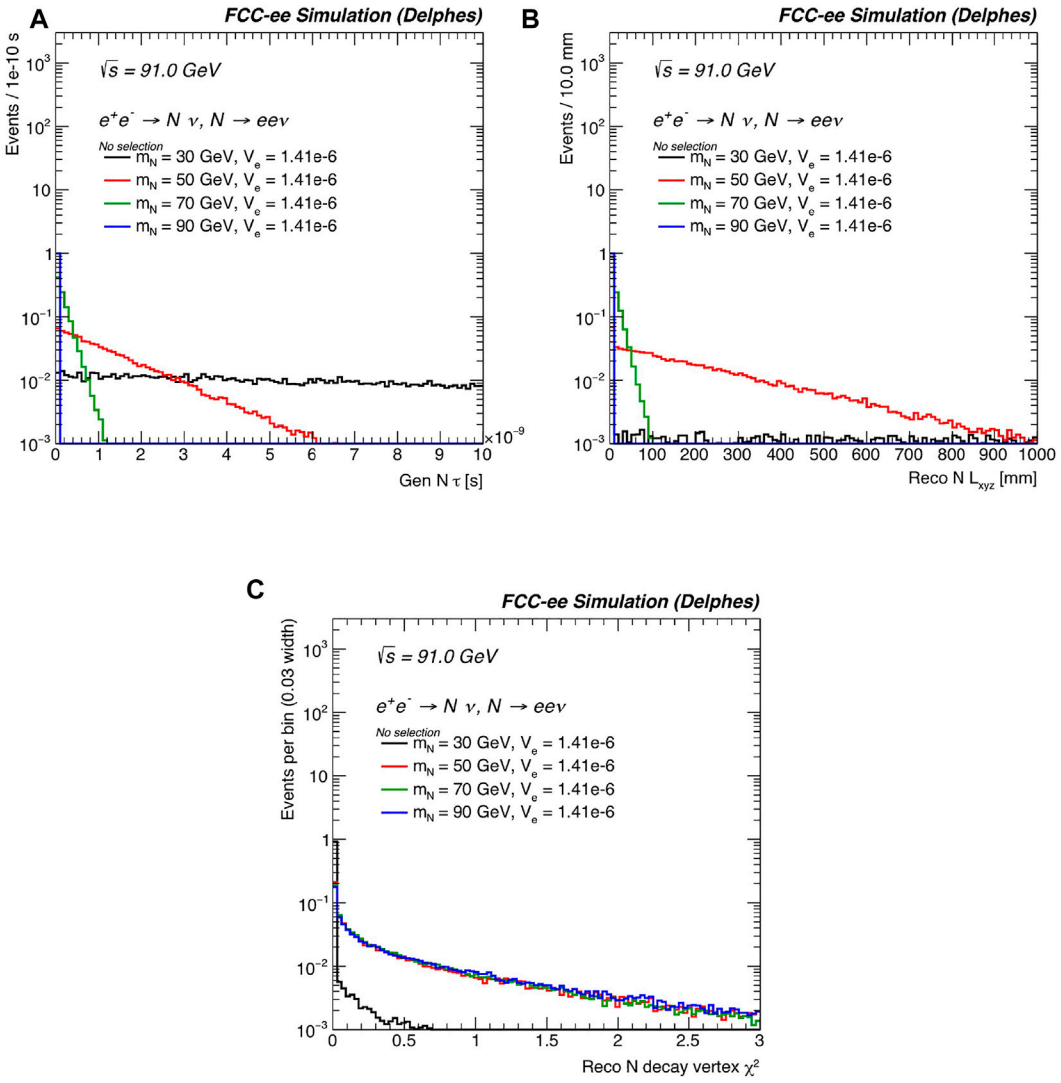


FIGURE 14
For the same processes and benchmark mass and $|V_{eN}|$ choices as in Figure 12, the differential distributions with respect to **(A)** the generator-level lifetime of N in the lab frame; **(B)** the reconstruction-level three-dimensional decay length of the N; and **(C)** the χ^2 of the reconstructed decay vertex of the HNL are shown. No selection criteria have been applied. The distributions are normalized to unit area.

e^+e^+X decay, the distribution of m_{ee} is dictated by the properties of N itself. For instance: for each of the mass benchmarks, the value of the observable m_{ee} does not exceed m_N itself, i.e., $\max(m_{ee}) < m_N$. This can be understood from momentum conservation:

$$m_N^2 = \left(p_{e^+} + p_{e^-} + p_\nu\right)^2 = p_\nu^2 + 2\left(p_\nu \cdot p_{e^+}\right) + 2\left(p_\nu \cdot p_{e^-}\right) + m_{ee}^2 \gtrsim m_{ee}^2 \,. \quad (28)$$

When m_{ee} is close to m_N, one can infer that the final-state neutrino carries little-to-no energy. For $m_N = 90$ GeV, kinematic peculiarities arise due to threshold effects. More specifically, since $\sqrt{s} = 91$ GeV, one can consider N to be essentially at rest when $m_N = 90$ GeV. For such masses, the two-body decay $N \to e^\pm W^\mp$ becomes kinematically favored. The energy of this first electron and W are given approximately by formulae similar to Eq. 26, and come out to be $E_1 \approx 9.4$ GeV and $E_W \approx 80.6$ GeV.

Assuming that the decay products of the W boson are configured in the lab frame such that the second electron carries away all the energy of W, i.e., $E_2 \approx E_W$, then the formula for m_{ee} shows that the maximum invariant mass for $m_N = 90$ GeV is about $\max(m_{ee}) \approx \sqrt{4\,(9.4\text{ GeV})\,(80.6\text{ GeV})} \approx 55$ GeV. This is in agreement with Figure 13A.

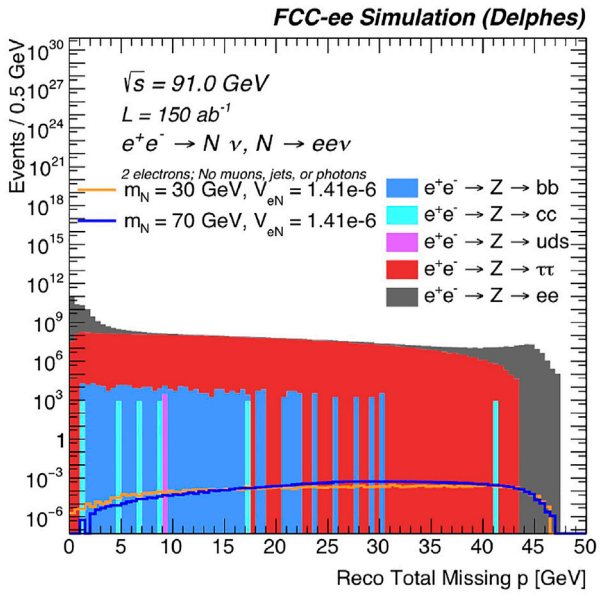


FIGURE 15
The normalized, reconstructed-level total missing momentum, for representative HNL signal benchmark mass and $|V_{eN}|$ choices, as well as background processes. Exactly two reconstructed electrons are required, as well as that there are no reconstructed muons, jets or photons in each event.

Comparing Figures 13A,B demonstrates some impact of the event reconstruction. Importantly, many of the kinematic features found at the generator level survive at the reconstruction level. In particular, the endpoints of m_{ee} are preserved. Likewise, the means of each distribution, which span about $m_{ee}^{\text{mean}} \approx 14 - 28$ GeV, remain unaltered at the reconstruction level. The relatively small impact of reconstruction effects can be tied to the high requirements of the FCC sub-detector systems.

In the absence of additional new physics, HNLs with masses below the EW scale and active-sterile mixing much smaller than unity are generically long-lived. To explore this at FCC-ee, Figure 14 shows (a) the generator-level lifetime (s) of N, given by $\tau = \gamma_N \tau_N$, where $\gamma_N = E_N/m_N$ is the Lorentz boost of N in the lab frame, and τ_N is the proper lifetime; (b) the reconstructed three-dimensional decay length (mm) of the HNL (L_{xyz}); and (c) the χ^2 of the reconstructed displaced vertex.

For a fixed width of $|V_{eN}| = 1.41 \times 10^{-6}$, different qualitative features can be observed for the representative m_N. For instance, at $m_N = 30$ GeV, characteristic generator-level lifetimes readily exceed several seconds. This implies displaced vertices can be well beyond one or more meters, and therefore outside the fiducial coverage of the IDEA detector. In these instances, a large region of the event's phase space corresponds to long-lived HNLs that ostensibly appear as missing momentum.

For heavier N, lifetimes are drastically smaller, with most HNL events exhibiting a lifetime of less than 1–2 s for $m_N \gtrsim 50$ GeV. For $m_N = 50$ (70) GeV, such lifetimes correspond to reconstructed displacements that are mostly within $L_{xyz} = 50$ (100) mm. Finally, in Figure 14C, the χ^2 curves indicate that the displaced vertices are well-reconstructed, with small χ^2 values.

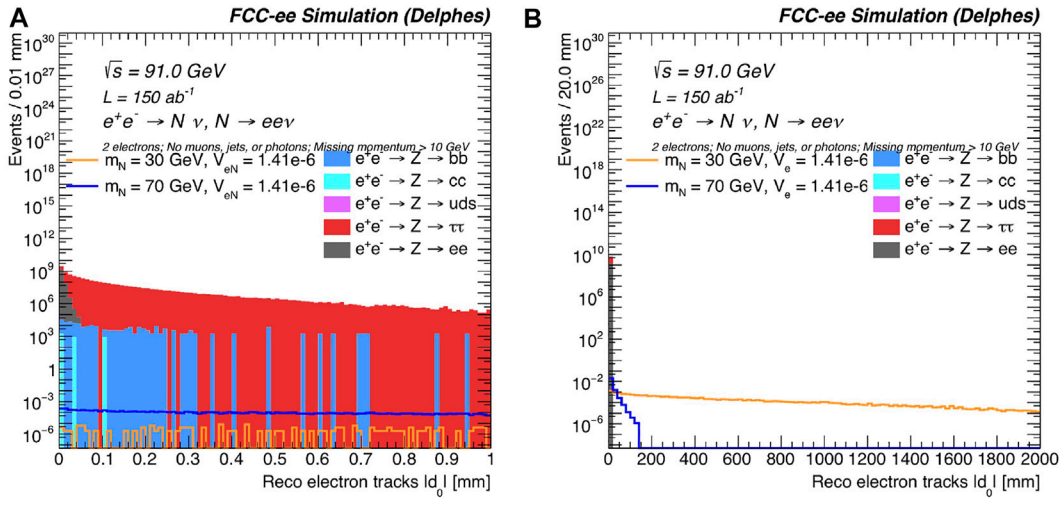


FIGURE 16
The normalized, reconstructed-level absolute value of the transverse impact parameter $|d_0|$, for representative HNL signal benchmark mass and $|V_{eN}|$ choices, as well as background processes, for **(A)** 0–1 mm in $|d_0|$ and **(B)** 0–2000 mm in $|d_0|$. Exactly two reconstructed electrons are required, as well as that there are no reconstructed muons, jets or photons in each event. The total missing momentum must be greater than 10 GeV.

TABLE 2 The expected number of events at an integrated luminosity of 150 ab^{-1} is shown for the background processes, for each selection criterion.

	Before selection	Exactly 2 reco e	Vetoes	$\not{p}$ > 10 GeV	$\lvert d_0\rvert$ > 0.5 mm
Z → ee	$2.19 \times 10^{11} \pm 6.94 \times 10^{7}$	$1.75 \times 10^{11} \pm 6.19 \times 10^{7}$	$1.53 \times 10^{11} \pm 5.80 \times 10^{7}$	$7.07 \times 10^{8} \pm 3.94 \times 10^{6}$	$\leq 3.94 \times 10^{6}$
Z → bb	$9.97 \times 10^{11} \pm 4.14 \times 10^{7}$	$5.64 \times 10^{8} \pm 9.85 \times 10^{5}$	$3.25 \times 10^{5} \pm 2.36 \times 10^{4}$	$1.22 \times 10^{5} \pm 1.45 \times 10^{4}$	$1.72 \times 10^{3} \pm 1.72 \times 10^{3}$
Z → ττ	$2.21 \times 10^{11} \pm 7.00 \times 10^{7}$	$5.49 \times 10^{9} \pm 1.10 \times 10^{7}$	$5.10 \times 10^{9} \pm 1.06 \times 10^{7}$	$2.52 \times 10^{9} \pm 7.47 \times 10^{6}$	$6.64 \times 10^{4} \pm 3.84 \times 10^{4}$
Z → cc	$7.82 \times 10^{11} \pm 2.61 \times 10^{7}$	$1.69 \times 10^{7} \pm 1.21 \times 10^{5}$	$5.22 \times 10^{3} \pm 2.13 \times 10^{3}$	$1.74 \times 10^{3} \pm 1.23 \times 10^{3}$	$\leq 1.23 \times 10^{3}$
Z → uds	$2.79 \times 10^{12} \pm 8.83 \times 10^{7}$	$2.30 \times 10^{7} \pm 2.54 \times 10^{5}$	$2.79 \times 10^{3} \pm 2.79 \times 10^{3}$	$\leq 2.79 \times 10^{3}$	$\leq 2.79 \times 10^{3}$

The cumulative number of events is shown. Only statistical uncertainty is taken into account.

TABLE 3 The expected number of events at an integrated luminosity of 150 ab^{-1} is shown for representative HNL signal benchmark masses and $|V_{eN}|$ choices, for each selection criterion.

	Before selection	Exactly 2 reco e	Vetoes	$\not{p}$ > 10 GeV	$\lvert d_0\rvert$ > 0.5 mm
m_N = 10 GeV, $\lvert V_{eN}\rvert = 2 \times 10^{-4}$	2534 ± 11	1006 ± 7	996 ± 7	951 ± 7	907 ± 7
m_N = 20 GeV, $\lvert V_{eN}\rvert = 9 \times 10^{-5}$	458 ± 2	313 ± 2	308 ± 2	293 ± 2	230 ± 1
m_N = 20 GeV, $\lvert V_{eN}\rvert = 3 \times 10^{-5}$	51.0 ± 0.2	34.7 ± 0.2	34.2 ± 0.2	32.6 ± 0.2	31.2 ± 0.2
m_N = 30 GeV, $\lvert V_{eN}\rvert = 1 \times 10^{-5}$	5.01 ± 0.02	3.85 ± 0.02	3.76 ± 0.02	3.54 ± 0.02	3.39 ± 0.02
m_N = 50 GeV, $\lvert V_{eN}\rvert = 6 \times 10^{-6}$	1.23 ± 0.01	0.99 ± 0.01	0.96 ± 0.01	0.92 ± 0.01	0.729 ± 0.004

The cumulative number of events is shown. Only statistical uncertainty is taken into account.

3.2.2 Backgrounds and event selection

Several backgrounds to the HNL processes described in Eq. 22a, Eq. 22bare considered, namely, Z bosons that decay to electron-positron pairs, to tau pairs, to light quarks, to charm quark pairs, and to b quark pairs. These background processes were simulated with the conditions described above.

Figures 15, 16 show distributions of variables that distinguish the HNL signal from these background processes. Figure 15 shows the total missing momentum $\not{p}$ in each event. Unlike in a hadron collider, where only the missing momentum in the transverse direction can be considered, the three-dimensional missing momentum can be used at FCC-ee. As can be seen from this figure, requiring $\not{p}$ > 10 GeV significantly reduces the background contributions while maintaining a high efficiency for the HNL signal.

Figure 16 shows the electron-track transverse impact parameter $|d_0|$ for each event. The transverse impact parameter is the distance of closest approach in the transverse plane of the helical trajectory of the track with respect to the beam axis; it is a measurement of the reconstructed electron's displacement. Requiring that both electron tracks have $|d_0|$ > 0.5 mm removes the vast majority of the background.

Taking these and other distributions into account, a simple event selection is developed, using reconstructed-level variables. Events must have exactly two electrons, and no photons, jets, or muons. These requirements substantially reduce the background from light and heavy quarks. We next require $\not{p}$ > 10 GeV, which is particularly effective at reducing $Z \to ee$ events with spurious missing momentum associated with finite detector resolution. Finally, we require that both electrons are displaced with $|d_0|$ > 0.5 mm to remove the vast majority of the remaining (prompt) backgrounds.

Table 2 shows the expected number of background events for each cumulative selection criterion, and Table 3 shows the same for representative HNL signal benchmark masses and $|V_{eN}|$ choices, assuming an integrated luminosity of 150 ab^{-1}. Within these limitations, these tables show that after all the selection criteria are applied, the background can be substantially reduced while the majority of the signal events are retained. After all the selection criteria are applied, we can expect about 1 event for an HNL with a mass of 50 GeV and $|V_{eN}| = 6 \times 10^{-6}$, with an integrated luminosity of 150 ab^{-1}. This benchmark point is illustrative of the maximum sensitivity to long-lived HNLs that can be achieved at FCC-ee, with the current study.

3.2.3 Majorana and Dirac nature of the HNL

If HNLs exist in nature, a chief goal is to ascertain whether they are Dirac or Majorana fermions. As discussed in Section 2.1 and elsewhere [33, 120–122], determining this is tantamount to observing processes that are mediated by N and exhibit LNV. However, at FCC-ee, the net lepton numbers of the processes $e^+e^- \to N\nu_\ell + N\overline{\nu_\ell}$ with $N \to$ (anything) are

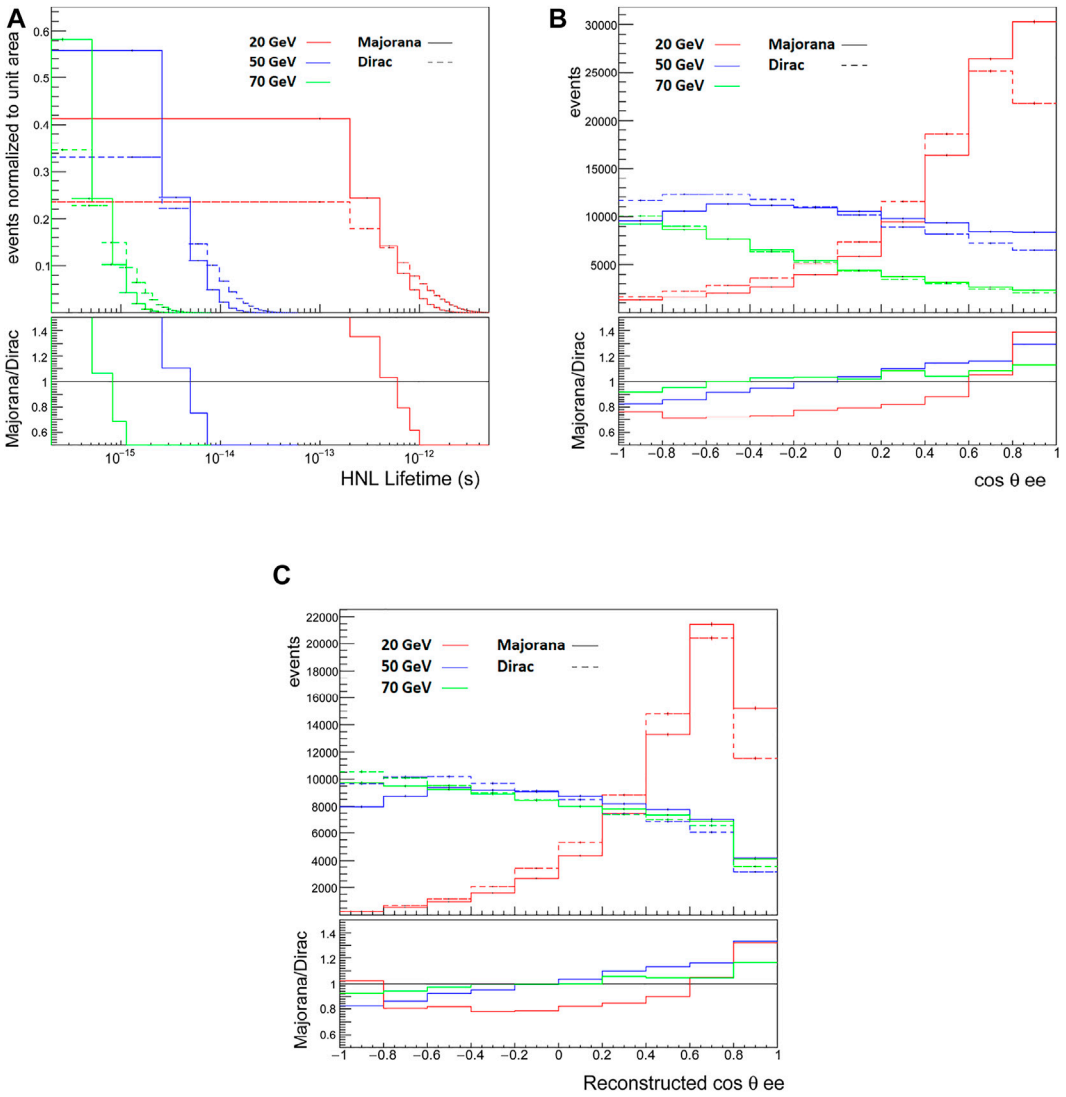


FIGURE 17
(A) The normalized distribution of lifetimes of Dirac (dashed) and Majorana (solid) HNLs in the processes defined in Eq. 22, for representative masses and assuming $|V_{eN}| = 10^{-3}$. **(B)** The generator-level angular separation $\cos\theta_{ee}$ for Dirac and Majorana HNLs under the same scenario. **(C)** Same as **(B)** but at the reconstruction level.

hidden because light neutrinos are not detected. This implies other metrics, such as angular distributions, are needed to disentangle the situation when lepton number violating states cannot be unambiguously identified.

To demonstrate the ability of FCC-ee to potentially disentangle the Dirac or Majorana nature of HNLs, the cleanest fully-leptonic decay channels are studied; the semileptonic decay channels have about twice as large a branching ratio and will be considered in future studies. Figure 17 shows the comparison of generator- and reconstruction-level observables for the two processes defined in Eq. 22. An important distinction to reiterate is that the Majorana HNL channel (solid line) includes final states that are both lepton number-conserving ($e^+e^-\nu_e\overline{\nu_e}$) as well as final states that are lepton number-violating ($e^+e^-\nu_e\nu_e$, $e^+e^-\overline{\nu_e\nu_e}$). On the other hand, the Dirac HNL channel (dashed line) consists only of final states that are lepton number-conserving ($e^+e^-\nu_e\overline{\nu_e}$). Therefore, kinematical differences amount to differences between LNV and LNC.

Figure 17A shows the normalized distribution of lifetimes of Dirac and Majorana HNLs for representative masses and

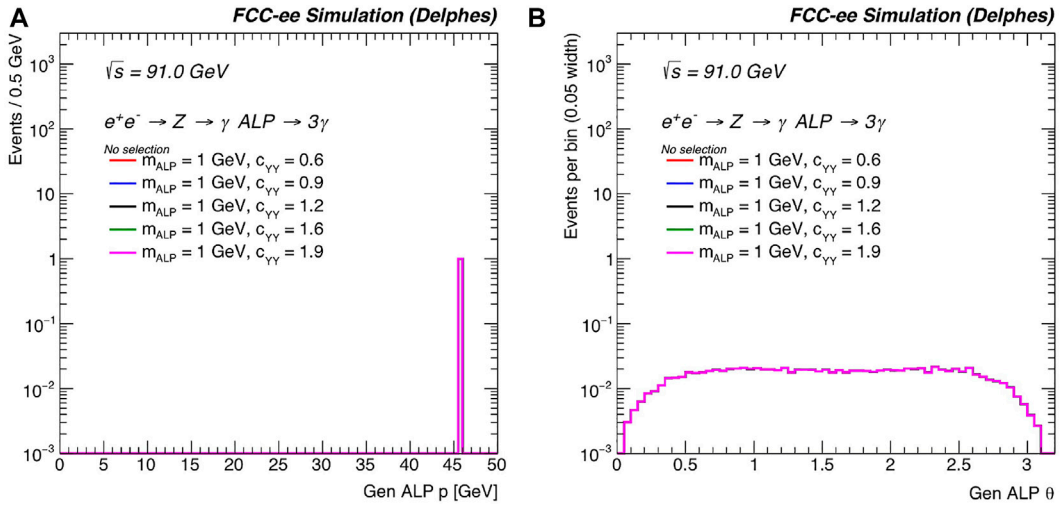


FIGURE 18
(A) Generated ALP momentum and **(B)** θ for m_{ALP} = 1 GeV and several benchmark choices of c_{YY}. The distributions are normalized to unit area.

assuming $|V_{eN}| = 10^{-3}$. Systematically, the lifetimes for Dirac N are twice as large as for the Majorana case. For m_N = 20–70 GeV, the lifetimes are roughly $\tau \sim \mathcal{O}(10^{-11})$ – $\mathcal{O}(10^{-15})$ s. As shown in Eqs 24, 25, the lifetime measurement can be used together with the total cross section to distinguish between the Dirac or Majorana nature of the observed HNL, because the combination of mixing angles that appears in both quantities is the same. For this to be done correctly, two more conditions must be met. First, the mass of the HNL must be known; this can be done by direct reconstruction, possibly combined with a kinematic fit if the HNL decays within the good quality tracker and calorimeter volumes. For longer lifetimes, a combination of decay length and laboratory decay time should be sufficient. Second, the event selection must have similar and well understood efficiencies for the three lepton flavors e, μ, τ, so that the differences can be corrected.

For the same scenario, Figure 17B shows the angular separation $\cos\theta_{ee}$ of the e^+e^- pair at the generator level. Here, several features can be observed. First, for small (large) m_N, the e^+e^- pair are largely collimated (back-to-back). This behavior can be understood from the kinematics: a heavier N is produced with less three-momentum, leading to three-body decays that are more isotropically distributed, whereas a lighter N is produced with more energy, which leads to more collimated decay products. The second feature that can be observed is that differences between the Majorana channel (LNC + LNV) and the Dirac channel (LNC) can reach $\mathcal{O}(\pm 30\%)$. Differences are largest when the e^+e^- pair are collimated ($\cos\theta_{ee} \approx 1$) or back-to-back ($\cos\theta_{ee} \approx -1$), and are smallest when they are orthogonal ($\cos\theta_{ee} \approx 0$).

Finally, Figure 17C shows the same angular separation at the reconstruction level. Again, several features can be observed. First is that reconstruction requirements markedly impact the $\cos\theta_{ee}$. In particular, isolation requirements significantly reduce cases where the e^+e^- pair are collimated ($\cos\theta_{ee} \approx 1$). Overall, the distribution for m_N = 50 GeV and m_N = 70 GeV become essentially indistinguishable. Moreover, differences between the Majorana channel (LNC + LNV) and the Dirac channel (LNC) can reduce to the $\mathcal{O}(\pm 20\%)$ level.

3.3 Axion-like particles

Figure 18 shows the generated ALP kinematics for m_{ALP} = 1 GeV and several benchmark choices of the hypercharge coupling c_{YY}. Figure 19 shows the generated ALP mass (m_{ALP}) and the invariant mass of the two-photon system ($m_{\gamma\gamma}$), and Figure 20 shows the generated three-dimensional lifetime τ_{xyz} and decay length L_{xyz} for the ALP signal. These variables are useful in distinguishing the ALP signal from background, and also for different values of the ALP mass and couplings. In addition, calorimeter and precision timing variables will be extremely helpful to include in this study

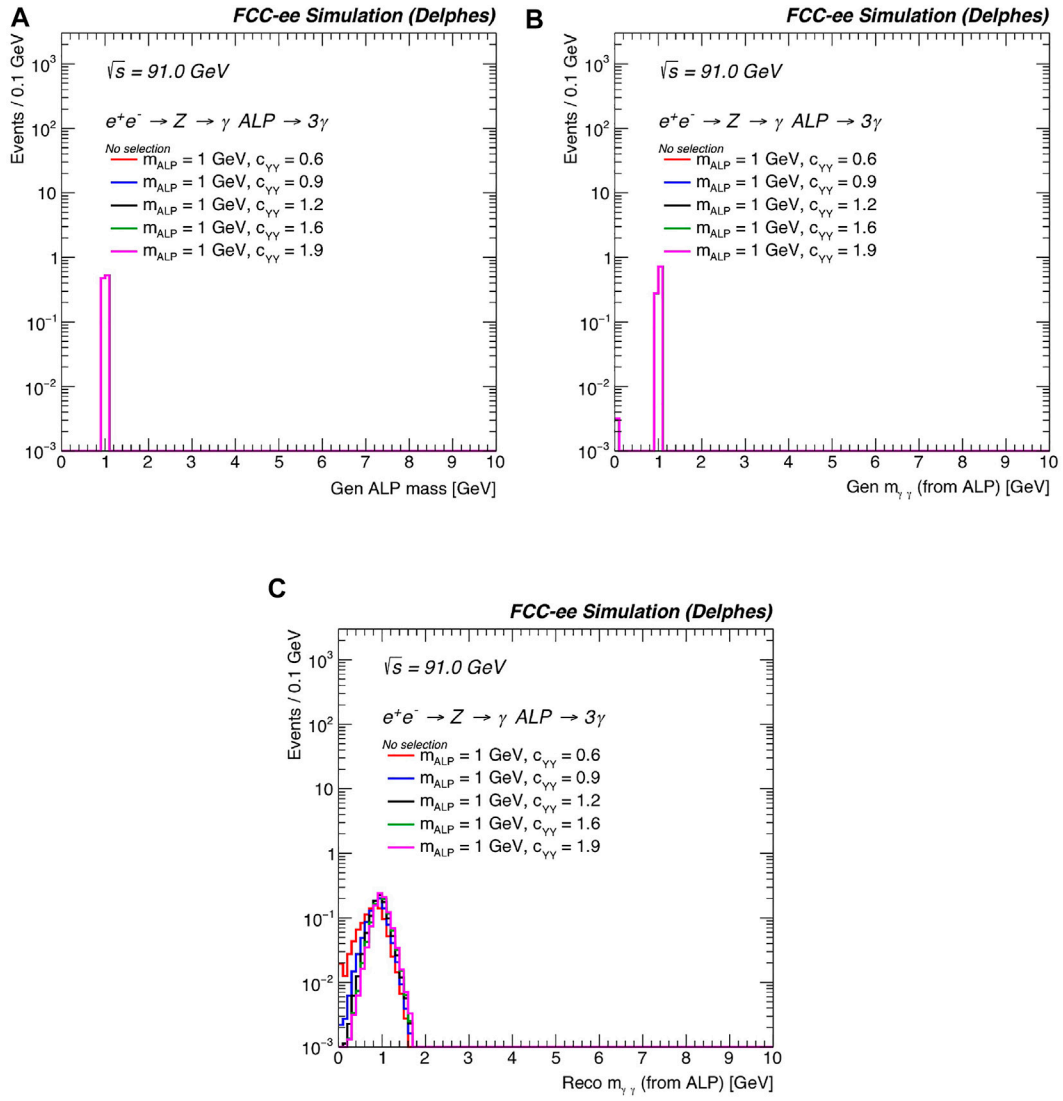


FIGURE 19
(A) Generated m_{ALP}, **(B)** generated $m_{\gamma\gamma}$, and **(C)** reconstructed $m_{\gamma\gamma}$ for m_{ALP} = 1 GeV and several benchmark choices of $c_{\gamma\gamma}$. The distributions are normalized to unit area.

of ALPs that decay to photons. We leave these studies to a later date.

3.4 Exotic Higgs boson decays

Exotic decays of Higgs bosons to LLPs are also an interesting experimental case study at FCC-ee. As was pointed out in Section 2.3, hadronic final states play a significant role, and so we plan to simulate this physics benchmark in a future paper.

3.5 Additional detectors for long-lived particles

It is possible to envisage up to four FCC-ee detectors, two of which sitting in the very large caverns foreseen from the start for the subsequent hadron collider detectors. The caverns are foreseen to be deep (200–300 m) underground, considerably reducing the cosmic ray backgrounds. A detector fully optimized for this important discovery possibility can thus be considered [83, 110, 194].

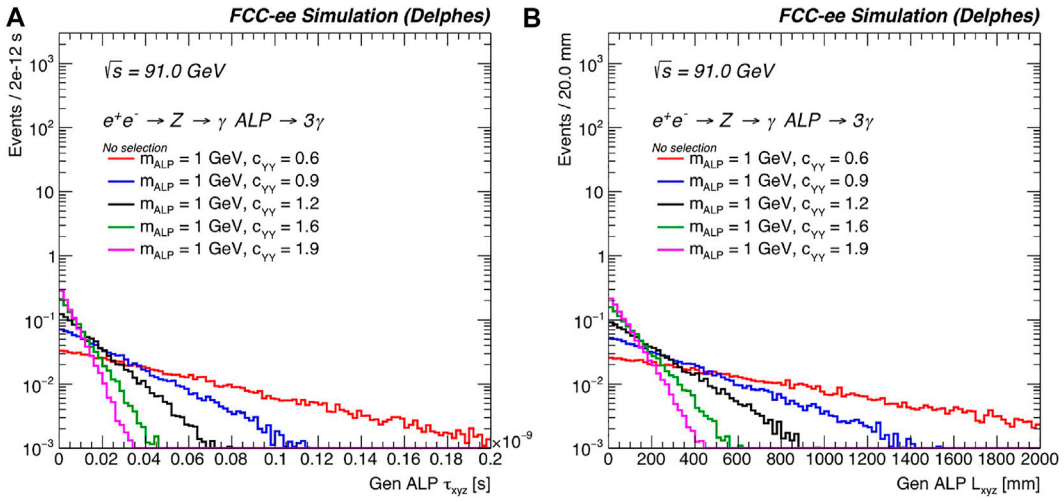


FIGURE 20
(A) Generated ALP τ_{xyz} and **(B)** L_{xyz} for m_{ALP} =1 GeV and several benchmark choices of $c_{\gamma\gamma}$. The distributions are normalized to unit area.

4 Summary and conclusion

In this paper, we discuss three key BSM cases at the future FCC-ee that experimentally can display long-lived signatures: heavy neutral leptons, axion-like particles, and exotic Higgs boson decays. While FCC-ee is primarily envisioned as a precision collider, the discussed scenarios are examples of direct searches that could be performed and could answer central questions of particle physics.

The three cases are carefully discussed from a theoretical perspective, representing the state-of-the-art and current best expected limits. Simulation studies are then presented for HNLs and ALPs. These two BSM cases can present displaced signatures: a displaced vertex for the former, and a displaced photon pair in the latter.

Different HNL signals—as well as a limited collection of background processes—are generated, kinematic variables are explored, and a first possible set of requirements is presented to isolate the signal from the SM backgrounds. Possible kinematic variables that could characterize an HNL as Dirac or Majorana are also explored experimentally.

For ALPs, signals are generated and validated and some key distributions are presented.

The work presented here can be expanded into more detailed studies, such as also including exotic Higgs boson decays, additional HNL decay channels, larger simulated samples, the use of timing information, and alternative detector designs. The simulation work presented here represents the first step towards a comprehensive evaluation of the experimental potential of FCC-ee in direct searches for BSM. Possible limitations could be solved by innovative experimental solutions that could boost the reach of FCC-ee for other non-standard signals.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

Some of the contributions are stated in the paper itself. Others involve collaboration and cross-checks. Work produced for the master theses of LR and TS is included in the paper.

Funding

RS is supported by the Swedish Research Council (VR 2017-05092). SK is supported by the Austrian Science Fund Elise-Richter grant project number V592-N27. RR acknowledges the support of the Polska Akademia Nauk (grant agreement PAN.BFD.S.BDN. 613. 022.

2021—PASIFIC 1, POPSICLE). This work has received funding from the European Union's Horizon 2020 research and innovation program under the Skłodowska-Curie grant agreement No. 847639 and from the Polish Ministry of Education and Science. AT was supported by the Australian Research Council through the ARC Discovery Project, DP210101900.

Acknowledgments

The authors would like to thank S. Petcov (SISSA/INFN) for the discussions during the writing of the paper and for providing comments to the draft.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Alimena J, Beacham J, Borsato M, Cheng Y, Vidal X. C., Cottin G et al. Searching for long-lived particles beyond the standard model at the Large Hadron Collider. *J Phys G* (2020) 47:090501. 1903.04497.

2. Group ES. Update of the European Strategy for particle physics (brochure). *Tech Rep Cern-esu-015* (2020). CERN, Geneva. doi:10.17181/CERN.JSC6.W89E

3. Benedikt M, Blondel A, Brunner O, Capeans Garrido M, Cerutti F, Gutleber J et al. FCC-ee: The lepton collider: Future circular collider conceptual design report volume 2. Future Circular Collider, Tech. Rep. CERN-ACC-2018-0057. 2, CERN, Geneva, 2018. doi:10.1140/epjst/e2019-900045-4

4. Chrzaszcz M, Gonzalez Suarez R, Monteil S. Hunt for rare processes and long-lived particles at FCC-ee. *Eur Phys J Plus* (2021) 136(1056). [2106.15459]. doi:10.1140/epjp/s13360-021-01961-4

5. Abdullahi AM, Alzas PB, Batell B, Boyarsky A, Carbajal S, Chatterjee A et al. The present and future status of heavy neutral leptons. *Snowmass Summer Study* (2022) 3:08039. 2203.,

6. Super-Kamiokande collaboration,Fukuda Y, Havakawa T, Ichihara I, Inoue K, Ishihara K, Ishino H et al. Evidence for oscillation of atmospheric neutrinos. *Phys Rev Lett* (1998) 81:1562–7. [hep-ex/9807003].

7. SNO collaboration,Ahmad QR, Allen RC, Andersen TC, D Anglin J, Barton JC, Beier EW et al. Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory. *Phys Rev Lett* (2002) 89:011301. [nucl-ex/0204008]. doi:10.1103/PhysRevLett.89.011301

8. Feruglio F. Pieces of the flavour puzzle. *Eur Phys J C* (2015) 75373. [1503.04071]. doi:10.1140/epjc/s10052-015-3576-5

9. Ma E. Pathways to naturally small neutrino masses. *Phys Rev Lett* (1998) 81:1171–4. [hep-ph/9805219]. doi:10.1103/physrevlett.81.1171

10. Agrawal P, Bauer M, Beacham J, Berlin A, Boyarsky A, Cebrian S et al. Feebly-interacting particles: FIPs 2020 workshop report. *Eur Phys J C* (2021) 81:1015. [2102.12143]. doi:10.1140/epjc/s10052-021-09703-7

11. Minkowski P. *$\mu \rightarrow e\gamma$ at a Rate of one Out of 10^9 muon decays? Phys Lett B* (1977) 67:421–8. doi:10.1016/0370-2693(77)90435-x

12. Glashow SL. The future of elementary particle physics. *NATO Sci Ser B* (1980) 61:687.

13. Gell-Mann M, Ramond P, Slansky R. Complex spinors and unified theories. *Conf Proc C* (1979) 790927:315–21. [1306.4669].

14. Mohapatra RN, Senjanovic G. Neutrino mass and spontaneous parity nonconservation. *Phys Rev Lett* (1980) 44:912–5. doi:10.1103/physrevlett.44.912

15. Yanagida T. Horizontal symmetry and masses of neutrinos. *Prog Theor Phys* (1980) 64:1103–5. doi:10.1143/ptp.64.1103

16. Schechter J, Valle JWF. Neutrino masses in SU(2) ⊗ U(1) theories. *Phys Rev D* (1980) 22:2227–35. doi:10.1103/physrevd.22.2227

17. Shrock RE. General theory of weak processes involving neutrinos. I. Leptonic pseudoscalar-meson decays, with associated tests for, and bounds on, neutrino masses and lepton mixing. *Phys Rev D* (1981) 24:1232–74. doi:10.1103/physrevd.24.1232

18. Konetschny W, Kummer W. Nonconservation of total lepton number with scalar bosons. *Phys Lett B* (1977) 70:433–5. doi:10.1016/0370-2693(77)90407-5

19. Magg M, Wetterich C. Neutrino mass problem and gauge hierarchy. *Phys Lett B* (1980) 94:61–4. doi:10.1016/0370-2693(80)90825-4

20. Cheng TP, Li L-F. Neutrino masses, mixings, and oscillations in SU(2)×U(1) models of electroweak interactions. *Phys Rev D* (1980) 22:2860–8. doi:10.1103/physrevd.22.2860

21. Lazarides G, Shafi Q, Wetterich C. Proton lifetime and fermion masses in an SO(10) model. *Nucl Phys B* (1981) 181:287–300. doi:10.1016/0550-3213(81)90354-0

22. Mohapatra RN, Senjanovic G. Neutrino masses and mixings in gauge models with spontaneous parity violation. *Phys Rev D* (1981) 23:165–80. doi:10.1103/physrevd.23.165

23. Foot R, Lew H, He XG, Joshi GC. Seesaw neutrino masses induced by a triplet of leptons. *Z Phys C - Particles Fields* (1989) 44:441–4. doi:10.1007/bf01415558

24. Atre A, Han T, Pascoli S, Zhang B. The search for heavy Majorana neutrinos. *JHEP* (2009) 05030. 0901.3589.

25. Deppisch FF, Bhupal Dev PS, Pilaftsis A. Neutrinos and collider physics. *New J Phys* (2015) 17:075019. [1502.06541]. doi:10.1088/1367-2630/17/7/075019

26. Cai Y, Han T, Li T, Ruiz R. Lepton number violation: Seesaw models and their collider tests. *Front Phys* (2018) 6:40. [1711.02180]. doi:10.3389/fphy.2018.00040

27. Bolton PD, Deppisch FF, Bhupal Dev PS. Neutrinoless double beta decay versus other probes of heavy sterile neutrinos. *JHEP* (2020) 03170. 1912.03058.

28. Wolfenstein L. Different varieties of massive Dirac neutrinos. *Nucl Phys B* (1981) B186:147–52. doi:10.1016/0550-3213(81)90096-1

29. Petcov ST. *On pseudo-Dirac neutrinos, neutrino Oscillations and neutrinoless double β-decay. Phys Lett B* (1982) 110B:245–9. doi:10.1016/0370-2693(82)91246-1

30. Shaposhnikov M. A possible symmetry of the. *Nucl Phys B* (2007) 763:49–59. [hep-ph/0605047]. doi:10.1016/j.nuclphysb.2006.11.003

31. Kersten J, Smirnov AY. Right-handed neutrinos at CERN LHC and the mechanism of neutrino mass generation. *Phys Rev D* (2007) 76:073005. [0705.3221]. doi:10.1103/physrevd.76.073005

32. Ibarra A, Molinaro E, Petcov ST. *TeV scale see-saw Mechanisms of neutrino mass generation, the Majorana Nature of the heavy singlet Neutrinos and $(\beta\beta)_{0\nu}$-Decay. JHEP* (2010) 09108. [1007.2378].

33. Moffat K, Pascoli S, Weiland C, (2017) Equivalence between massless neutrinos and lepton number conservation in fermionic singlet extensions of the Standard Model, 1712.07611.

34. Anamiati G, Hirsch M, Nardi E. Quasi-Dirac neutrinos at the LHC. *JHEP* (2016) 10010. [1607.05641].

35. Tsai Y-S. *Decay Correlations of Heavy Leptons in e+ e- —> Lepton+ Lepton-. Phys Rev D* (1971) 4:2821.

36. Gronau M, Leung CN, Rosner JL. Extending limits on neutral heavy leptons. *Phys Rev D* (1984) 29:2539–58. doi:10.1103/physrevd.29.2539

37. Petcov ST. Possible signature for production of Majorana particles in e^+e^- and p-$\bar{p}$collisions Collisions. *Phys Lett B* (1984) B 139:421–6. doi:10.1016/0370-2693(84)91844-6

38. Dittmar M, Santamaria A, Gonzalez-Garcia MC, Valle JWF. *Production Mechanisms and Signatures of isosinglet neutral heavy Leptons in Z^0 decays*. *Nucl Phys B* (1990) 332:1–19. doi:10.1016/0550-3213(90)90028-c

39. FCC-ee study Team collaboration,Blondel A, Graverini E, Serra N, Shaposhnikov M. Search for heavy right handed neutrinos at the FCC-ee. *Nucl Part Phys Proc* (2016) 1883–90. [1411.5230]. doi:10.1016/j.nuclphysbps.2015.09.304

40. Duarte L, Peressutti J, Sampayo OA. Majorana neutrino decay in an effective approach. *Phys Rev D* (2015) 92:093002. [1508.01588]. doi:10.1103/physrevd.92.093002

41. Banerjee S, Dev PSB, Ibarra A, Mandal T, Mitra M. Prospects of heavy neutrino searches at future lepton colliders. *Phys Rev D* (2015) 92:075002. doi:10.1103/physrevd.92.075002

42. Duarte L, Romero I, Peressutti J, Sampayo OA. Effective Majorana neutrino decay. *Eur Phys J C* (2016) 76453. [1603.08052]. doi:10.1140/epjc/s10052-016-4301-8

43. Antusch S, Cazzato E, Fischer O. Displaced vertex searches for sterile neutrinos at future lepton colliders. *JHEP* (2016) 12:007. [1604.02420].

44. Antusch S, Cazzato E, Drewes M, Fischer O, Garbrecht B, Gueter D et al. Probing leptogenesis at future colliders. *JHEP* (2018) 09:124. [1710.03744].

45. Duarte L, Zapata G, Sampayo OA. *Final taus and initial state polarization signatures from effective interactions of Majorana neutrinos at future e^+e^- colliders*. *Eur Phys J C* (2019) 79240. [1812.01154].

46. Ding J-N, Qin Q, Yu F-S. *Heavy neutrino searches at future Z-factories*. *Eur Phys J C* (2019) 79:766. 1903.766. doi:10.1140/epjc/s10052-019-7277-3

47. Barducci D, Bertuzzo E, Caputo A, Hernandez P, Mele B. The see-saw portal at future Higgs Factories. *JHEP* (2021) 03:117. 2011.04725.

48. Blondel A, de Gouvêa A, Kayser B. Z-boson decays into Majorana or Dirac heavy neutrinos. *Phys Rev D* (2021) 104:055027. [2105.06576]. doi:10.1103/physrevd.104.055027

49. Zapata G, Urruzola T, Sampayo OA, Duarte L. Lepton collider probes for Majorana neutrino effective interactions. *Eur Phys J C* (2022) 82(544). [2201.02480]. doi:10.1140/epjc/s10052-022-10448-0

50. Shen Y-F, Ding J-N, Qin Q. Monojet search for heavy neutrinos at future Z-factories. *Eur Phys J C* (2022) 82:398. [2201.05831]. doi:10.1140/epjc/s10052-022-10301-4

51. Barducci D, Bertuzzo E. The see-saw portal at future Higgs factories: The role of dimension six operators. *JHEP* (2022) 06:077. [2201.11754].

52. Pati JC, Salam A. Lepton number as the fourth color. *Phys Rev D* (1974) 10:275–89. doi:10.1103/physrevd.10.275

53. Mohapatra RN, Pati JC. Left-right gauge symmetry and an isoconjugate model ofCPviolation. *Phys Rev D* (1975) 11:566–71. doi:10.1103/physrevd.11.566

54. Mohapatra RN, Pati JC. Natural left-right symmetry. *Phys Rev D* (1975) 11:2558–61. doi:10.1103/physrevd.11.2558

55. Senjanovic G, Mohapatra RN. Exact left-right symmetry and spontaneous violation of parity. *Phys Rev D* (1975) 12:1502–5. doi:10.1103/physrevd.12.1502

56. Senjanovic G. Spontaneous breakdown of parity in a class of gauge theories. *Nucl Phys B* (1979) 153:334–64. doi:10.1016/0550-3213(79)90604-7

57. Langacker P. Grand unified theories and proton decay. *Phys Rep* (1981) 72:185–385. doi:10.1016/0370-1573(81)90059-4

58. Hewett JL, Rizzo TG. Low-energy phenomenology of superstring inspired E(6) models. *Phys Rep* (1989) 183:193–381. doi:10.1016/0370-1573(89)90071-9

59. Faraggi AE, Nanopoulos DV. *A SUPERSTRING Z' AT O (1-TeV) ? Mod Phys Lett A* (1991) 6:61–8. doi:10.1142/s0217732391002621

60. Buchmuller W, Greub C, Minkowski P. Neutrino masses, neutral vector bosons and the scale of B-L breaking. *Phys Lett B* (1991) 267:395–9. doi:10.1016/0370-2693(91)90952-m

61. Grzadkowski B, Iskrzynski M, Misiak M, Rosiek J. Dimension-Six terms in the standard model Lagrangian. *JHEP* (2010) 10:085. [1008.4884.

62. Brivio I, Trott M. The standard model as an effective field theory. *Phys Rep* (2019) 793:1–98. [1706.08945]. doi:10.1016/j.physrep.2018.11.002

63. Gauld R, Pecjak BD, Scott DJ. One-loop corrections to h\to b\bar{b} and h\to \tau\bar \tau *decays in the standard model dimension-6 EFT: Four-fermion operators and the large-m_t limit*. *J High Energ Phys* (2016) 05:080. [1512.02508]. doi:10.1007/jhep05(2016)080

64. del Aguila F, Bar-Shalom S, Soni A, Wudka J. Heavy Majorana neutrinos in the effective Lagrangian description: Application to hadron colliders. *Phys Lett B* (2009) 670:399–402. [0806.0876]. doi:10.1016/j.physletb.2008.11.031

65. Aparici A, Kim K, Santamaria A, Wudka J. Right-handed neutrino magnetic moments. *Phys Rev D* (2009) 80:013010. [0904.3244]. doi:10.1103/physrevd.80.013010

66. Elgaard-Clausen G, Trott M. On expansions in neutrino effective field theory. *JHEP* (2017) 11(088). [1703.04415.

67. Hernández P, Kekic M, López-Pavón J, Racker J, Salvado J. Testable baryogenesis in seesaw models. *JHEP* (2016) 08157. [1606.06719].

68. Drewes M, Garbrecht B, Gueter D, Klaric J. Testing the low scale seesaw and leptogenesis. *JHEP* (2017) 08:018. 1609.09069.

69. Drewes M, Hajer J, Klaric J, Lanfranchi G. NA62 sensitivity to heavy neutral leptons in the low scale seesaw model. *JHEP* (2018) 07(105). 1801.04207.

70. Chrzaszcz M, Drewes M, Gonzalo TE, Harz J, Krishnamurthy S, Weniger C. A frequentist analysis of three right-handed neutrinos with GAMBIT. *Eur Phys J C* (2020) 80(569). 1908.569. doi:10.1140/epjc/s10052-020-8073-9

71. DUNE collaboration,Abi B, Acciarri R, Acero MA, Adamov G, Adams D, Adinolfi M et al. Deep underground neutrino experiment (DUNE), far detector technical design report, volume II: DUNE physics, 2002.03005.

72. Caputo A, Hernandez P, Kekic M, López-Pavón J, Salvado J. The seesaw path to leptonic CP violation. *Eur Phys J C* (2017) 77(258). [1611.05000]. doi:10.1140/epjc/s10052-017-4823-8

73. King SF, Luhn C. Neutrino mass and mixing with discrete symmetry. *Rep Prog Phys* (2013) 76:056201. [1301.1340]. doi:10.1088/0034-4885/76/5/056201

74. Feruglio F, Romanino A. Lepton flavor symmetries. *Rev Mod Phys* (2021) 93:015007. 1912.06028. doi:10.1103/revmodphys.93.015007

75. Xing Z-z. Flavor structures of charged fermions and massive neutrinos. *Phys Rep* (2020) 854:1–147. 1909. doi:10.1016/j.physrep.2020.02.001

76. Caputo A, Hernandez P, Lopez-Pavon J, Salvado J. The seesaw portal in testable models of neutrino masses. *JHEP* (2017) 06112. 1704.08721.

77. Esteban I, Gonzalez-Garcia M, Maltoni M, Schwetz T, Zhou A, et al. The fate of hints: Updated global analysis of three-flavor neutrino oscillations. *JHEP* (2020) 09:178. 2007.14792.

78. DUNE collaboration,Abi B, Acciarri R, Acero MA, Adamov G, Adams D, Adinolfi M et al. Long-baseline neutrino oscillation physics potential of the DUNE experiment. *Eur Phys J C* (2020) 80(978). 2006.16043.

79. Antusch S, Cazzato E, Fischer O. *Sterile neutrino searches at future e^-e^+, pp, and e^-p colliders*. *Int J Mod Phys A* (2017) 32:1750078. [1612.02728]. doi:10.1142/s0217751x17500786

80. Ellis RK, Physics briefing Book: Input for the European Strategy for particle physics update 2020, 1910.11775.

81. Lou XL, Circular electron-positron collider status and progress (presented at the hkust-ias conference on high energy physics 2022), 2022.

82. CEPC Study Group collaboration,Dong M. CEPC conceptual design report: Volume 2. *Phys Detector* (2018) 1811:10545.

83. Chrząszcz M, Drewes M, Hajer J. Hecate: A long-lived particle detector concept for the FCC-ee or CEPC. *Eur Phys J C* (2021) 81:546. 2011.01005. doi:10.1140/epjc/s10052-021-09253-y

84. CHARM collaboration,Dorenbosch J, Allaby JV, Amaldi U, Barbiellini G, Berger C, Bergsma F et al., A search for decays of heavy neutrinos in the mass range 0.5-GeV to 2.8-GeV. *Phys Lett B* (1986) 166:473–8.

85. Abela R, Daum M, Eaton GH, Frosch R, Jost B, Kettle PR et al. Search for an admixture of heavy neutrino states in pion decay. *Phys Lett B* (1981) 105:263–6. doi:10.1016/0370-2693(81)90884-4

86. Yamazaki T. Search for heavy neutrinos in kaon decay. *Conf Proc C* (1984) 840719:262.

87. E949 collaboration,Artamonov AV. *Search for heavy neutrinos in $K^+ \to \mu^+\nu_H$ decays*. *Phys Rev D* (2015) 91:052001. [1411.3963].

88. Bernardi G, Carugno G, Chauveau J, Dicarlo F, Dris M, Dumarchez J et al. Further limits on heavy neutrino couplings. *Phys Lett B* (1988) 203:332–4. doi:10.1016/0370-2693(88)90563-1

89. NuTeV, E815 collaboration,Vaitaitis A, Drucker RB, Formaggio J, Koutsoliotas S, Adams T, Alton A et al. Search for neutral heavy leptons in a high-energy neutrino beam. *Phys Rev Lett* (1999) 83:4943–6. [hep-ex/9908011]. doi:10.1103/physrevlett.83.4943

90. Vaitaitis AG. Search for neutral heavy leptons in a high-energy neutrino beam. PhD thesis. New York, NY: Columbia University (2000). doi:10.2172/1421441

91. CMS collaboration,Sirunyan AM. Search for heavy neutral leptons in events with three charged leptons in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *Phys Rev Lett* (2018) 120:221801. 1802.02965.

92. DELPHI collaboration,Abreu P. Search for neutral heavy leptons produced in Z decays. *Z. Phys. C.* (1997) 74:57–71. doi:10.1007/s002880050370

93. ATLAS collaboration,Aad G. *Search for heavy neutral leptons in decays of W bosons produced in 13 TeV pp collisions using prompt and displaced signatures with the ATLAS detector. JHEP* (2019) 10:265. 1905.09787.

94. CMS collaboration,Tumasyan A. Search for long-lived heavy neutral leptons with displaced vertices in proton-proton collisions at $\sqrt{s}$ = 13. *TeV* (2022) 2201: 05578.

95. Izaguirre E, Shuve B. Multilepton and lepton jet probes of sub-weak-scale right-handed neutrinos. *Phys Rev D* (2015) 91:093010. [1504.02470]. doi:10.1103/physrevd.91.093010

96. Drewes M, Hajer J. Heavy Neutrinos in displaced vertex searches at the LHC and HL-LHC. *JHEP* (2020) 02:070. 1903.06100.

97. Pascoli S, Ruiz R, Weiland C. Heavy neutrinos with dynamic jet vetoes: Multilepton searches at $\sqrt{s}$ = 14, 27, and 100 TeV. *JHEP* (2019) 06:049. 1812.08750.

98. Ballett P, Boschi T, Pascoli S. Heavy neutral leptons from low-scale seesaws at the DUNE near detector. *JHEP* (2020) 03:111. 1905.00284.

99. FASER collaboration,Ariga A, Ariga T, Boyd J, Cadoux F, Casper DW, Favre Y et al. FASER's physics reach for long-lived particles. *Phys Rev D* (2019) 99:095011. [1811.12522]. doi:10.1103/physrevd.99.095011

100. SHiP collaboration,Ahdida C, Albanese R, Alexandrov A, Anokhina A, Aoki S, Arduini G et al. Sensitivity of the SHiP experiment to heavy neutral leptons. *JHEP* (2019) 04:077. [1811.00930].

101. Gorbunov D, Krasnov I, Kudenko Y, Suvorov S. Heavy Neutral Leptons from kaon decays in the SHiP experiment. *Phys Lett B* (2020) 810:135817. 2004.07974. doi:10.1016/j.physletb.2020.135817

102. Curtin D, Drewes M, McCullough M, Meade P, Mohapatra RN, Shelton J et al. Long-lived particles at the energy frontier: The MATHUSLA physics case. *Rep Prog Phys* (2019) 82:116201. 1806.07396. doi:10.1088/1361-6633/ab28d6

103. Aielli G, Ben-Haim E, Cardarelli R, Charles MJ, Cid Vidal X, Coco V et al. Expression of interest for the CODEX-b detector. *Eur Phys J C* (2020) 801177. [1911.00481]. doi:10.1140/epjc/s10052-020-08711-3

104. Bay Nielsen S. Prospects of sterile neutrino search with the FCC-ee. Master's thesis. Copenhagen U. (2017).

105. Antusch S, Cazzato E, Fischer O, Hammad A, Wang K. Lepton flavor violating dilepton dijet signatures from sterile neutrinos at proton colliders. *JHEP* (2018) 10067. 1805.11400.

106. Boyarsky A, Mikulenko O, Ovchynnikov M, Shchutska L, Exploring the potential of FCC-hh to search for particles from B mesons, 2022, 2204.01622.

107. Mękała K, Reuter J, Żarnecki AF. *Heavy neutrinos at future linear e^+e^- colliders. JHEP* (2022) 06:010. 2202.06703.

108. Antusch S, Fischer O, Hammad A. Lepton-trijet and displaced vertex searches for heavy neutrinos at future electron-proton colliders. *JHEP* (2020) 03110. 1908.02852.

109. Li P, Liu Z, Lyu F-F. progress.Heavy neutral leptons at muon colliders.

110. Wang ZS, Wang K. Physics with far detectors at future lepton colliders. *Phys Rev D* (2020) 101:075046. [1911.06576]. doi:10.1103/physrevd.101.075046

111. Giffin P, Gori S, Tsai Y-D, Tuckler D, Heavy neutral leptons at beam dump experiments of future lepton colliders, 2022, 2206.13745.

112. Canetti L, Drewes M, Shaposhnikov M. Matter and antimatter in the universe. *New J Phys* (2012) 14:095012. [1204.4186]. doi:10.1088/1367-2630/14/9/095012

113. Akhmedov EK, Rubakov VA, Smirnov AY. Baryogenesis via neutrino oscillations. *Phys Rev Lett* (1998) 81:1359–62. [hep-ph/9803255]. doi:10.1103/physrevlett.81.1359

114. Asaka T, Shaposhnikov M. *The vMSM, dark matter and baryon asymmetry of the universe. Phys Lett B* (2005) 620:17–26. [hep-ph/0505013]. doi:10.1016/j.physletb.2005.06.020

115. Pilaftsis A, Underwood TEJ. Resonant leptogenesis. *Nucl Phys B* (2004) 692: 303–45. [hep-ph/0309342]. doi:10.1016/j.nuclphysb.2004.05.029

116. Drewes M, Georis Y, Klarić J. Mapping the viable parameter space for testable leptogenesis. *Phys Rev Lett* (2022) 128:051801. [2106.16226]. doi:10.1103/physrevlett.128.051801

117. Klarić J, Shaposhnikov M, Timiryasov I. Reconciling resonant leptogenesis and baryogenesis via neutrino oscillations. *Phys Rev D* (2021) 104:055010. [2103.16545]. doi:10.1103/physrevd.104.055010

118. Sabti N, Magalich A, Filimonova A. An extended analysis of heavy neutral leptons during big bang nucleosynthesis. *JCAP* (2020) 11(056). 2006.07387.

119. Boyarsky A, Ovchynnikov M, Ruchayskiy O, Syvolap V. Improved big bang nucleosynthesis constraints on heavy neutral leptons. *Phys Rev D* (2021) 104: 023517. 2008.00749. doi:10.1103/physrevd.104.023517

120. Schechter J, Valle JWF. Neutrinoless double-βdecay in SU(2)×U(1) theories. *Phys Rev D* (1982) 25:2951–4. doi:10.1103/physrevd.25.2951

121. Hirsch M, Kovalenko S, Schmidt I. Extended black box theorem for lepton number and flavor violating processes. *Phys Lett B* (2006) 642:106–10. [hep-ph/0608207]. doi:10.1016/j.physletb.2006.09.012

122. Duerr M, Lindner M, Merle A. On the quantitative impact of the schechter-valle theorem. *JHEP* (2011) 06091. [1105.0901].

123. Bray S, Lee JS, Pilaftsis A. Resonant CP violation due to heavy neutrinos at the LHC. *Nucl Phys B* (2007) 786:95–118. [hep-ph/0702294]. doi:10.1016/j.nuclphysb.2007.07.002

124. Ruiz R. Quantitative study on helicity inversion in Majorana neutrino decays at the LHC. *Phys Rev D* (2021) 103:015022. 2008.01092. doi:10.1103/physrevd.103.015022

125. de Gouvêa A, Fox PJ, Kayser BJ, Kelly KJ. Three-body decays of heavy Dirac and Majorana fermions. *Phys Rev D* (2021) 104:015038. [2104.05719]. doi:10.1103/physrevd.104.015038

126. Hernández P, Jones-Pérez J, Suarez-Navarro O. Majorana vs pseudo-Dirac neutrinos at the ILC. *Eur Phys J C* (2019) 79220. [1810.07210]. doi:10.1140/epjc/s10052-019-6728-1

127. Asaka T, Blanchet S, Shaposhnikov M. The νMSM, dark matter and neutrino masses. *Phys Lett B* (2005) 631:151–6. [hep-ph/0503065]. doi:10.1016/j.physletb.2005.09.070

128. Drewes M. The phenomenology of right handed neutrinos. *Int J Mod Phys E* (2013) 22:1330019. [1303.6912]. doi:10.1142/s0218301313300191

129. Fukugita M, Yanagida T. Barygenesis without grand unification. *Phys Lett B* (1986) 174:45–7. doi:10.1016/0370-2693(86)91126-3

130. Davidson S, Ibarra A. A Lower bound on the right-handed neutrino mass from leptogenesis. *Phys Lett B* (2002) 535(25–32):25–32. [hep-ph/0202239]. doi:10.1016/s0370-2693(02)01735-5

131. Pilaftsis A, Underwood TEJ. Electroweak-scale resonant leptogenesis. *Phys Rev D* (2005) 72:113001. [hep-ph/0506107]. doi:10.1103/physrevd.72.113001

132. Chun EJ, Cvetic G, Dev PSB, Drewes M, Fong CS, Garbrecht B et al. Probing leptogenesis. *Int J Mod Phys A* (2018) 33(1842005). [1711.02865]. doi:10.1142/s0217751x18420058

133. Dodelson S, Widrow LM. Sterile-neutrinos as dark matter. *Phys Rev Lett* (1994) 72:17–20. [hep-ph/9303287]. doi:10.1103/physrevlett.72.17

134. Drewes M, Adhikari R, Agostini M, Anh Ky N, Araki T, Archidiacono M et al. A white paper on keV sterile neutrino dark matter. *JCAP* (2017) 01:025. 1602.04816.

135. Boyarsky A, Drewes M, Lasserre T, Mertens S, Ruchayskiy O. Sterile neutrino dark matter. *Prog Part Nucl Phys* (2019) 104:1–45. [1807.07938. doi:10.1016/j.ppnp.2018.07.004

136. Shi X-D, Fuller GM. New dark matter candidate: Nonthermal sterile neutrinos. *Phys Rev Lett* (1999) 82:2832–5. [astro-ph/9810076]. doi:10.1103/physrevlett.82.2832

137. Laine M, Shaposhnikov M. Sterile neutrino dark matter as a consequence of nuMSM-induced lepton asymmetry. *JCAP* (2008) 06(031). 0804.4543.

138. Ghiglieri J, Laine M. Improved determination of sterile neutrino dark matter spectrum. *JHEP* (2015) 11171. [1506.06752].

139. Venumadhav T, Cyr-Racine F-Y, Abazajian KN, Hirata CM. Sterile neutrino dark matter: Weak interactions in the strong coupling epoch. *Phys Rev D* (2016) 94: 043515. [1507.06655]. doi:10.1103/physrevd.94.043515

140. Shaposhnikov M. The nuMSM, leptonic asymmetries, and properties of singlet fermions. *JHEP* (2008) 08008. 0804.4542.

141. Canetti L, Drewes M, Shaposhnikov M. Sterile neutrinos as the origin of dark and baryonic matter. *Phys Rev Lett* (2013) 110:061801. [1204.3902]. doi:10.1103/physrevlett.110.061801

142. Canetti L, Drewes M, Frossard T, Shaposhnikov M. Dark matter, baryogenesis and neutrino oscillations from right handed neutrinos. *Phys Rev D* (2013) 87:093006. [1208.4607]. doi:10.1103/physrevd.87.093006

143. Ghiglieri J, Laine M. Sterile neutrino dark matter via coinciding resonances. *JCAP* (2020) 07(012). 2004.10766.

144. Kusenko A, Takahashi F, Yanagida TT. Dark matter from split seesaw. *Phys Lett B* (2010) 693:144–8. [1006.1731]. doi:10.1016/j.physletb.2010.08.031

145. Li T, Chao W. Neutrino masses, dark matter and symmetry at the LHC. *Nucl Phys B* (2011) 843:396–412. [1004.0296]. doi:10.1016/j.nuclphysb.2010.10.004

146. Escudero M, Rius N, Sanz V. *Sterile neutrino portal to Dark Matter I: The* $U(1)_{B-L}$ *case*. *JHEP* (2017) 02:045. 1606.01258.

147. Bezrukov F, Hettmansperger H, Lindner M. keV sterile neutrino Dark Matter in gauge extensions of the Standard Model. *Phys Rev D* (2010) 81:085032. [0912.4415]. doi:10.1103/physrevd.81.085032

148. Bezrukov F, Kartavtsev A, Lindner M. Leptogenesis in models with keV sterile neutrino dark matter. *J Phys G: Nucl Part Phys* (2013) 40:095202. [1204.5477]. doi:10.1088/0954-3899/40/9/095202

149. Nemevsek M, Senjanovic G, Zhang Y. Warm dark matter in low scale left-right theory. *JCAP* (2012) 07:006. 1205.0844.

150. Shaposhnikov M, Tkachev I. The vMSM, inflation, and dark matter. *Phys Lett B* (2006) 639:414–7. [hep-ph/0604236]. doi:10.1016/j.physletb.2006.06.063

151. Petraki K, Kusenko A. Dark-matter sterile neutrinos in models with a gauge singlet in the Higgs sector. *Phys Rev D* (2008) 77:065014. [0711.4646]. doi:10.1103/physrevd.77.065014

152. Boyanovsky D. Clustering properties of a sterile neutrino dark matter candidate. *Phys Rev D* (2008) 78:103505. [0807.0646]. doi:10.1103/physrevd.78.103505

153. Frigerio M, Yaguna CE. Sterile neutrino dark matter and low scale leptogenesis from a charged scalar. *Eur Phys J C* (2015) 75(31). [1409.0659]. doi:10.1140/epjc/s10052-014-3252-1

154. Drewes M, Kang JU. Sterile neutrino Dark Matter production from scalar decay in a thermal bath. *JHEP* (2016) 05(051). [1510.05646].

155. Merle A, Niro V, Schmidt D. New production mechanism for keV sterile neutrino dark matter by decays of frozen-in scalars. *JCAP* (2014) 03028. 1306.3996.

156. König J, Merle A, Totzauer M. keV sterile neutrino dark matter from singlet scalar decays: The most general case. *JCAP* (2016) 11:038. [1609.01289.

157. Peccei RD, Quinn HR. CPConservation in the presence of pseudoparticles. *Phys Rev Lett* (1977) 38:1440–3. doi:10.1103/physrevlett.38.1440

158. Peccei RD, Quinn HR. Constraints imposed byCPconservation in the presence of pseudoparticles. *Phys Rev D* (1977) 16:1791–7. doi:10.1103/physrevd.16.1791

159. Weinberg S. A new light boson? *Phys Rev Lett* (1978) 40:223–6. doi:10.1103/physrevlett.40.223

160. Wilczek F. Problem of StrongPandTInvariance in the presence of instantons. *Phys Rev Lett* (1978) 40:279–82. doi:10.1103/physrevlett.40.279

161. Preskill J, Wise MB, Wilczek F. Cosmology of the invisible axion. *Phys Lett B* (1983) 120:127–32. doi:10.1016/0370-2693(83)90637-8

162. Dine M, Fischler W. The not so harmless axion. *Phys Lett B* (1983) 120:137–41. doi:10.1016/0370-2693(83)90639-1

163. Kamionkowski M, March-Russell J. Planck scale physics and the Peccei-Quinn mechanism. *Phys Lett B* (1992) 282:137–41. [hep-th/9202003]. doi:10.1016/0370-2693(92)90492-m

164. Barr SM, Seckel D. Planck scale corrections to axion models. *Phys Rev D* (1992) 46:539–49. doi:10.1103/physrevd.46.539

165. Holman R, Hsu SDH, Kephart TW, Kolb EW, Watkins R, Widrow LM. Solutions to the strong CP problem in a world with gravity. *Phys Lett B* (1992) 282:132–6. [hep-ph/9203206]. doi:10.1016/0370-2693(92)90491-l

166. Ghigna S, Lusignoli M, Roncadelli M. Instability of the invisible axion. *Phys Lett B* (1992) 283:278–81. doi:10.1016/0370-2693(92)90019-z

167. Rubakov VA. Grand unification and heavy axion. *JETP Lett* (1997) 65:621–4. [hep-ph/9703409]. doi:10.1134/1.567390

168. Holdom B, Peskin ME. Raising the axion mass. *Nucl Phys B* (1982) 208:397–412. doi:10.1016/0550-3213(82)90228-0

169. Berezhiani Z, Gianfagna L, Giannotti M. Strong CP problem and mirror world: The Weinberg-Wilczek axion revisited. *Phys Lett B* (2001) 500:286–96. [hep-ph/0009290]. doi:10.1016/s0370-2693(00)01392-7

170. Hook A. Anomalous solutions to the StrongCPProblem. *Phys Rev Lett* (2015) 114:141801. [1411.3325]. doi:10.1103/physrevlett.114.141801

171. Fukuda H, Harigaya K, Ibe M, Yanagida TT. Model of visible QCD axion. *Phys Rev D* (2015) 92:015021. [1504.06084]. doi:10.1103/physrevd.92.015021

172. Gherghetta T, Nagata N, Shifman M. A visible QCD axion from an enlarged color group. *Phys Rev D* (2016) 93(115010). [1604.01127]. doi:10.1103/physrevd.93.115010

173. Dimopoulos S, Hook A, Huang J, Marques-Tavares G. A collider observable QCD axion. *JHEP* (2016) 11(052). [1606.03097].

174. Agrawal P, Howe K. Factoring the strong CP problem. *JHEP* (2018) 12(029). [1710.04213].

175. Gaillard MK, Gavela MB, Houtz R, Quilez P, Del Rey R. Color unified dynamical axion. *Eur Phys J C* (2018) 78(972). [1805.06465]. doi:10.1140/epjc/s10052-018-6396-6

176. Nelson AE, Seiberg N. R symmetry breaking versus supersymmetry breaking. *Nucl Phys B* (1994) 416:46–62. [hep-ph/9309299]. doi:10.1016/0550-3213(94)90577-0

177. Goh H-S, Ibe M. R-axion detection at LHC. *JHEP* (2009) 03:049. 0810.5773.

178. Bellazzini B, Mariotti A, Redigolo D, Sala F, Serra J. R-axion at colliders. *Phys Rev Lett* (2017) 119:141804. [1702.02152]. doi:10.1103/physrevlett.119.141804

179. Gripaios B, Pomarol A, Riva F, Serra J. Beyond the minimal composite Higgs model. *JHEP* (2009) 04:070. 0902.1483.

180. Ferretti G, Karateev D. Fermionic UV completions of Composite Higgs models. *JHEP* (2014) 03077. [1312.5330].

181. Belyaev A, Cacciapaglia G, Cai H, Flacke T, Parolini A, Serôdio H. Singlets in composite Higgs models in light of the LHC 750 GeV diphoton excess. *Phys Rev D* (2016) 015004. [1512.07242].

182. Ferretti G. Gauge theories of partial compositeness: Scenarios for run-II of the LHC. *JHEP* (2016) 06:107. 1604.06467.

183. Jeong KS, Jung TH, Shin CS. Axionic electroweak baryogenesis. *Phys Lett B* (2019) 790:326–31. [1806.02591]. doi:10.1016/j.physletb.2019.01.036

184. Yue C-X, Zhang H-Y, Wang H. Production of axion-like particles via vector boson fusion at future electron-positron colliders. *Eur Phys J C* (2022) 82(88). [2112.11604]. doi:10.1140/epjc/s10052-022-10007-7

185. Zhang H-Y, Yue C-X, Guo Y-C, Yang S. Searching for axionlike particles at future electron-positron colliders. *Phys Rev D* (2021) 104:096008. [2103.05218]. doi:10.1103/physrevd.104.096008

186. Coelho RO, Goncalves VP, Martins DE, Rangel MS. *Production of axionlike particles in PbPb collisions at the LHC, HE–LHC and FCC: A phenomenological analysis*. *Phys Lett B* (2020) 806:135512. 2002.135512. doi:10.1016/j.physletb.2020.135512

187. Bauer M, Neubert M, Thamm A. Collider probes of axion-like particles. *JHEP* (2017) 12:044. 1708.00443.

188. Bauer M, Heiles M, Neubert M, Thamm A. Axion-like particles at future colliders. *Eur Phys J C* (2019) 79:74. 1808.74. doi:10.1140/epjc/s10052-019-6587-9

189. Cacciapaglia G, Deandrea A, Iyer AM, Sridhar K. Tera-Z stage at future colliders and light composite axionlike particles. *Phys Rev D* (2022) 105:015020. [2104.11064]. doi:10.1103/physrevd.105.015020

190. Spira M, Djouadi A, Graudenz D, Zerwas PM. Higgs boson production at the LHC. *Nucl Phys B* (1995) 453:17–82. [hep-ph/9504378]. doi:10.1016/0550-3213(95)00379-7

191. Bauer M, Neubert M, Renner S, Schnubel M, Thamm A. The low-energy effective theory of axions and ALPs. *JHEP* (2021) 04:063. 2012.12272.

192. Knapen S, Thamm A. Direct discovery of new light states at the FCCee. *Eur Phys J Plus* (2021) 136:936. 2108.08949. doi:10.1140/epjp/s13360-021-01874-2

193. Schäfer R, Tillinger F, Westhoff S, Near or far detectors? Optimizing long-lived particle searches at electron-positron colliders, 2202.11714.

194. Tian M, Wang K, Wang ZS, Search for long-lived axions with far detectors at future lepton colliders, 2201.08960.

195. Frugiuele C, Fuchs E, Perez G, Schlaffer M. Relaxion and light (pseudo) scalars at the HL-LHC and lepton colliders. *JHEP* (2018) 10:151. 1807.10842.

196. de Blas J, Ciuchini M, Franco E, Mishima S, Pierini M, Reina L et al. Electroweak precision observables and Higgs-boson signal strengths in the standard model and beyond: Present and future. *JHEP* (2016) 12(135). [1608.01509].

197. de Blas J, Cepeda M, D'Hondt J, Ellis RK, Grojean C, Heinemann B et al. *Higgs Boson Studies at Future Particle Colliders*, 01139 (2020). 1905.03764.*JHEP*.

198. Cepeda M, Gori S, Outschoorn VM, Shelton J *Exotic Higgs Decays*, 2111. p. 12751.

199. Alipour-Fard S, Craig N, Jiang M, Koren S. Long live the Higgs factory: Higgs decays to long-lived particles at future lepton colliders. *Chin Phys C* (2019) 43: 053101. [1812.05588]. doi:10.1088/1674-1137/43/5/053101

200. Strassler MJ, Zurek KM. Echoes of a hidden valley at hadron colliders. *Phys Lett B* (2007) 651:374–9. [hep-ph/0604261]. doi:10.1016/j.physletb.2007.06.055

201. Strassler MJ, Zurek KM. Discovering the Higgs through highly-displaced vertices. *Phys Lett B* (2008) 661:263–7. [hep-ph/0605193]. doi:10.1016/j.physletb.2008.02.008

202. Han T, Si Z, Zurek KM, Strassler MJ. Phenomenology of hidden valleys at hadron colliders. *JHEP* (2008) 07:008. [0712.2041.

203. Chacko Z, Goh H-S, Harnik R. The Twin Higgs: Natural electroweak breaking from mirror symmetry. *Phys Rev Lett* (2006) 96:231802. [hep-ph/0506256]. doi:10.1103/physrevlett.96.231802

204. Burdman G, Chacko Z, Goh H-S, Harnik R. Folded supersymmetry and the LEP paradox. *JHEP* (2007) 02:009. [hep-ph/0609152].

205. Cai H, Cheng H-C, Terning J. A quirky little Higgs model. *JHEP* (2009) 05: 045. 0812.0843.

206. Craig N, Katz A, Strassler M, Sundrum R. Naturalness in the dark at the LHC. *JHEP* (2015) 07:105. [1501.05310].

207. Curtin D, Essig R, Gori S, Jaiswal P, Katz A, Liu T et al. Exotic decays of the 125 GeV Higgs boson. *Phys Rev D* (2014) 90:075004. [1312.4992]. doi:10.1103/physrevd.90.075004

208. Haisch U, Kamenik JF, Malinauskas A, Spira M. Collider constraints on light pseudoscalars. *JHEP* (2018) 03:178. 1802.02156.

209. Fuchs E, Matsedonskyi O, Savoray I, Schlaffer M. Collider searches for scalar singlets across lifetimes. *JHEP* (2021) 04:019. 2008.12773.

210. Cui Y, Shuve B. Probing baryogenesis with displaced vertices at the LHC. *JHEP* (2015) 02:049. 1409.6729.

211. Batell B, Pospelov M, Shuve B. Shedding light on neutrino masses with dark forces. *JHEP* (2016) 08:052. 1604.06099.

212. Accomando E, Delle Rose L, Moretti S, Olaiya E, Shepherd-Themistocleous CH. Novel SM-like Higgs decay into displaced heavy neutrino pairs in U(1)' models. *JHEP* (2017) 04:081. 1612.05977.

213. Accomando E, Delle Rose L, Moretti S, Olaiya E, Shepherd-Themistocleous CH. Extra Higgs boson and Z' as portals to signatures of heavy neutrinos at the LHC. *JHEP* (2018) 02:109. 1708.03650.

214. Curtin D, Essig R, Gori S, Shelton J. Illuminating dark photons with high-energy colliders. *JHEP* (2015) 02157. [1412.0018].

215. Stelzer T, Long WF. Automatic generation of tree level helicity amplitudes. *Comput Phys Commun* (1994) 81:357–71. [hep-ph/9401258]. doi:10.1016/0010-4655(94)90084-1

216. Alwall J, Frederix R, Frixione S, Hirschi V, Maltoni F, Mattelaer O et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP* (2014) 07(079). [1405.0301].

217. Sjöstrand T, Ask S, Christiansen JR, Corke R, Desai N, Ilten P et al. An introduction to PYTHIA 8.2. *Comput Phys Commun* (2015) 191(159)–77. [1410.3012]. doi:10.1016/j.cpc.2015.01.024

218. DELPHES 3 collaboration,de Favereau J, Delaere C, Demin P, Giammanco A, Lemaître V, Mertens A et al. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP* (2014) 02:057. 1307.6346.

219. collaboration,Antonello M. Idea: A detector concept for future leptonic colliders. *Nuovo Cim C* (2020) 43:27.

220. Volkl V, Madlener T, Helsens C, key4hep/k4simdelphes: v00-01-07, 2021. doi:10.5281/zenodo.5585287

221. Volkl V, Madlener T, Gaede F, Sailer A, Helsens C, Declara PF et al. key4hep/edm4hep: v00-04, 2021. doi:10.5281/zenodo.5585967

222. Guiraud E, Naumann A, Piparo D. *Tdataframe: Functional chains for root data analyses* (2017). doi:10.5281/zenodo.260230

223. Brun R, *root-project/root: v6.18/02*, 2019. doi:10.5281/zenodo.3895860

224. Ai X, Allaire C, Calace N, Gessinger P, Grasland H, Gray H et al. Acts project: v8.3.0, 2021. doi:10.5281/zenodo.4818137

225. Cacciari M, Salam GP, Soyez G. FastJet user manual. *Eur Phys J C* (2012) 72: 1896. [1111.6097]. doi:10.1140/epjc/s10052-012-1896-2

226. Pivarski J, Osborne I, Das P, Hollands A, Schreiner H, Biswas A et al. scikit-hep/awkward-1.0: 1.8.0rc3, 2022. doi:10.5281/zenodo.5885291

227. Alva D, Han T, Ruiz R. *Heavy Majorana neutrinos from Wγ fusion at hadron colliders. JHEP* (2015) 02:072. [1411.7305].

228. Degrande C, Mattelaer O, Ruiz R, Turner J. Fully-Automated precision predictions for heavy neutrino production mechanisms at hadron colliders. *Phys Rev D* (2016) 94:053002. [1602.06957]. doi:10.1103/physrevd.94.053002

229. Christensen ND, Duhr C. FeynRules - Feynman rules made easy. *Comput Phys Commun* (2009) 180:1614–41. [0806.4194]. doi:10.1016/j.cpc.2009.02.018

230. Degrande C, Duhr C, Fuks B, Grellscheid D, Mattelaer O, Reiter T. UFO - the universal FeynRules output. *Comput Phys Commun* (2012) 183:1201–14. [1108.2040]. doi:10.1016/j.cpc.2012.01.022

231. Alloul A, Christensen ND, Degrande C, Duhr C, Fuks B. FeynRules 2.0 -a complete toolbox for tree-level phenomenology. *Comput Phys Commun* (2014) 185: 2250–300. [1310.1921]. doi:10.1016/j.cpc.2014.04.012

232. Bacchetta N, Cld – a detector concept for the FCC-ee, (2019) 1911.12230.

233. Ruiz RE. PhD thesis, 1509. Pittsburgh U. (2015). p. 06375.*Hadron Collider Tests of Neutrino Mass-Generating Mechanisms*.

frontiers | Frontiers in Physics

TYPE Review
PUBLISHED 12 October 2022
DOI 10.3389/fphy.2022.935196

Check for updates

OPEN ACCESS

EDITED BY
Alessandro Tricoli,
Brookhaven National Laboratory (DOE), United States

REVIEWED BY
Akira Yamamoto,
High Energy Accelerator Research Organization, Japan
Bernhard Auchmann,
Paul Scherrer Institut (PSI), Switzerland

*CORRESPONDENCE
Alexander V. Zlobin,
zlobin@fnal.gov

SPECIALTY SECTION
This article was submitted to Radiation Detectors and Imaging, a section of the journal Frontiers in Physics

RECEIVED 03 May 2022
ACCEPTED 29 August 2022
PUBLISHED 12 October 2022

CITATION
Bottura L, Prestemon S, Rossi L and Zlobin AV (2022), Superconducting magnets and technologies for future colliders.
Front. Phys. 10:935196.
doi: 10.3389/fphy.2022.935196

COPYRIGHT
© 2022 Bottura, Prestemon, Rossi and Zlobin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Superconducting magnets and technologies for future colliders

Luca Bottura[1], Soren Prestemon[2], Lucio Rossi[3,4] and Alexander V. Zlobin[5]*

[1]European Organization for Nuclear Research (CERN), Geneva, Switzerland, [2]Lawrence Berkeley National Laboratory, Berkeley, CA, United States, [3]Physics Department, Università di Milano, Milan, Italy, [4]LASA Laboratory, INFN-Milan, Segrate, Italy, [5]Fermi National Accelerator Laboratory (Fermilab), Batavia, IL, United States

The implications of accelerator magnet R&D towards future colliders are reviewed and discussed. It starts with a brief overview of the present and future accelerator facilities which rely on the significant advances and innovations in key technologies. Then advances and needs for present key projects and studies are expanded on specific examples. This provides the lead to discuss the recent progress in accelerator magnet R&D and the future plans. We conclude with a summary of our view of the major development drivers and future perspectives.

KEYWORDS

accelerator magnet, collider, magnetic field, R&D, superconductor

1 Introduction

The last century of extraordinary progress in fundamental physics, seeking answers to the most fundamental questions on the nature and structure of our Universe, has relied on experiments of increasing scale and complexity. Among those experiments, particle accelerators are an example of large-scale scientific endeavors, where several areas of science and engineering are seamlessly integrated into spectacular scientific instruments. One such example is the Large Hadron Collider (LHC) [1], a world's premier facility presently in operation at the European Organization for Nuclear Research (CERN) near Geneva in Switzerland and uniting the efforts of the international High Energy Physics (HEP) community. The LHC was successful in discovering and measuring properties of the Higgs boson [2, 3], the particle completing the present view of the Standard Model (SM), concluding a chase that lasted half a century. And yet, it is well known that other physics phenomena Beyond the Standard Model (BSM) must exist to explain properties of the Universe not yet understood. Therefore, there is strong interest from the international HEP community in improving the measurement statistics and accessing energy scales well beyond the current reach of almost 14 TeV provided by the LHC. Several projects and studies are on-going in this direction. In this paper we review the implications for accelerator magnet R&D towards future colliders. We give first a brief overview of the future accelerator facilities which demand the most challenging advances in accelerator magnet technologies. On specific examples we expand advances and needs for a selected projects and studies. This provides the lead

to discuss the recent progress in accelerator magnet R&D and the future plans. We finally conclude with a summary of our view of the major development drivers and future perspectives.

2 Overview of future accelerator facilities

Several facilities have been proposed, are in discussion and study, or are being constructed to extend the energy reach and explore BSM physics. This follows a process whereby progress in physics discovery and understanding feeds the strategy for future experiments and investments in new or enhanced facilities enabling new physics reach. In Europe, the community process occurs every five to 7 years through the proceedings of a European Strategy Group, resulting in a document on the European Strategy for Particle Physics (ESPP) that provides a shared roadmap and guides European investments in this area [4]. The last ESPP update was recently completed with the publication of an Accelerator R&D Roadmap in 2022 [5]. In the United States, the Strategy for High Energy Physics, including Particle Physics, is also reviewed and revised within eight to 10 years. The process involves a year-long community discussion period known as "The Snowmass Community Planning Exercise" [6]. Snowmass is then followed by a Particle Physics Project Prioritization Panel (P5) which is tasked with the final strategic planning that guides U.S. HEP investments. The last completed Snowmass process was in 2013, culminating in a community planning report, which was then followed by a P5 report in 2014 [7]. A new Snowmass process is going to its final phase as we write.

A concrete result of the European process is the High Luminosity upgrade of the LHC (HL-LHC), described later in more detail. The HL-LHC, which aims at improving the precision of the LHC measurements by a factor 3 or more, was initiated as a project at CERN in 2015, and involves the world-wide accelerator physics and technology community.

A second example, also considered in more detail later, is the Electron-Ion Collider (EIC), to be built at BNL over the 10–15 years as a U.S.-centered effort started in 2020. The Nuclear Science Advisory Committee (NSAC) recommended the EIC as the next major nuclear physics facility for the U.S. Nuclear physics program, with luminosity up to 10^{34} $cm^{-2}s^{-1}$ and center of mass (c.o.m.) energy in the range of 20–100 GeV initially, and highly spin polarized electrons, protons, and light ion beams. The EIC will use electrons to probe hadrons, protons and neutrons, and their interactions through the gluons that bind their constituents. It will provide unprecedented capability for understanding the interaction of elemental quarks and gluons that form the basic structure of atoms and nuclei.

Looking at the future of HEP, and beyond existing installations, particle colliders have a prominent role, being considered the most promising means for the next step in both precision and energy reach. A number of candidates are known and relatively well defined. The Future Circular Collider (FCC) proposed by CERN [8] foresees an integrated plan, with a first step, the FCC-ee, consisting of a 100 km lepton (e + e-) collider at c. o.m. Energies in the range of 90–350 GeV to probe with high precision the properties of the Z, W and H bosons, as well as the top quark. The electro-weak factory machine will then make place to a hadron machine, FCC-hh, in the same tunnel repeating the successful scheme of LEP/LHC. By strong high field on the level of 16 T, indeed, the FCC-hh can reach 100 TeV c. o.m. Energy with pp collisions, probing new physics well beyond the LHC reach. The FCC-hh will also provide ion-ion collisions and the possibility of precision physics by means of p-e or ion-e collisions.

The Institute of High Energy Physics (IHEP) of the Chinese Academy of Science (CAS) has a similar proposal of lepton and hadron machines [9] consisting of two steps. The Circular Electron-Positron Collider (CEPC) would first probe the SM heavy particles with precision measurements in the range of 90–240 GeV c. o.m. Energy, to be followed by a Super Proton-Proton Collider (SPPC) installed in the same tunnel, which will collide high energy hadrons in the range of 75 TeV (CDR) to 125–150 TeV (ultimate) c. o.m. Energy. The first step in both proposals foresees circular lepton colliders for precision physics.

Linear colliders are an alternative in the range of energy targeted for precision measurements. The International Linear Collider (ILC) is a 20 km long superconducting RF accelerator proposal, aiming at colliding leptons (e + e-) in the range of 250 GeV (TDR updated) to one TeV (upgrades) c. o.m. Energy, targeting precision measurements of heavy SM particles and their coupling. CERN has proposed a Compact Linear Collider (CLIC) that would serve both for precision measurements, as well as physics at the energy Frontier. The CLIC proposal is an attempt to bring lepton colliders to the energy Frontier, which is presently dominated by circular hadron colliders. The CLIC plans to use copper RF cavities to accelerate and collide leptons (e + e-) in three steps, from a c. o.m. Energy of 380 GeV, to one TeV, and eventually to 3 TeV. The accelerator length would grow correspondingly, in the three steps, from a minimum 11 km (380 GeV) to a maximum of 50 km (3 TeV).

Lepton colliders benefit from the fact that leptons are fundamental particles, and a head-on collision makes the full particle c. o.m. Energy available for a given physics process. By contrast, hadrons are composite particles, and only a fraction of the particle energy is available in each process "channel", typically one-sixth to one-10th of the c. o.m. Energy. The downside of circular electron-positron colliders is synchrotron radiation, which severely limits the energy reach, as can be seen looking at the planned energy range of the proposed FCC-ee and CEPC. Linear electron-positron colliders avoid the synchrotron radiation problem but require the beam to be accelerated to full energy in a single passage, hence resulting in a very long

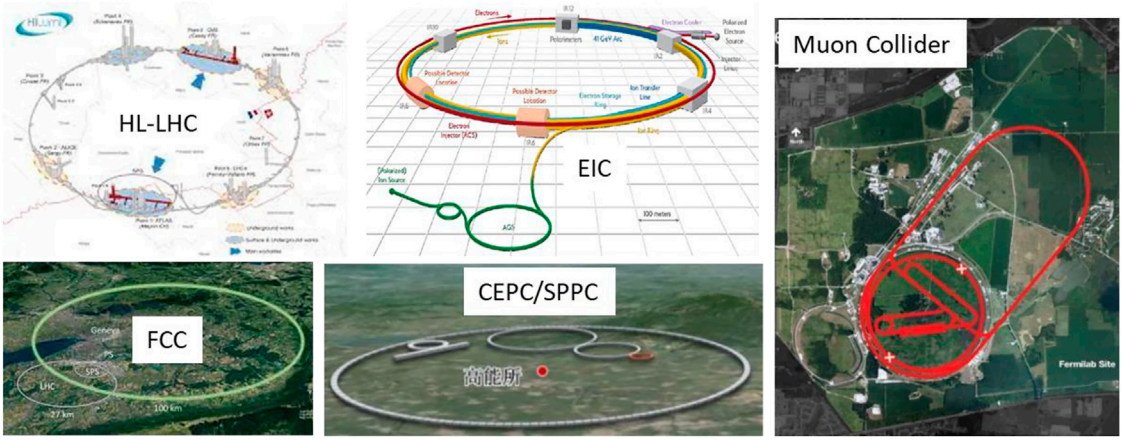


FIGURE 1
Proposed layouts of the HL-LHC (CERN), EIC (BNL), FCC (CERN), CEPC/SPPC (China) and Muon Collider (3 TeV version, FNAL).

accelerator and in a waste of energy for a given luminosity (most of the particle are lost without collisions). Scientific reasons (limited energy reach vs. circular hadron colliders, limited luminosity as Higgs and EW factory vs. circular *e-e* colliders) and energy consumption considerations are now disfavoring linear collider like CLIC, at least as energy Frontier machine. With the highest c. o.m. Energy of 3 TeV, even the final stage of the CLIC (necessitating of almost 600 MW of electric power) is not quite in reach of FCC-hh or SPPC at 100 TeV c. o.m. Energy.

A solution to this conundrum may be provided by a Muon Collider (MC). Muons are leptons, and profit from the same advantage of being point-like particles, but they mitigate the synchrotron radiation because their mass is over 200 times larger than that of electrons. This allows accelerating and colliding muon beams in rings and using technology developed for proton colliders. Since muons are not stable particles, the challenge is to do it within the lifetime of muons, i.e. 2.2 μs in the laboratory frame. Several options of MC were proposed and are being studied. The U.S. Muon Accelerator Program has produced a detailed study of a muon collider at a c. o.m. Energy of 3 TeV [10], while present activities are concentrating on higher c. o.m. Energies, from 6 TeV [11], to 10 TeV and beyond [12, 13].

The proposed layouts of the HL-LHC, EIC, FCC, CEPC/SPPC, and 3 TeV MC are shown in Figure 1 (not in scale).

Though rather cursory, the above review of ongoing collider projects and upcoming proposals gives a clear impression of a very active community, providing a strong and consistent pull. The implications are major and wide-ranging, in many fields beyond HEP, and especially on magnet technology. Accelerator magnets are the preferred means to form, shape and guide high-energy particle beams, and it is not surprising that accelerator magnet technology has gone hand in hand with the progress in particle colliders. Magnets are required in any type of accelerator, be it circular or linear. However, among the projects and proposals listed above, the challenges posed by future circular hadron and muon colliders are by far the most demanding. In return, progress in magnet technology is a vital ingredient to future colliders. In the following sections we review needs, advances and perspective for future accelerator magnets in direct connection with future colliders.

3 Superconducting magnets for HL-LHC

The High Luminosity LHC is the upgrade of the LHC aiming at the ultimate goal of increasing the collision rate by a factor of 7, in order to attain an integrated luminosity of more than 4,000 fb^{-1}, as ultimate goal. These values are an order of magnitude larger than the similar ones in the LHC [14–16]. The cornerstone of the upgrade are the insertions magnets and especially the new Inner Triplet (IT), consisting of large-aperture low-β quadrupoles, which are designed to reach transverse betatron function at the ATLAS and CMS collision points of $\beta^* = 15$ cm (the ultimate β^* limit of approximately 10 cm). The IT consists of quadrupoles, dipoles and corrector magnets. Additional high-field dipoles are needed for the Dispersion Suppressor (DS), known as 11 T dipoles. A compilation of the cross-sections of the HL-LHC magnets is shown in Figure 2 while Table 1 summarizes the main parameters of the HL-LHC Nb_3Sn magnets, the IT quadrupoles and DS (11 T) dipole, compared to the present Nb-Ti LHC inner triplet quadrupoles and the LHC main superconducting dipoles.

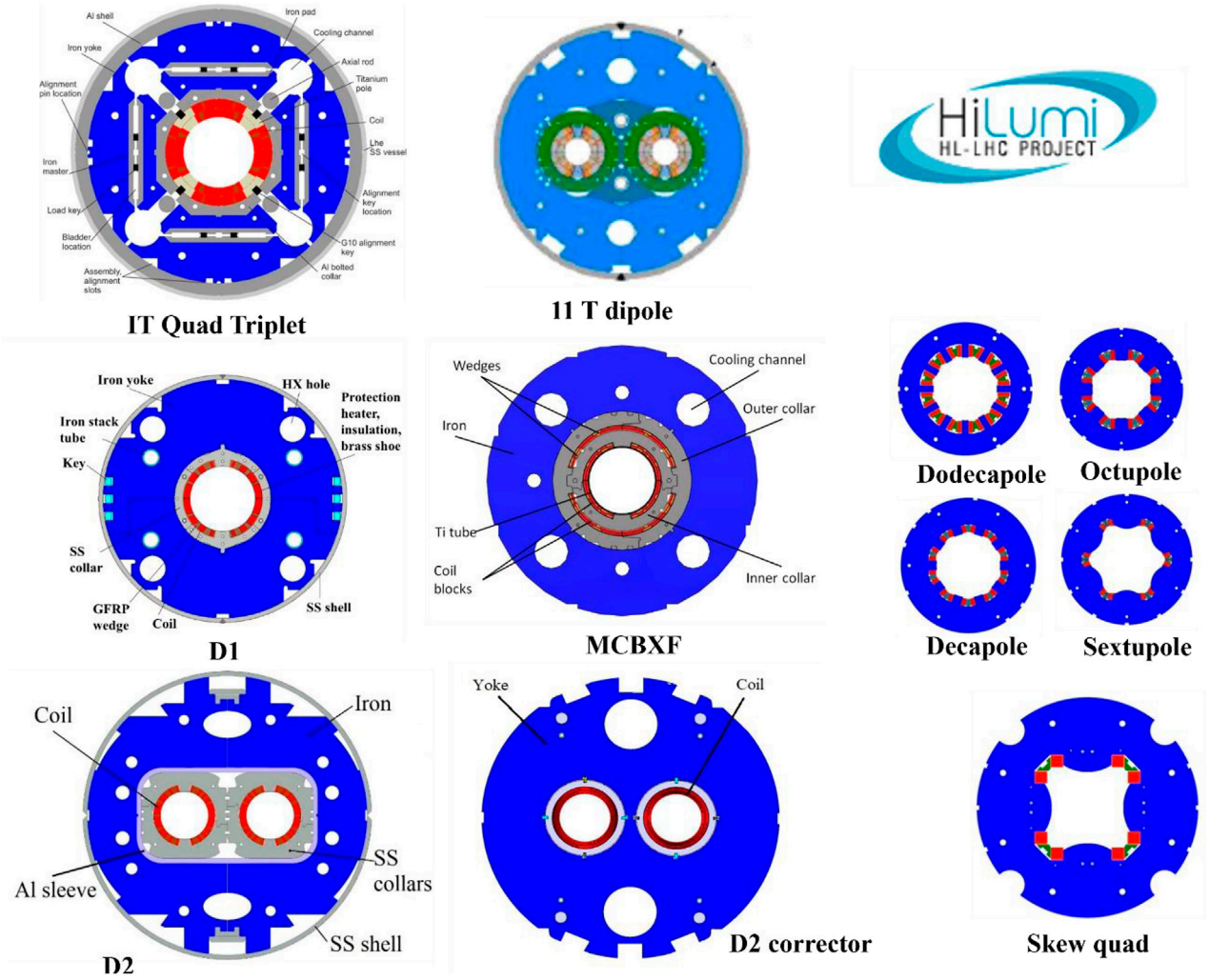


FIGURE 2
The cross section of the various magnets developed for High Luminosity LHC.

TABLE 1 Parameters of the HL-LHC IT quadrupoles and 11 T dipoles and the LHC main dipoles.

Parameter	IT quadrupole	11 T dipole	LHC dipole
B0 (T)	11.4 (B_{coil})	11.2	8.3
$L_{magnetic}$ (m) type A or type B	2 × 4.2 or 7.15	2 × 5.31	14.3
Superconductor	Nb_3Sn	Nb_3Sn	Nb-Ti
Coil aperture (mm)	150	60	56
Strand diameter (mm) × number of strands	0.85 × 40	0.70 × 40	1.065 × 28 (IL) 0.825 × 36 (OL)
Cable mid-thickness × width (mm)	1.525 × 18.37	1.25 × 14.70	1.9 × 15.1
Coil current density (A/mm^2)	462	594	368 (IL), 460 (OL)
Coil cooling	Indirect (solid conduction) at 1.9 K	Indirect (solid conduction) at 1.9 K	Direct (1.9 K HeII), porous insulation
Stored energy (MJ)	2×(4.91 ÷ 8.37)	2×5	6.87
Stored energy per unit length (MJ/m)	1.17	0.9	0.48

3.1 Nb_3Sn magnets

To reach such small β*, a set of four quadrupoles (Q1, Q2a, Q2b and Q3) is placed left and right of the LHC high luminosity Insertion Regions, IR1 (ATLAS) and IR5 (CMS). The large 150-mm aperture of the new IT quadrupoles is the main ingredient to reaching small β*. It is much larger than the one in the present 70-mm LHC IT quadrupoles. The large aperture is necessary to accommodate the beam waist, while leaving sufficient space for a thick tungsten shield that decreases the radiation dose to the coils of the IT magnets by a factor of ten. At the same time, the IT quadrupole field gradient must be as high as practical to minimize the IT length. The chosen design gradient of 132 T/m generates a peak of magnetic field of nearly 11.5 T in the coils, which is by far above the reach of Nb-Ti and requires the use of Nb_3Sn superconductor. The quadrupoles construction is a joint venture between a team of three U.S. laboratories and CERN. The U.S. laboratories (BNL, FNAL and LBNL) have collaborated for long time in the US-LARP program, especially devised for the R&D required for the LHC upgrade (2004-2018) [17]. Nowadays, those laboratories are federated under the U.S. HL-LHC Accelerator Upgrade Project (US-AUP) for magnet fabrication and test. CERN started R&D for the HL-LHC in 2010, after LHC entered operation, providing guidance as to the parameters necessary for the initial R&D phase in USA, and being the ultimate responsible of HL-LHC [18]. A second challenge of HL-LHC is the increased beam intensity, twice that of the LHC beam. This calls for additional collimators to be placed in the dispersion suppressor around the LHC Point 7 (DS7). To avoid loss-driven quench of the main dipoles around DS7. Due to the tight filling factor of LHC, it is very difficult to recover space by displacing the main dipoles (to do so one would need to remove and replace some 27 dipoles per side!) the only practical way to gain the space necessary for the collimators without a large change of the machine layout, is to remove one main dipole, which is worth of 8.33 T field over 14.3 m magnetic length, and then to replace it with an 11 T long dipole [16], to maintain the same bending strength of about 121 Tm at collision energy with a reduced magnetic length of 11 m. As for the IT quadrupoles, the new higher field value imposes to use Nb_3Sn superconductor for the coil. An 11 T dipole unit is composed of two magnets of about 5.5 m length with an approximate 4 m space to place the collimator in between the two Nb_3Sn magnets. The 11 T dipole R&D was launched at the start of the HL-LHC project, at the end of 2010 when FNAL and CERN initiated a collaboration. FNAL supported the construction of a few short model magnets and played a key role in transferring to CERN the experience in Nb_3Sn technology [19, 20]. CERN, after the construction of a few models, proceeds to the construction of full size 11 T dipoles.

Both the IT quadrupoles and 11 T dipoles were designed and manufactured in accordance with some simple decisions taken at the project outset, adapted to the small series of magnets planned (few tens). Operation at 1.9 K was chosen as a reference condition, compatible with LHC cryogenics. The operating point was taken at 80% of critical conditions along the magnet load line, where the critical surface is defined by strand measurement extracted from the Rutherford cable used in the magnet. For Nb_3Sn in the range of interest, this corresponds to a fraction of critical current of about 50%, and a temperature margin of at least 4.5 K. For the IT quadrupoles the operating point was reduced to increase margin and adjusted to 78% of the load line. The peak stress was limited to 150 MPa in nominal conditions at any location in the coil. This level of stress was found to be the upper limit for brittle Nb_3Sn superconductor, based on experimental data on wires and cables [21, 22]. The maximum allowable hot-spot temperature was set at 350 K, considering all nominal and off-nominal protection scenarios (with up to two independent failures).

The coils are based on a cos ($n\vartheta$) layout with two layers and no grading. The benefits of not grading are the use of a one type of cable, made with one kind of wire, simplifying conductor procurement and coil technology, a reduction of stress and hot-spot temperature due to a lower engineering current density, and avoiding internal joints. The coils are manufactured using the wind-and-react technique by winding an insulated cable, heat treating, and impregnating the coils with epoxy resin. This is the only practical solution to wind the brittle conductor around the tight radius at the coil ends. The turn and ground insulations are made of a S2 glass, which is braided directly on the Rutherford cable (turn), or applied as sheets (ground), acting as a spacer during winding and heat treatment. The impregnation is performed after heat treatment using the epoxy resin CTD101K, casting the fragile strands in a composite with sufficient dielectric strength.

The main characteristic of the HL-LHC strands is the critical current density, that is specified to be 2450 A/mm^2 at 12 T and 4.2 K. This value is significantly below the best performance achieved in short batches, of 3000 A/mm^2 at 12 T and 4.2 K. This was done to limit the cost (the 3000 A/mm^2 is not yet consolidated in production) and to cope with the fact that the HL-LHC IT quadrupoles and 11 T dipoles are the first large application of Nb_3Sn superconductor in accelerator magnets. A second important remark is on the filament diameter D_{eff} which is relatively large, within 40–70 μm. The magnetization associated with such values is large, and the conductor is prone to flux jumps at low fields. For the IT quadrupoles, field quality is important only at high energy, where the magnetization is forcibly small. For the 11 T dipoles, the planned number of magnets is small (2–4 units at most), and the expected perturbation to the beam due to a sextupole error of about 20 units at injection can be corrected by the regular correction system. For these reasons, a relatively large D_{eff}, a more readily scaled for production with the requisite quality control, was considered sufficient.

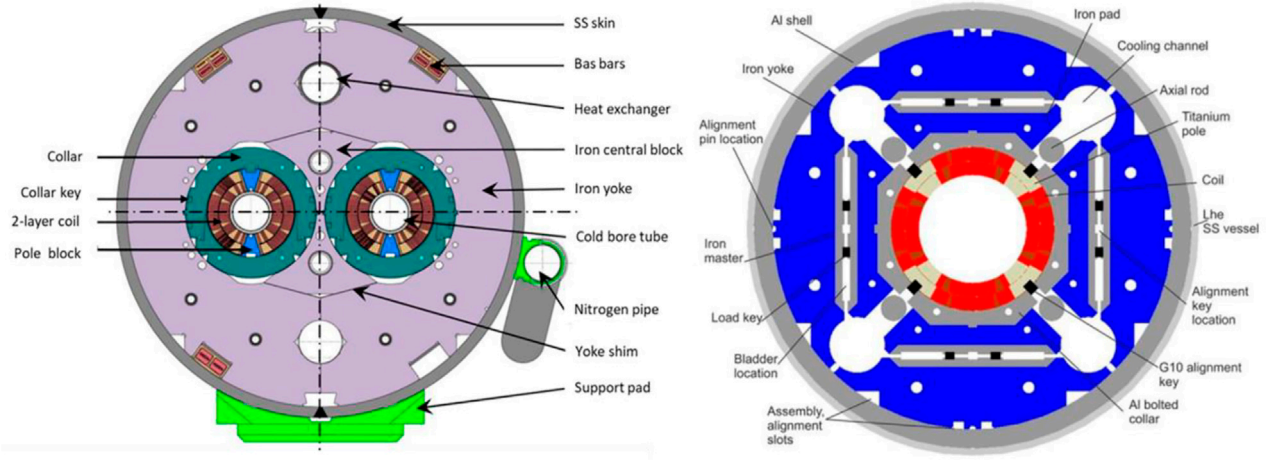


FIGURE 3
Left: cross-section of the 11 T dipole cold mass. Right: cross-section of the MQXF cold mass. The two cross-sections are not in scale.

The work on HL-LHC conductor has focused on industrial yield and production homogeneity, with piece lengths above 1 km. In total, about 25 tons of Nb_3Sn strands have been produced for the HL-LHC magnets. While highly successful, the magnet experience has confirmed that the mechanical properties and limits of the brittle Nb_3Sn wires are key parameters in the wire optimization. For a given wire, the irreversible longitudinal strain was found to depend on the heat treatment schedule. It is evident that these results, and the performance degradation experienced in some short model and long prototype magnets, point to the need to revisit the Nb_3Sn wires design and manufacturing for a next step in accelerator magnet R&D.

The coil layout of the IT quadrupoles was mainly optimized with the goal of meeting the severe field quality (FQ) requirement at collision energy when the beam size reaches its maximum. The coil has an aperture of 150 mm. It is wound with an insulated cable consisting of 40 strands, 0.85 mm in diameter, an average dimension of bare cable of 1.59 mm by 18.36 mm, and a keystone angle of 0.4°. Two magnetic lengths are employed: 4.2 m to be assembled in a doublet, making a cold mass of nearly 10 m, and 7.15 m. The overall current density at the operating temperature and nominal peak field of 11.5 T in the coil is 460 A/mm^2, while stored energy is 1.2 MJ/m. For comparison the stored energy in the LHC dipole is 0.5 MJ/m over the two magnetic channels. The nominal peak stress is 110 MPa.

The 11 T coil design is based on an insulated cable consisting of 40 strands, 0.7 mm in diameter, with average dimension of 1.25 mm by 14.7 mm, and a keystone angle of 0.8°. The coil is wound with an aperture of 60 mm and magnetic length of 5.21 m. The overall current density at operating temperature and nominal field of 11.3 T in the coil is 600 A/mm^2, while stored energy is 0.9 MJ/m. The nominal peak stress is 130 MPa.

The wound coil is placed in a steel mold and the Nb-Sn precursors are reacted by heating to 650°C in Argon atmosphere to form superconducting Nb_3Sn phase. Superconducting filaments expand by about 4–5% in volume once the Nb_3Sn compound is formed. Transverse clearance and mobile parts with gaps are used to avoid straining the coils during the cool-down from 650°C.

The reacted coil is spliced to flexible Nb-Ti leads and instrumented. In this phase the quench protection heaters, made of steel-copper thin strips encapsulated between layers of polyimide, are placed over the coils and covered with glass sheets. The coil is then impregnated under vacuum with CTD101 K epoxy resin, which can withstand about 30 MGy of radiation dose.

The 11 T dipoles use the classical collar structure inside the cold yoke surrounded by thick stainless-steel skin (Figure 3, left) to give to the coil a pre-stress sufficient to balance the effect of electromagnetic loads, as in the Tevatron, HERA and LHC magnets. The coils of the twin-aperture 11 T dipole are collared as separate single units to reduce the manufacturing risk. For the IT quadrupoles, the structure is based on the novel "bladders and keys" (B&K) principle [23], originally developed specifically to address concerns of overstressing brittle superconductors during magnet fabrication. The prestress is applied during assembly at room temperature by opening gaps in the components of the yoke structure, loading an external Al shell using the bladders, and locking the mechanical load state using keys (see Figure 3, right). The pre-stress at room temperature is further increased during cool down due to the shrinking of the external cylinder around the yoke, reaching the desired design value at the operating temperature.

Both 11T dipole and IT quadrupole have large and comparable engineering current density, though also in this respect the 11 T dipole is more pushed. Also, the values of stored energy per unit length and per coil mass, owing to the combination of high field and large aperture, are decidedly larger than in previous LHC magnets (see Table 1). For both magnets, the basic concept of protection is the same as in the LHC, relying on firing heaters positioned at the coil outer surfaces all along their length, the outer quench heaters (QH). A consequence of the high current density is that quench detection needs to be significantly faster than in the LHC dipoles (10 ms vs. 20–30 ms), and that the active quench protection needs to act within 20 ms. The QHs are subdivided into multiple circuits for redundancy. Bank of capacitors are kept charged and then discharged into the QH, when required, which is in principle sufficient for the protection. Due to the requirement of redundancy for the IT quadrupoles and their critical position, highly exposed to radiation, it has been decided to add the recently developed Coupling Loss Induced Quench (CLIQ). This system provides heating by forcing an oscillating current through the coils, *via* the AC loss mechanism in the conductor. CLIQ is then an independent protection system and can compensate the loss of one or two OQH circuits [24]. Although the HL-LHC magnet protection is much more demanding than in the LHC magnets as concerning detection and actively induced quench propagation, the solution is very reliable.

In the various projects for the LHC luminosity upgrade, form LARP to HL-LHC, about 20 short model magnets and more than 10 long magnets have been manufactured and tested. The training to the nominal field takes 5 to 10 quenches, whereas the ultimate values, set at 8% above the nominal value, are reached with 10–20 quenches, which is almost three times more than for the LHC main dipoles. However, the training memory is very good, even better than LHC magnets made of ductile Nb-Ti. Once an ultimate quench value is reached, the first quench after a thermal cycle is usually consistently not below the last reached value which makes long magnet training acceptable.

The results of the most recent 11 T dipole magnets have shown that performance retention through powering and thermal cycles is still an issue. It is likely related to changes in the stress and strain state of the coil over the entire cycle of construction and operation. Indeed, this seems to be the main issue for both the 11 T dipoles and IT quadrupoles. This is a topic where future development is necessary. Also, operation at 1.9 K only yields a few % increase in quench current, well short of the expected 10% in accordance with critical current scaling. Although the benefits He-II are clear, and, in particular, the superior heat transport as well as the operating margin in a radiation environment, this choice could be questioned, especially for impregnated magnets for future developments.

A recent review of the design and performance of the Nb_3Sn IR magnets for High Luminosity LHC can be found in [25].

3.2 Nb-Ti magnets

Besides the IT quadrupoles, the HL-LHC upgrade involves a significant change in the configuration of the whole magnetic system at the IR. A description of all these magnets and their performance can be found in [25].

The six resistive dipoles that separate the beams at the IT, just in front of the IT quadrupoles, with an aperture of 70 mm and a total length of 18 m are replaced by a new single Nb-Ti D1 magnet, 7 m long, 150 mm aperture, and generating a field of 6 T. The D1 magnets are being produced under the responsibility of KEK (Japan). The existing recombination dipole D2, with 90 mm aperture, is replaced by a two-in-one Nb-Ti dipole with two apertures of 105 mm. As the field in the two apertures in D2 has the same direction, causing strong iron saturation, this calls for an innovative design with a small left-right asymmetry in the coils to correct the field errors, similar to what was done in the single aperture J-PARC superconducting beam line for neutrino generation designed by KEK. The HL-LHC D2 magnets are produced under the responsibility of INFN (Italy).

A sizeable number of Nb-Ti corrector magnets of large aperture are also being designed and manufactured for HL-LHC, all with characteristics that will likely be used in future colliders. A very large, single aperture magnet combines vertical and horizontal dipole field corrector; it is built as two nested dipole coils of 150- and 200-mm inner diameter, each generating up to 4.5 Tm integral bending strength. These correctors are being produced under the responsibility of CIEMAT (Spain). The development of high order (HO) corrector magnets based on a Super Ferric (SF) design with 150 mm aperture, suitable for the generation of modest multipole field and compatible with a high radiation environment, is also under way. The HO correctors are being produced under the responsibility of INFN (Italy). A Canted Cos-Theta (CCT) dipole corrector magnet has been developed to be assembled in the cold mass of the D2 separation dipole, with an aperture of 105 mm, with a rated integral bending strength of 5 T·m. The CCT correctors are produced under the responsibility of IHEP (China).

4 Superconducting magnets for Electron-Ion Collider

Brookhaven National Laboratory (BNL) in the U.S. has started construction of the Electron-Ion Collider (EIC) with the c. o.m. Nominal energies within ~20–100 GeV with a future upgrade to ~140 GeV, and a luminosity of up to 10^{34} cm^{-2} s^{-1} for electron-proton collisions [26]. The decision to proceed with the project was taken in January 2020 with the goal to prepare it for commissioning and operation by the end of 2030. The EIC will advance the present particle colliders and enable the U.S. and international nuclear physics community to

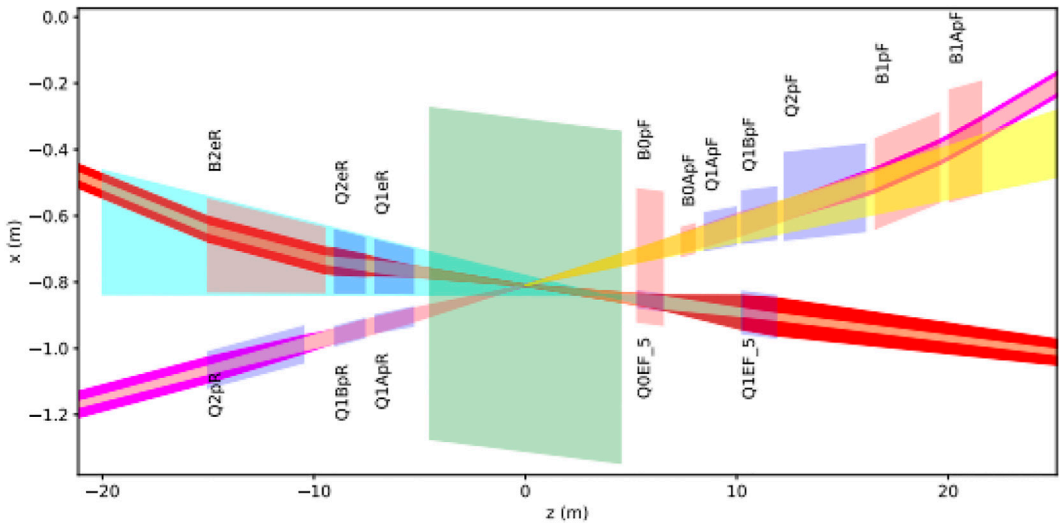


FIGURE 4
IR1 layout. The central detector is shown in light green, the hadron and electron beams are shown in magenta and light red respectively. Magnet apertures are shown in light red for dipoles or light blue for quadrupoles. The forward neutron cone is shown in yellow.

progress towards fundamental understanding of the nature of matter, providing a more clear picture of how quarks and gluons interact to form the basic structure of atoms and nuclei. The EIC design concept substantially leverages existing infrastructure from the RHIC accelerator complex at BNL. Nevertheless, significant new infrastructure has to be added to provide polarized electrons and strong hadron cooling.

All the magnets in the EIC electron storage ring (ESR) are water-cooled iron-dominated electromagnets. RHIC and its injector complex remain, for the most part, unchanged. The arcs with the Nb-Ti superconducting magnets are the same.

The EIC interaction region includes superconducting low-beta quadrupoles and spectrometer dipoles near the detector, and both normal-conducting and superconducting magnets for the matching section into the arcs of the electron and hadron storage rings, respectively. The IR layout is shown in Figure 4. It requires fifteen new unique superconducting magnets. Cross-sections of the key IR superconducting magnets are presented in Figure 5 and their parameters are summarized in Table 2. All the magnets are based on traditional Nb-Ti technology. Most of them operate at a temperature of 4.2 K, except for three magnets (Q1APF, Q1BPF and Q2PF) which due to the relatively high magnetic field need to be cooled to 1.9 K. Nine of the new superconducting magnets will be made with direct wind method and equipment previously developed at BNL [27]. This technology provides excellent field quality for a variety of magnets. Four magnets (B0aPF, Q1APF, Q1BPF, Q2PF and B1PF) are based on the collared cos-theta superconducting coils made of a Rutherford cable as the RHIC magnets.

Some of the IR magnets are twin-aperture magnets with separate coils for the hadron and electron beams inside a common iron yoke. A main challenge for these magnets is the small crossing angle, which leads to a small separation of the magnet apertures. It is vital for this design to diminish magnetic crosstalk, which is achieved by using soft iron between the two apertures and avoiding its saturation. It is helpful in this context to taper some of these magnets, thus increasing the amount of iron between the apertures to minimize its saturation.

In tapered magnets without special provisions field components change along the magnet length. Since the gradient change is not desirable, a novel winding concept has been developed to address this problem [28]. The winding pattern which allows the coil radius to change while preserving the gradient is based on a coil concept known since the 1960s [29] as double-helix or canted cosine-theta coils. A unique feature of double-helix coils is in their flexibility that allows smoothly varying field components along the coil axis as needed. Tapered coils are foreseen for Q1ER, Q2ER, B2ER and Q1APR magnets. Practice tapered coils have been made using the direct wind method and tested to prove the concept.

Discussions in the nuclear physics community show that there is a significant interest in a complementary second IR. The complementarity is provided through the energy range where the second IR is optimized to have the highest luminosity and through the detection capabilities that are alternative and complementary to the first IR. Consistently with user needs and the complementarity approach, the second IR is optimized to provide the highest luminosity at a lower CM

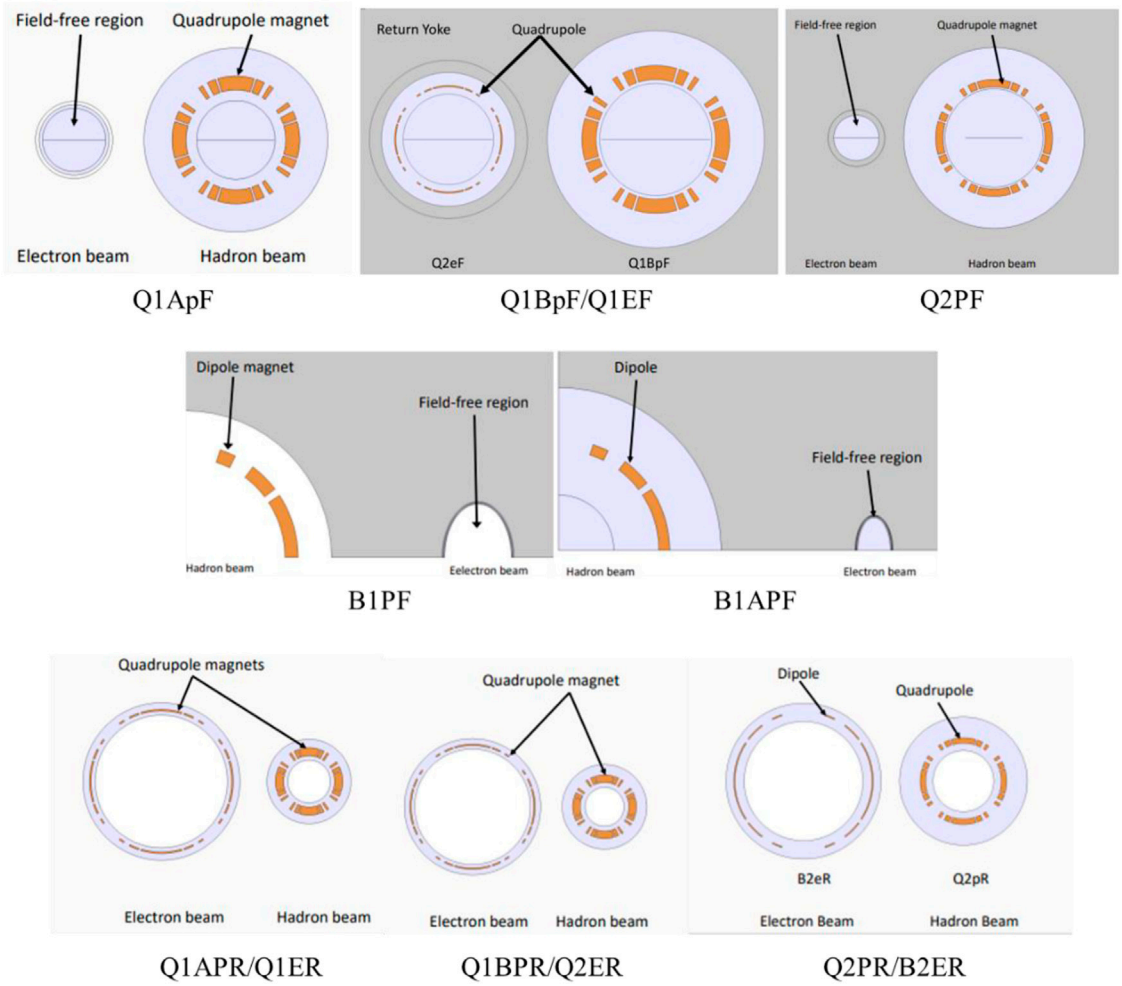


FIGURE 5
Cross-sections of IR1 superconducting magnets.

TABLE 2 IR1 Nb-Ti magnet parameters.

	Q1ApF	Q1BpF/Q1EF	Q2PF	Q1APR/Q1ER	Q1BPR/Q2ER	B1PF	B1APF	Q2PR/B2ER
Aperture, mm	112	156/128	262	40-51/130-160*	56	262	336	108/194-278*
Max field or gradient, T or T/m	72.6	66.2/8	40.7	78.38/14	78.38	3.4	2.7	34/0.2
Magnetic length, m	1.46	1.61	3.8	1.8	1.4	3	1.5	4.5
Nominal current, kA	20	20/1	15	0.236/0.073	0.265	-	7.67	0.215/0.075
Load line margin, %	32	36	32	142/32	21	58	60	50/104

Front-rear.

energy than the first IR. This enables leveling of the EIC luminosity curve over a wider energy range. The second IR can also provide a different acceptance coverage than the first IR.

Technical feasibility of all the second IR magnets is being studied. All the second IR magnets meet the nominal requirement of their fields being less than ~5 T at the maximum beam energy. However, some of the quadrupole

and dipole magnets may still present engineering challenges due to their large apertures. There are also certain advantages of increasing the magnetic field of some IR2 quadrupoles and dipoles to the level of 8–10 T which would require using brittle Nb_3Sn superconductor. The stress and strain sensitivity of this superconductor in combination with high magnetic field and large coil aperture will require using stress management techniques [30, 31]. These approaches are being studied by the US-MDP (see section 6.2.1 below).

5 SC magnets for future hadron colliders

Two major hadron machines are under consideration–the Future Circular Collider (FCC) in Europe and the Super Proton-Proton Collider (SPPC) in China. Brief descriptions of machine and magnet parameters are presented below.

5.1 Future circular collider

The main goal of the FCC hadron collider, FCC-hh, is to advance the energy and intensity limits of particle colliders to search for new physics, aiming at proton collision c. o.m. Energies of 100 TeV, which is about seven times of those at the LHC. The hadron collider would be placed in a tunnel of approximately 100 km length that would initially host the lepton collider FCC-ee, the first step in the integrated FCC plan. Although the project is centered around CERN, an international collaboration of more than 150 universities, research institutions and industrial organizations from all over the world have joined in the development of this circular collider.

Superconducting magnets are among the grand challenges of the FCC-hh. Given the tunnel dimensions and the target beam energy, the arc dipoles, approximately 15 m long, need to generate 16 T in an aperture of 50 mm [32]. The present lattice design is based on the extrapolation of the LHC and HL-LHC lattices, and uses arc quadrupoles of 7.2 m length generating 360 T/m in the same bore as the dipoles [33], with peak field in the coil of about 10.5 T. The Interaction Region (IR) also calls for quadrupoles of high field gradient and large aperture, capable of withstanding the energy deposition and dose generated by the collision debris. The range of gradients in the IR quadrupoles is from 106 to 130 T/m, with respective apertures in the range of 210 to 164 mm, corresponding to a peak field in the coil of about 11 T.

The values of peak field have been the motivation for choosing Nb_3Sn as the baseline superconductor for the majority of the FCC-hh magnets, as for the HL-LHC IR quadrupoles and 11 T dipoles. Still, this is a significant extrapolation of magnet technology beyond the work presently on-going for the HL-LHC, and especially for the dipoles. The peak field in the quadrupoles is close to the values of the HL-LHC IR quadrupoles, but larger apertures imply larger forces and stresses, thus exacerbating the issues of coil mechanical support, not to mention the increasing difficulty of magnet protection. For the dipoles, on the other hand, achieving a 16 T bore field at operation implies significant development of conductor [34], as well as magnet design and technology [35].

The most relevant advancement within the FCC magnet R&D program was brought by the exploratory effort to identify the magnet design with best properties in terms of performance, efficiency of material usage and cost. Four coil design layouts were examined in detail: cos-theta at INFN, block-type at CEA, common-coil at CIEMAT, canted cos-theta at PSI (Figure 6). A key aspect of this work is that all designs were produced using the same conductor parameters such as the critical current density J_c of 1500 A/mm^2 at 4.2 K and 16 T consistent with the conductor R&D target, the copper matrix RRR of 100, the maximum strand diameter of 1.2 mm, and the minimum Cu/nonCu ratio of 0.8:1. A critical assumption was to take the operation margin of 14% along the load-line which is below the value assumed for HL-LHC. A nominal field in aperture of 16 T corresponds to a short sample bore field of 18.6 T. For the quench protection parameters, the assumptions were an allowable time window for detection and quench initiation (i.e. the time it takes for the quench heater or CLIQ to induce a widespread transition) of 40 ms, maximum hot spot temperature of 350 K, and maximum voltage to ground of 2.5 kV. Finally, the four designs were developed using identical coil mechanical properties and, in general, the parameters of the structural materials used in the magnet.

The main results of this study are that compact configurations, based on cos-theta and block coils, are very close to each other in terms of operating margin and material cost. This is because the ratio of coil width to diameter is relatively large. Their stress distribution is very different, though. While cos-theta coils have peak stress at high field locations, at the midplane, in block coils the peak stress is generally located in moderate field region, on the outer boundary, which may be an advantage given the strain sensitivity of Nb_3Sn. Common coils appear to offer benefits of simpler end geometry, and a definite premium if a react-and-wind manufacturing procedure can be proven. The price is an increase of material, and corresponding cost, by about 20%. The canted-cos-theta is in a similar situation, offering stress management (not yet fully proved) at the price of increased material cost, in this case of about 25%.

The second major advance of the magnet R&D driven by FCC-hh is the construction of small-scale demonstration magnets to explore performance boundaries. Two configurations were produced and tested within the scope of the FCC magnet R&D program called the extended Racetrack Model Coil (eRMC), and Racetrack Model Magnet (RMM).

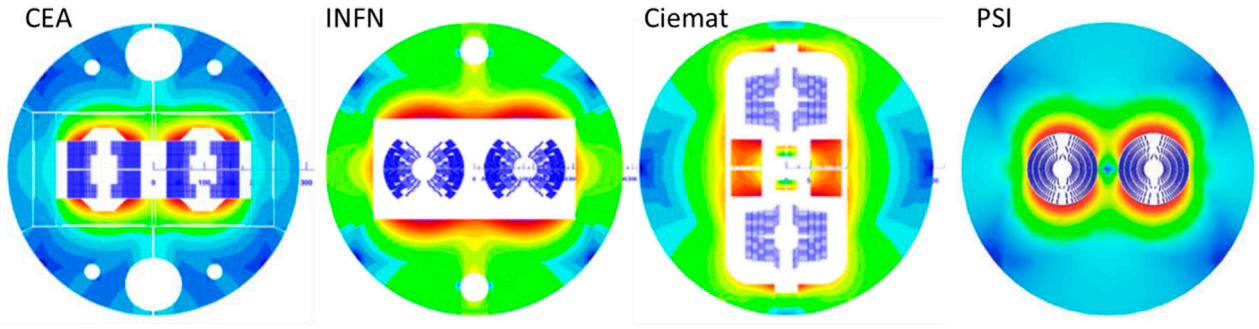


FIGURE 6
FCC dipole conceptual designs (EuroCircol).

These magnet models are assembled out of racetrack coils wound with cables of size representative of the FCC-hh magnet design, although at reduced performance. eRMC and RMM have shown that in this simplified configuration a Nb_3Sn coil can achieve operating fields in the range of 16 T, as discussed later.

5.2 Super proton-proton collider

To achieve the 75–150 TeV c. o.m. Energy, SPPC needs thousands of accelerator magnets with nominal magnetic fields within 12–24 T and apertures of 40–50 mm to guide and focus the high-energy proton beams. These magnets have to provide the field uniformity on the level of 10^{-4} in ~65% of their aperture. As the LHC magnets, these magnets have two separate apertures with opposite field direction inside the common iron yoke to minimize the magnet transverse size and reduce its cost. The aperture separation in the main dipoles is presently estimated on the level of 200–300 mm. In the final magnet design this parameter will be optimized to achieve the acceptable crosstalk between two apertures and minimize the overall magnet cross-section. The outer diameter of the arc dipole and quadrupole cold masses is limited by 900 mm to be installed inside vacuum vessels with an outer diameter of 1.5 m. The total magnetic length of the main dipoles is around 65.4 km out of the total collider ring circumference of 100 km. For the dipole length of about 15 m, approximately 4,360 dipole magnets will be needed [36, 37].

SPPC magnets with magnetic fields up to 15–16 T consider using advanced Nb_3Sn superconductors. To provide nominal operation fields up to 24 T High Temperature Superconducting materials are needed. These materials should have acceptable cost and be capable of operating in high fields at large mechanical stresses. Special attention at the present time is being paid to Iron Based Superconductors (IBS) discovered in 2008 and promising high J_c at high fields.

Conceptual design studies of twin-aperture 12 T dipole magnets for the SPPC based on the IBS technology are being performed to achieve the SPPC target parameters. The design studies are based on the high level of current density, about 10 times higher than its present level, assuming that it will be achieved within the next 10 years. In addition to the significant increase of the current carrying capability, it is also anticipated that the IBS will have better mechanical properties than the present high field superconductors such as Nb_3Sn, REBCO or Bi2212, and substantially lower cost.

Cross-sections of the two design options under consideration for the 12 T SPPC dipole based on the IBS conductor are shown in Figure 7. The coil aperture in both cases is 45 mm and the nominal field in the two magnet apertures is 12 T with the relative geometrical field errors smaller than 10^{-4}. The coil layout uses the common-coil configuration due to its simple structure compatible with the tape-type conductor and simplicity for fabrication. Two types of coil ends based on soft-way bending and hard-way bending are being studied and compared with respect to the field quality and structure design and parameters. The magnet design details and parameters are reported in [38].

5.3 High energy muon collider

Muon colliders (MC) have emerged in the past few years as potential game changers, with the promise of unique opportunities for particle physics applications. They can offer various options and possibly a staged approach towards the next step in HEP. The main difficulty of muon colliders is the short muon lifetime of 2.2 μs, which, however, increases thanks to relativistic dilation when muons are accelerated to high energies. The decay of a muon beam within a storage ring can yield pure, intense, and well controlled beams of neutrinos for neutrino-oscillation studies on short- and long-baseline. The muon mass is approximately 207 times larger than that of electrons and

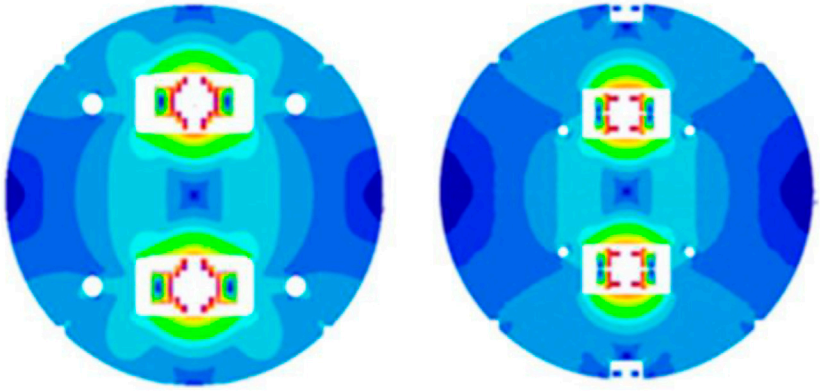


FIGURE 7
Cross-sections of the two design options for the 12 T SPPC dipole based on IBS.

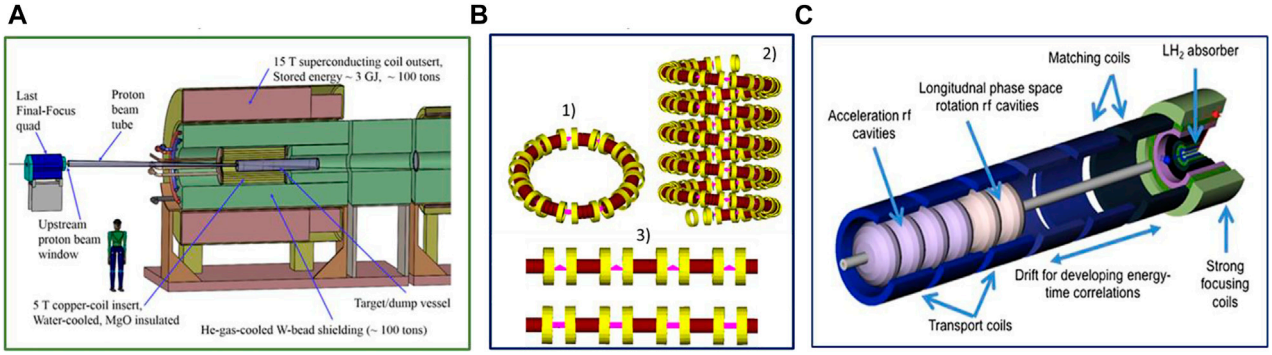


FIGURE 8
MC solenoid concepts: **(A)** muon production solenoid; **(B)** various possible solenoid configurations in 6D muon cooling system; **(C)** final cooling system with high-field solenoids.

positrons, hence, the beams are not subject to the stringent synchrotron radiation and bremsstrahlung limits of *e-p* colliders. Therefore, in principle muon beams can be accelerated to high energies, in circular colliders and stored in rings for a time sufficiently long to produce high luminosity collisions. Indeed, TeV-class muon colliders are considered as the most effective path for a high-luminosity lepton collider.

The concept of the muon collider was first proposed by G.I. Budker in 1969 [39]. Since then, innovative MC concepts have been developed in the framework of several design studies and R&D programs in the U.S., such as the Neutrino Factory and Muon Collider Collaboration (NFMCC) [40] followed by the Muon Accelerator Program (MAP) [41], and now the work will continue by the recently-created International Muon Collider Collaboration (IMCC) [42] to form the foundation for the next steps of High Energy Physics.

The muon collider requires several types of superconducting magnets with various geometries, apertures, field level and configuration, powering mode, *etc.* To provide the required field parameters in each area. Concepts for MC solenoids, devised for muon production and cooling, and dipole and quadrupoles concepts, developed for muon acceleration, storage, and final focus, are shown in Figures 8, 9.

The MC target area and the front end consist of a series of large-aperture high-field solenoids [43]. The MAP study has devised a target area consisting of a 20 T solenoid which is composed of an assembly of a 15 T superconducting outsert with 2 m aperture and a 5 T normal conducting insert with 0.3 m aperture. This high-field solenoid captures the pions produced by the impact of an intense proton beam on a target (Figure 8A). It followed by the 12 m long decay channel made of solenoids with tapered apertures from 2 m to ~0.6 m and

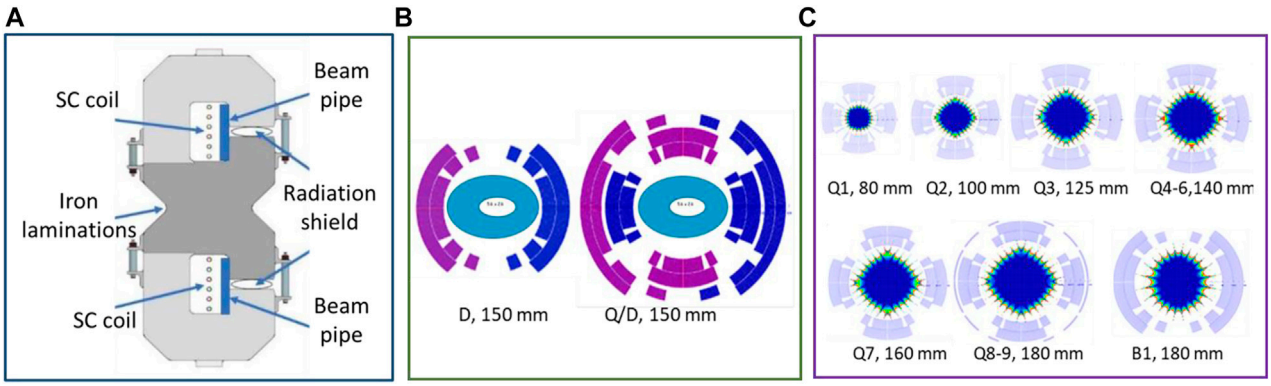


FIGURE 9
MC acceleration and SR magnets: **(A)** fast-cycling super-ferric dipole magnet with HTS coil, **(B)** collider dipole and combined function dipole-quadrupole magnets with thick elliptical absorbers in the apertures, **(C)** IR quadrupole and dipole magnets (good field quality areas in the apertures are shown in blue color).

field from 20 T to ~2 T. Based on the field level, the outer superconducting part of capture solenoid and the decay channel solenoids have been designed using Nb_3Sn and Nb-Ti coils.

Then the muon beam enters the beam cooling system which consists of six-dimensional (6D) cooling channel and the final cooling stage. In this system the 6D phase space volume of muon beams is reduced before further acceleration and injection into the storage ring. Several designs of the 6D cooling channel have been proposed and studies [44] (Figure 8B). In the 6D beam cooling area, large-aperture 10–20 T solenoids with normal RF cavities inside or in-between are considered. The field level in the 6D cooling solenoids can be achieved using contemporary Nb-Ti or Nb_3Sn superconductors. The final cooling stage (Figure 8C) consists of a series of 50-mm aperture solenoids with high magnetic field ideally in the range of 50–60 T [45]. These final cooling solenoids could be devised based on experience with the record field 32 T hybrid solenoid at NHMFL and high-field solenoids in Europe, which are using hybrid coils with HTS inner and LTS outer sections.

The accelerator concept is based on a sequence of linacs, recirculating linacs and rapid cycling synchrotrons (RCS). The RCS are designed as hybrid system, consisting in interleaved resistive fast-cycling magnets [46], and high-field steady-state superconducting magnets. The advantage of this configuration is that it reduces the length of the RCS and, thus, the time lost accelerating the muon beams. In the first RCS, the fast-ramping dipole magnets for the muon acceleration has to provide a field variation within ±2 T during 0.4 ms in a rectangular aperture of 80 mm by 40 mm, which corresponds to a field ramp rate of 10 kT/s. These magnets will use traditional fast ramping magnet technology based on resistive coils and warm iron yoke. The last RCS requires a similar swing of ±2 T in about 10 ms, or a field ramp rate of 400 T/s. Higher field swing would be welcome. These magnets can use superconducting coils (Figure 9A).

The last major part of the MC complex is the collider Storage Ring (SR) and Interaction Region (IR). A specific feature of a muon collider is the neutrino flux resulting from muon decay, taking place continuously along the storage ring. One of the measures to reduce the resulting radiation flux is to use combined function magnets in the arc, thus avoiding straight sections as would be the case if separate arc quadrupoles were used. The MAP baseline 3 TeV collider is based on 150-mm aperture dipoles and combined function magnets with a nominal dipole field of 10.5 T and a field gradient of 85 T/m in the arc [47]. The IR uses focusing quadrupoles with field gradients up to 250 T/m and apertures from 80 to 180 mm, and 180-mm aperture 8 T dipoles [48]. Magnet protection from radiation in the collider ring arc is achieved by using thick Tungsten absorbers in magnet apertures and by masks in between magnets. The cross-sections of collider ring dipoles and combined quadrupole/dipole magnets with inner absorbers, and IR quadrupoles and dipoles designed within the scope of the MAP study are shown in Figures 9B,C. All the magnets are based on Nb_3Sn superconductor operating at 4.5 K which provide sufficient temperature margin for the radiation heat load.

A 6 TeV IR design, and on-going studies for a MC at 10 TeV, assume HTS technology to achieve nominal operation fields up to 20 T, and possibly a further increase of magnet aperture to accommodate for even larger shielding. Hybrid HTS/LTS coils with larger operating fields and margins as well as curved magnets are also being considered.

6 Superconducting accelerator magnet R&D

6.1 Achievements

Having reached the limit of Nb-Ti magnet technology, all the above projects and studies have been turning towards advanced superconducting materials and novel magnet technologies, targeting increasing fields. The main motivations for a long-term program are the following:

- Development of high-field magnets is an activity with long lead times. Time scales up to a decade are required to master new technology and bring novel ideas into application. This stresses the importance that R&D runs in parallel with the study of new accelerator concepts and options, so that specific magnet technology is available for a HEP realization, anticipating the demand for the moment when the decision of construction is taken.
- SC magnet technology for future HEP colliders, in particular high field magnets, requires large scale infrastructure and considerable investment. Best use of such infrastructure is made if an R&D program is run as a continuous progression.
- Similar to the research infrastructure, the development of high field magnets requires stable proficient teams with broad competencies, spanning over many fields of science, often assembled in collaborations ranging from academia to industry. Building one such research team needs considerable investment, and continuity is an asset.

The present state-of-the-art in high-field magnets for accelerators is the result of a coordinated efforts that has spanned the past 30 years. The largest effort was put in the development of Nb_3Sn conductor and magnet technology. A strong focus was provided by the US-DOE programs devoted to Nb_3Sn conductor and magnet development, initiated at the end of 1990's. The resulting program, a broad collaboration among the US-DOE accelerator Laboratories and associated Institutions, is now continuing under the auspices of the U.S. Magnet Development Program [49, 50]. European-wide activities in high-field accelerator magnets took form under the EU-FP6 initiative for Coordinated Accelerator Research in Europe (CARE) [51]. The Next European Dipole Joint Research Activity (NED-JRA) [52], which enfolded from 2004 to 2009, was followed by the EU-FP7 EuCARD [53]. The above activities were instrumental to the development of the HL-LHC Nb_3Sn magnet technology, presently used in the QXF quadrupoles and 11 T dipoles described earlier, and demonstrating the importance of a continuous R&D.

In parallel, the interest in the exceptional high-field potential of High-Temperature Superconductors (HTS) has spurred much excitement for accelerator magnets. Although copper oxide compounds containing rare-earths (REBCO) and bismuth (BSCCO) are only in a stage of early technical maturity, laboratories and industry have shown that HTS are capable to produce fields in excess of 28 T in commercial NMR solenoids to record values of 45.5 T in small experimental solenoids. This is an area where we expect to see fast progress. Activities are in various EU laboratories are following the program initiated with the EuCARD [53] and EuCARD2 [54] collaborations, now fostered by the on-going I-FAST [55] EU project. In the US, HTS accelerator magnet development is an integral part of the US-MDP.

6.1.1 Nb_3Sn magnets

Progress in the development of Nb_3Sn magnets for accelerators can be appreciated by the suggestive presentation in Figure 10, where the steady increase of field produced by dipole magnets over the past 4 decades is shown. The plot contains results from short demonstrator magnets, built with racetrack coils with the main purpose to test field reach, short model magnets with most final features but reduced length in the range of one to 2 m, as well as full-size accelerator magnets containing all final features for operation in an accelerator.

The first significant results can be traced back to the 1980's, at BNL and LBNL [56]. This work eventually led to the achievements of D20 in the 1990's, and the 16 T field attained with the demonstrator HD1 at LBNL in the 2000's. Fields in the same range were reached at CERN in 2015 and exceeded as of 2020 as a result of the push provided by FCC-hh. As we mentioned earlier, work in the 1990's and 2000's has laid the foundations for the HL-LHC Nb_3Sn magnets. And yet, the R&D program itself was largely funded by HEP in the US, as well as initiatives in Europe, i.e., essentially independent of a specific HEP project.

Figure 10 demonstrates that the timeline for progress in Nb_3Sn magnet technology is slow. It took about a decade for CERN and collaborators to reproduce the results obtained in the U.S., going from conductor R&D to the 16.2 T dipole field obtained with RMC03 in September 2015. This gives a good benchmark for the time scale necessary to enter into the field, including the required infrastructure as well as design and manufacturing skills. The final result of this work is the record magnet FRESCA2, built in collaboration between CERN and CEA, and generating a field of 14.6 T in an aperture of 100 mm diameter [57]. This field level has been achieved also at FNAL by a high-field model dipole built within the scope of the US-MDP program [58] as a step towards the highest field that can be attained with a cos-theta coil configuration. Figure 11 shows the 100-mm aperture block-type FRESCA2 and 60-mm aperture cos-theta MDPCT1 14.6 T record dipole magnets.

The progression is not yet complete, as demonstrated by the last point in the plot, corresponding to the 16.4 T reached at CERN in 2021 in a special Racetrack Model Magnet or RMM.

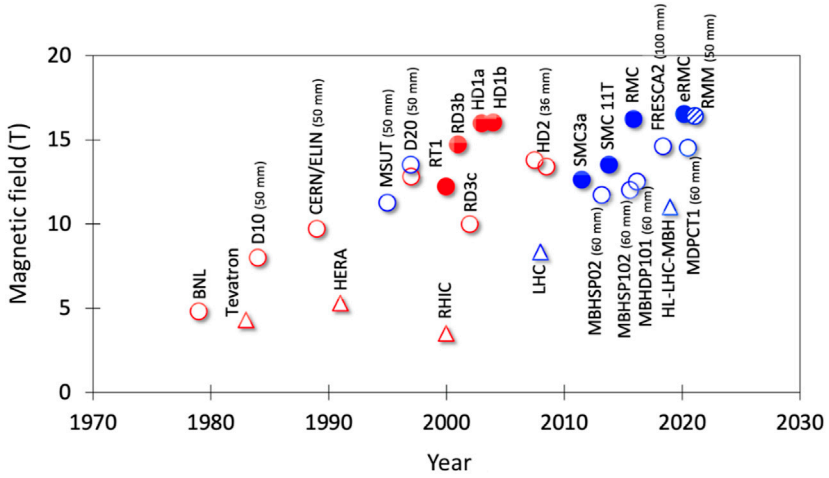


FIGURE 10
Progression of fields reached by Nb_3Sn dipole magnets of various configurations and dimensions. Open symbols represent magnets with a free bore, solid symbols are for demonstrators built with racetracks and no free bore. The hatched symbol (RMM) is a magnet with a free cavity. Tests were performed in liquid helium (4.2 K, red symbols) or superfluid helium (1.9 K, blue symbols). Superconducting collider dipole magnets are shown as triangles for comparison.

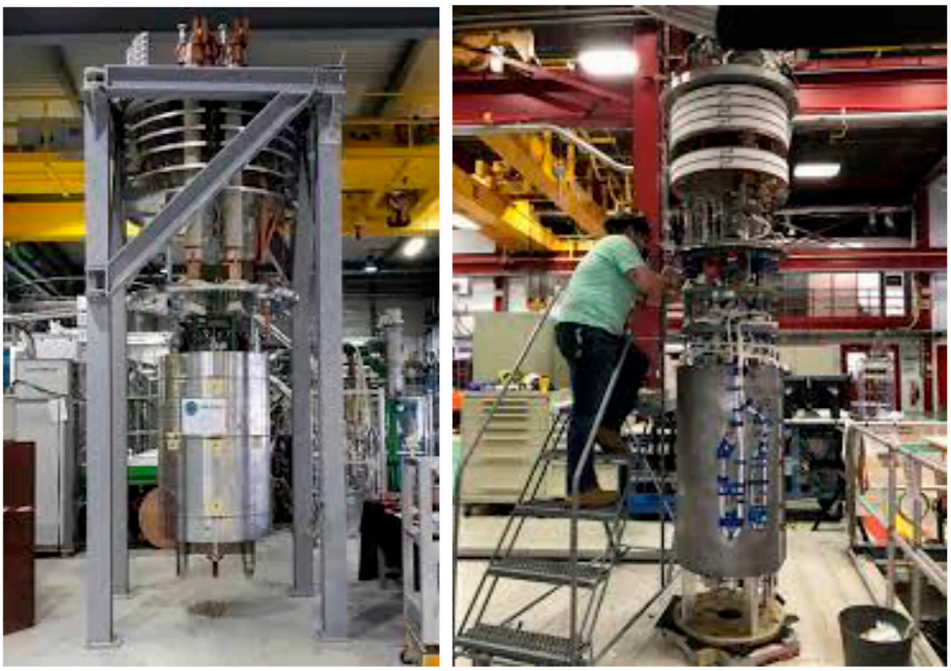


FIGURE 11
The 14.6 T record dipole magnets: 100-mm aperture block-type FRESCA2 (left) and 60-mm aperture cos-theta MDPCT1.

RMM reproduces the 2D cross-section of a block-coil magnet but has flat ends. It results in a cavity in the center of the magnet, which is the case of the RMM is 50 mm, the nominal aperture for FCC-hh.

Figure 10 also shows the remarkable achievement of the Nb_3Sn MBH 11 T dipole for HL-LHC. This dipole was built at CERN in collaboration with industry (GE-Alstom) [59]. Also here, we remark that the time scale of the project, from inception

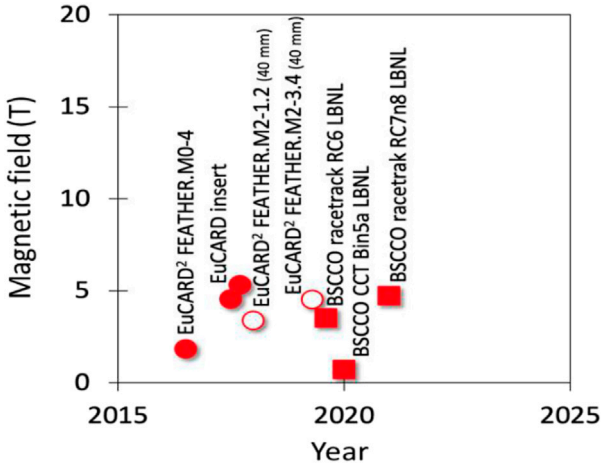


FIGURE 12
Maximum fields obtained with demonstrator HTS short magnets, producing a transverse field of dipole type. All reported maximum value were reached in liquid helium (4.2 K). Solid symbols are coils with no bore (e.g., racetracks), while open symbols are real magnets, i.e., with a usable bore. Round symbols refer to coils built with REBCO, while square symbols refer to BSCCO-2212 coils.

in 2010 to the test of the first accelerator unit in 2019, is a decade. The first such magnet, MBHB002, was tested in July 2019 and also detains the performance record for its class. As we mentioned earlier, the 11 T dipole program has pointed out that there are still questions to be resolved beyond magnet performance, on long-term reliability and robustness. These questions are being addressed so that Nb_3Sn can be used in an operating accelerator.

6.1.2 HTS magnets

HTS accelerator magnets are expected to be the next step in SC accelerator magnet technology. As reported in the graph of Figure 12, the HTS technology is making it first trials for accelerator-class magnets.

The general interest in the potential of this class of material with spectacular performance coagulated at about the same time in the mid of the 2000's in the U.S. and Europe. On the U.S. side, efforts were coordinated by the US-DOE sponsored Very High Field Superconducting Magnet Collaboration [60], which targeted BSCCO-2212 as HTS conductor suitable for high field. This activity is part of the scope of the US-MDP [61] now addressing both BSCCO-2212 and REBCO in various cables (Rutherford and CORC™) and magnet (racetracks, cos-theta and canted cos-theta) configurations. As anticipated, in the EU the first seeds initiated already with the EU-FP7 EuCARD collaboration and were pursued intensely with the follow-up EU-FP7 EuCARD2 [62] and EU-H2020 ARIES [63] programs.

In Europe the HTS conductor activity was directed to REBCO, whose selection was mainly driven by the perceived potential and by consideration of relatively easy coil fabrication technology with a conductor that does need heat treatment of the whole coil [64]. After the attempt of flat racetrack with simple stacked tape with a fairly good amount of copper as used in EuCARD insert, EuCARD2 had the ambition to make to test HTS in a real accelerator magnet configuration. So EuCARD2 dipole magnets were designed with an open bore and wound with a cable. A multi-tape cable, rather than a simple tape, was selected as conductor, in order to operate the magnet in the10 kA range, very much like the usual LTS high field dipoles, to limit inductance to cope with quench protection and in order to profit of current sharing between cable sub-elements. In addition, transposed topology for the cable has been used, ruling out the use of simple stacked tape [54]. This allowed to explore the potential of Roebel cable whose potential for HTS tape composed cable was just being demonstrated [65]. The result of these activities are small demonstrator magnets. A simple flat coil, M0-4 was wound and assembled to check windability and behavior of Roebel cable. It reached almost 13 kA at the minimum temperature of 15 K (test at higher amperage was not possible due to limitation of the test station). Then two real 35 mm open bore (coil gap of 40 mm) dipoles, M2-1-2 and M2-3-4, were manufactured that reached 3.5 T and 4.5 T, respectively, at 5 K in stand-alone mode [66]. This last result is a very encouraging figure, just 10% short of the design field of 5 T, probably due to damage to the leads outside the coils. However, the use of Roebel cable proved to be cumbersome and the use of coil technology with non-insulated conductor, as well as new considerations on possible futility of transposition with HTS tapes, open the way to a single tape or a stacked tape cable as conductor for magnet winding. This might change completely the direction to go.

Meanwhile at LBNL in the U.S., two lines continue to be developed. The first one is pursuing BSCCO-2212 conductor in form of Rutherford cable, first with race-track coil geometry and then with CCT coil geometry [67]. The other line is pursuing REBCO round cable (so-called CORC™) in a CCT layout coil [68]. Recently R&D activities on the HTS insert coils have started also at FNAL. The work is focused on both REBCO [69] and BSCCO-2212 [70] conductor and shell-type coils with stress management. The R&D is progressing steadily, though slowly, with results similar to those achieved in Europe. Figures 13, 14 show the cables, coil cross-section and pictures of these small-scale pioneering HTS dipole coils. As mentioned above, the collection of the main results recently achieved in Europe and the U.S. in Fig. 6.1.2-1 shows only the initial part of a path that will hopefully lead to similar, or even better, results obtained for Nb_3Sn.

The next step beyond the further advance of the HTS conductor and coils technology is to study the insertion of

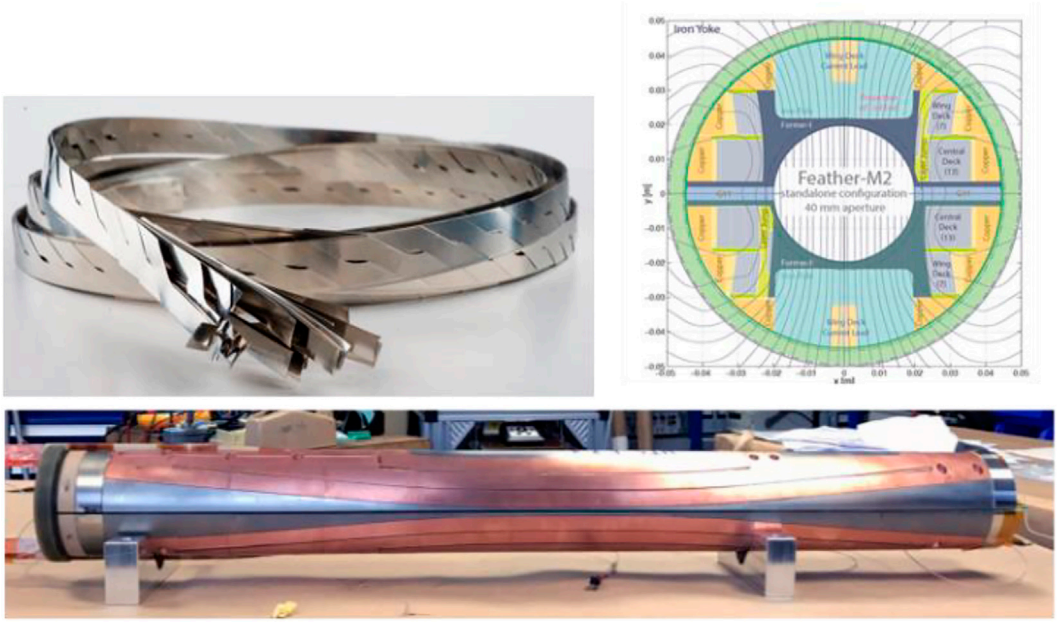


FIGURE 13
EuCARD2 Roebel cable (top left), aligned block dipole sketch (top right), and first real insert coil before test at CERN (bottom).

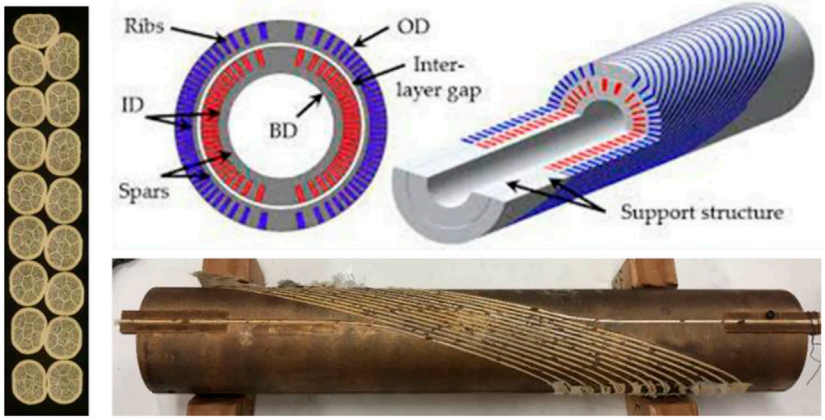


FIGURE 14
BSCCO-2212 Rutherford cable (side left), CCT coil cross-section and 3D view (top right), and initial CCT BSCCO-2212 insert coil at LBNL (bottom right).

these small-size demonstrators in large bore, LTS background magnets. This is to increase the central field and demonstrate the ability to break the barrier of LTS magnet performance, while at the same time exploring this new range of field. One EuCARD2 dipole, after an attempt to be powered in a background field underwent a severe degradation, whose origin is not yet clear.

6.3 Regional magnet development programs

6.3.1 US magnet development program

The US Magnet Development Program (MDP), formed by the US DOE Office of High Energy Physics in 2015 [49], following the 2013 “Snowmass” decadal community planning

process, is composed of researchers from leading DOE Laboratories and Universities focused on developing accelerator magnets for colliders.

The overarching goals of the program include 1) exploration the performance limits of Nb_3Sn accelerator magnets, focusing on minimizing the required operating margin and significantly reducing or eliminating magnet training; 2) development and demonstration of HTS accelerator magnets with a self-field of 5 T or greater, compatible with operation in a hybrid HTS/LTS magnet for operation fields beyond 16 T; 3) investigation of fundamental aspects of magnet design and technology that can lead to substantial performance improvements and magnet cost reduction; and 4) performing Nb_3Sn and HTS conductor R&D with clear targets of increasing performance, understanding present performance limits, and reducing the cost of accelerator magnets.

There are many paths and opportunities for high-field accelerator magnet development. To provide focus, the MDP has centered its current research around specific themes. First, the program explores stress-managed structures as a means to enable high-field accelerator magnets. By intercepting the large azimuthal and radial forces that accumulate in the magnet midplane area in conventional accelerator magnet designs, the stress-managed structures have the potential to mitigate degradation to the strain-sensitive Nb_3Sn and HTS superconductors in high-field environments. Second, it explores the use of hybrid HTS/LTS magnets for high field accelerator magnets. The use of HTS opens a path to high fields beyond that achievable with LTS materials, and by using LTS material in the low-field region the technology seeks to optimize the cost and size of the overall system. Third, it strives to improve our understanding and control of magnet technology and magnet performance for accelerators through the development of advanced diagnostics and modeling tools. Integrating these into the magnet design and testing provides critical feedback for further improvements and refinement, mitigating risk and ultimately leading to more cost-effective magnet. Fourth, to inform further promising avenues for magnet development, design studies are underway on high-field accelerator magnet concepts. Finally, in support of the high-field magnet needs, the program seeks to further develop superconductors, focusing on improving characteristics critical to accelerator applications such as enhanced transport current, improved production quality, and ultimately a reduction in material cost.

The themes described above are aligned with the goals established when the MDP was founded in 2016. The program has made progress towards its goals in advancing superconductors, developing core HTS magnet technologies, and demonstrating record Nb_3Sn accelerator magnet performance. The MDP roadmap has been updated in 2020, the updated plan is detailed in [50]. The program is structured around the themes described above, and has four components: Nb_3Sn magnets, HTS magnets, technology, and SC materials.

The development of Nb_3Sn magnet technology is a core element of the program. The successful four-layer cosine-theta magnet [71], designed and built by the FNAL team as part of the MDP, serves as the foundation and motivation for an effort to investigate and develop stress-managed magnet concepts, where the magnetic body forces (Lorentz forces) - that in traditional magnet designs accumulate in the dipole midplane area - are instead captured locally and transmitted to a mechanical structure. Two complementary approaches to stress-management are being pursued. The first, known as the Canted Cos-Theta (CCT) [72], is effectively the limiting case, where each turn in the coil is captured and in principle no force accumulation occurs. The second, known as the stress-managed cosine-theta (SMCT), is somewhat similar to a traditional cosine-theta dipole layout, but where the magnetic force acting on groups of turns are transferred to a support structure, leading to a more efficient use of conductor as compared to the CCT. The CCT approach is currently under rapid development using subscale prototypes that allow for systematic development and testing of design details. Similarly, the SMCT concept is being developed using the mirror-structure concept, which enables a single coil to be tested in a realistic magnet environment for more rapid development of the concept. By pursuing both concepts in parallel, the program can most efficiently explore the potential of stress-managed structures for accelerator magnet application. We note that in traditional dipole magnet designs the mechanical stress scales with field and radius; the stress-managed structures seek to break that scaling, enabling access to higher magnetic field and to larger bore magnets.

A second major thrust of the MDP is the development of HTS magnet technology. The program is pursuing the use of both REBCO and BSCCO-2212; each has advantages and disadvantages. Since both materials are strain-sensitive and prone to damage under mechanical stress, the stress-managed mechanical concepts described above are highly relevant and are being adapted and applied for the HTS magnets. Significant progress has been made over the last few years, both in developing the basic HTS magnet technology and in fabricating and testing prototype HTS dipole magnets, and the program expects to achieve a 5 T stand-alone HTS dipole demonstration in the near future. In parallel, the program is preparing for first tests of hybrid magnets wherein an HTS insert dipole is positioned inside a Nb_3Sn outsert. Such an approach is anticipated to be the most efficient and effective means of achieving dipole fields above 16 T. To accomplish these goals, the MDP leverages the expertise and facilities available at all the partner institutions.

A third major element of the program is the development of fundamental magnet technologies, encompassing critical

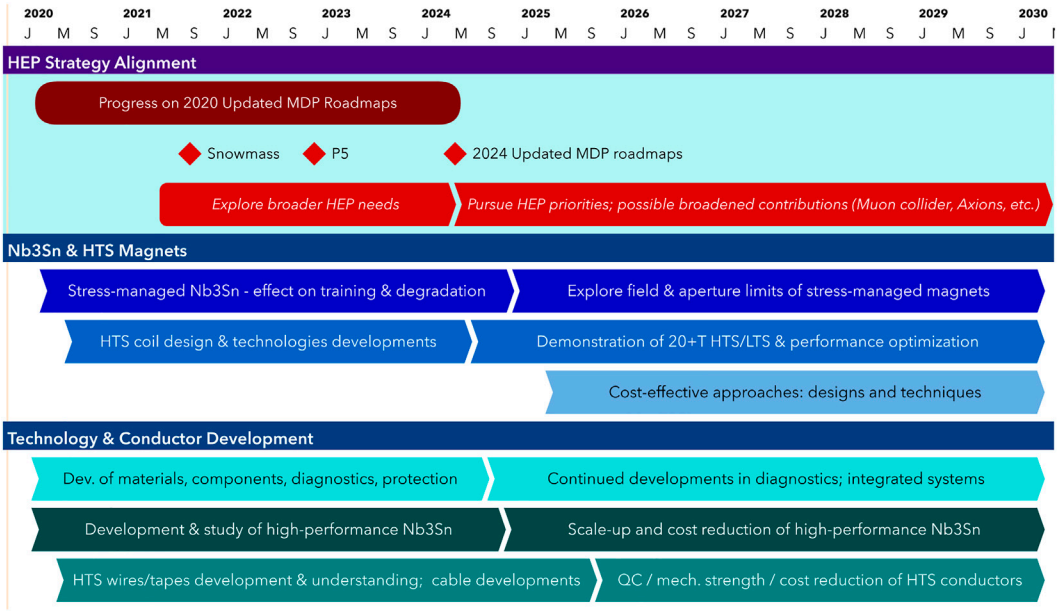


FIGURE 15
A 10-year roadmap for the US-MDP.

elements such as advanced modeling techniques, novel diagnostics, the exploration and characterization of novel magnet materials, and new data analysis techniques. Tightly integrated with the first two thrusts of the program, this element is central to advancing our communities understanding of magnet behavior and performance limitations. This knowledge is used to explore future paths for the program, and in particular is being applied to scope out paths towards ~20 T hybrid magnet designs.

Finally, at the heart of high field accelerator magnets is the superconductor. The MDP invests in commercial conductors, working closely with industry to make sure the material properties are matched to accelerator magnet needs. Investments are also made in conductor R&D, where new superconductor architectures tailored to the accelerator magnet communities' needs are explored. The development of new BSCCO-2212 powder manufacturing, which, when coupled with overpressure-processing, has led to dramatic increase in overall current density achieved in BSCCO-2212 magnets over the last decade, is one example. Similarly, reductions in REBCO substrate thickness, coupled with improved pinning *via* Zr doping, has led to major advances in REBCO cable performance. In the Nb_3Sn realm, novel concepts that introduce vortex-pinning enhancement through doping with ZrO_2 particles and/or Hf [73] have led to significant increases in current density at high fields, further expanding the potential reach of Nb_3Sn. There have also been very intriguing developments in the use of high heat capacity ("high-Cp") materials either internal to Nb_3Sn wires [74] or as part of the cable [75], with the potential to significantly improve Nb_3Sn magnet training.

A 10-year roadmap for the US-MDP aligned with the US High Energy Physics community planning process is shown in Figure 15.

6.3.2 European high field magnet program

Present and future demands from HEP, as discussed above, were included in the process of updating the European Strategy for Particle Physics (ESPP). The ESPP consultation and synthesis process started with the Open Symposium of Granada in May 2019 and completed with the endorsement of the ESPP update by the CERN Council, in June 2020 [76, 77]. The references quoted contain strong statements supporting R&D activities on high-field accelerator magnets. The above statements have been translated in the two long-term technical goals of the HFM R&D program:

Goal 1. The search for the maximum practical operating field that can be separated from the development of accelerator technology, defining two sub-goals:

I. Quantify and demonstrate Nb_3Sn ultimate field, with a projected target of 16 T, developing conductor and magnet technology towards *ultimate* Nb_3Sn performance. The target field may be revised based on the performance of demonstrators and model magnets, as well as practical considerations of operating margin and cost.

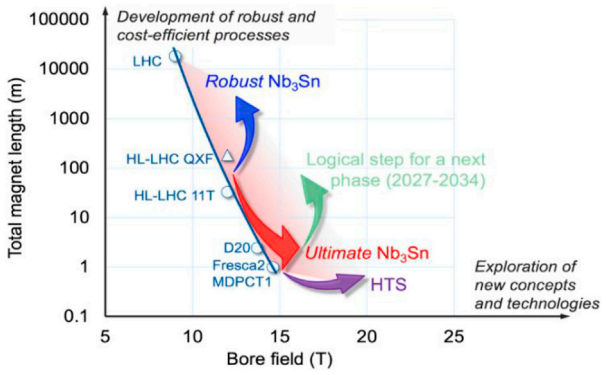


FIGURE 16
Schematic representation of the objectives of the European HFM R&D program, intended to advance in synchronism on the front of maximum field (red for Nb_3Sn, purple for HTS) as well as large-scale production (blue). The possible evolution for the longer term is represented in green.

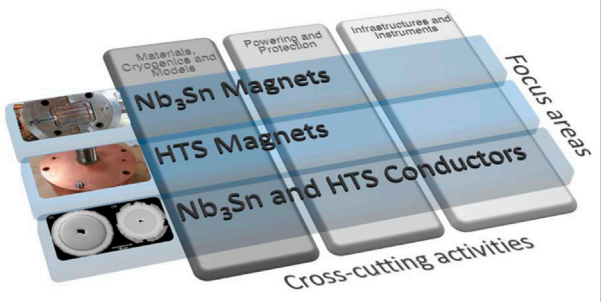


FIGURE 17
Schematic representation of the proposed EU R&D program for high field magnets.

II. Develop Nb_3Sn magnet technology for a large-scale production, simplifying manufacturing and striving for cost reduction. This is done taking HL-LHC with an ultimate field in the range of 12 T and a relatively small production as suitable benchmark. Nb_3Sn magnets of this class should be *robust* from manufacturing, through test, installation and commissioning up to accelerator operation. This development will be measured against the performance of long demonstrator magnets.

Goal 2. Develop HTS magnet technology for accelerators, providing a proof-of-principle of beyond the reach of Nb_3Sn. This program breaks the evolutionary changes of LTS magnet technology, calling for a number of significant innovations. The target dipole field is set for 20 T, well above the reach of Nb_3Sn. HTS should be considered for applications where not only high field is sought, but also higher operating temperature than liquid helium, large operating margin, and radiation tolerance. Finally, for HTS the possibility of full HTS magnets to operate the accelerator at 14–16 T, i.e., the same field that Nb_3Sn, but at much higher temperature of 15–20 K will be explored. The main driver of this study is the possibility to operate the full accelerator at higher temperature with significant saving factor on the electric power needed by cryogenics.

A suggestive graphical representation of the main objectives above is shown in Figure 16, where the total length of magnets produced with a certain technology plotted vs. the maximum field reached. Note that the HL-LHC quadrupoles QXF are included in the plot. The line passing through the points on the plot defines a boundary of the state-of-the-art and goes from the large-scale end of the nearly 20 km of Nb-Ti LHC double-aperture magnets in the range of 9 T ultimate field, to the high-field end of single model magnets, each about 1 m long, and reaching 14.5 T maximum field. The HL-LHC magnets represent intermediate field, about 12 T, and total length scale in the range of 100 m. The objectives of the European HFM R&D program correspond to an extension of the field reach by moving the boundary along the horizontal axis (magnetic field), profiting from advances in both Nb_3Sn and HTS magnet technology, and an extension of the production capability by moving the boundary along the vertical axis (magnet length), through the development of robust design and manufacturing processes.

The target represented by the graphical representation of Figure 16 only defines a first step in the R&D, which should enfold in the 2021-2027 period. The parallelism in the development is necessary to provide significant advances within 5 years and feed the discussion for the next iteration of the European Strategy for Particle Physics with crucial deliverables. Advancing on both fronts of maximum field and large-scale capability will provide options for a decision on magnet technology towards the next hadron collider. A follow-up can already be imagined, depending on the successful outcome of the first phase. This should be dedicated to prototyping the new generation of high field magnets. This is represented by the green arrow in Figure 16, although the choice of the field level and magnet length to be realized are only indicative, and will depend on the results of the first phase of the R&D.

The R&D targets formulated above respond directly to the demands of the principal stakeholders. For Nb_3Sn magnets, the target field and large-scale capability is driven by the demands of an FCC-hh [78], and the staged approach is compatible with the allotted development time of the integrated FCC program [79].

Even if the main focus of the European HFM R&D is on dipole magnets for an FCC-hh, an R&D on high-field Nb_3Sn and HTS magnets along the lines outlined above will be highly relevant to develop suitable design and technology solutions also for other collider options, and in particular a muon collider. Specific examples are HTS conductor and coil winding technology, exploring partial- and no-insulation windings, relevant for the target and cooling solenoids, and the study of stress management in Nb_3Sn magnets, with

impact on the muon collider ring and IR magnets. Also, very important is considering HTS magnet operation at temperatures above liquid helium, which is relevant to understanding the operating margin in the high heat load and radiation environment of a high-energy muon collider ring.

The structure of the present European program towards high-field superconducting accelerator magnets, fruit of the consultation process driven by the LDG [80] is represented graphically in Figure 17. The work is divided in three focus areas, in foreground, supported by three cross-cutting R&D activities, in the background. The focus areas are: 1) Nb_3Sn magnets; 2) HTS magnets; and 3) Nb_3Sn and HTS conductors. Activities in the focus areas will directly provide elements for critical decisions for the field reach of a given magnet technology or for superconductor performance specifications. The cross-cutting R&D activities are grouped into 1) structural and composite materials, cryogenics and thermal management, and modeling; 2) powering and protection; and 3) infrastructure for production and test as well as instruments for diagnostics and measurement. These activities are intended to provide the background research needed to respond to the challenges identified in the focus areas.

6.3.3 Muon collider driven magnet R&D

The updated European Strategy for Particle Physics, produced by the Laboratory Directors Group (LDG) under the auspices of the CERN Council, includes a strong recommendation towards the development of muon beams, in particular in view of a high-energy, high luminosity Muon Collider. The focus is on two energy ranges, around 3 TeV and above 10 TeV. The International Muon Collider Collaboration (IMCC) is focusing on the development of a Muon Collider design and address the associated technical challenges. The collaboration aims at establishing feasibility of a muon collider to develop the concept and technology to a level of maturity that allows committing to its construction and to assess the physics reach.

The full conceptual design of a Muon Collider poses grand challenges and is expected to demand a considerable effort. In particular, the challenges for the key magnet systems of the complex are well beyond the reach of present technology. New concepts, demonstrators, models and prototypes will be needed to prove construction and performance in the extraordinary operating conditions expected. The MAP study produced a consistent baseline set of requirements and technology options [81]. The work of IMCC magnet working group will evolve from these configurations. Clearly, the hope is to profit from the advances in magnet technology in the past 10 years, and in particular HTS.

6.3.3.1 Muon target, capture and cooling solenoids

The challenges identified in this part of the muon collider complex can be represented well by the first and last magnets, namely the target solenoid and the final cooling solenoid. In the scheme considered at present, the target solenoid requires a field of 20 T in a 150 mm bore. The magnet needs a large bore, driven by shielding requirement, and has substantial residual heat load, resulting from the nuclear interactions of the protons in the target. Mechanics, magnet protection, and the ability to sustain and remove a high heat influx are the aspects that dominate the design. This may require the development of conductor reinforcements, e.g. inspired by work done in large detector magnets or pulsed solenoids for fusion. HTS may be an interesting option to reduce the magnet size and increase the operating temperature. Indeed, recent developments and achievements towards the use of large HTS cables for fusion makes this option particularly interesting.

The solenoid at the end of the cooling chain directly drives the final emittance of the muon beam, which is inversely proportional to the bore field. The highest bore field is hence required in the range of 40 T (minimum) to 60 T (target) in a 50 mm bore. The best technology for such a magnet is an all-superconducting coil, which has significant advantages on footprint, investment and operating costs if compared to a superconductive/resistive hybrid option. An ultra-high field solenoid of this class has not been built yet, and the challenges span from the choice of the HTS material, possibly in the form of a cable, the extraordinary forces and stresses, quench management at large stored energy, field stability (in case of non-insulated or partially insulated winding technology is adopted), and the integration of LTS and HTS windings (in case the magnet is built as an LTS/HTS hybrid). Going well beyond available technology, this solenoid will require considerable R&D and demonstration, including studies and measurements of effects that are specific to ultra-high fields, such as the internal forces developed in tapes hosting shielding currents, or the helium phase separation in large magnetic field gradients. It must be noted that development of solenoids with such performance is of high interest also for other fields of science and societal applications, such as the high magnetic field science and the NMR spectroscopy.

6.3.3.2 Dipoles and quadrupoles for muon acceleration, storage and collision

R&D for muon acceleration magnet will be focused primarily on reducing and managing the stored energy of the accelerator ring which for the dipole circuit is of the order of several tens of MJ. At the required acceleration rate, a full magnet swing takes a fraction of millisecond, which corresponds to the peak power in the range of tens of GW. The solution envisaged is to power the magnets using power converters relying on energy storage (e.g. capacitor banks), and only replenish the energy lost during one full cycle (mainly) due to the coil resistance and hysteresis losses in iron, Alternatives such as flux-pumps with high Q-factor may be considered to improve on the energy efficiency.

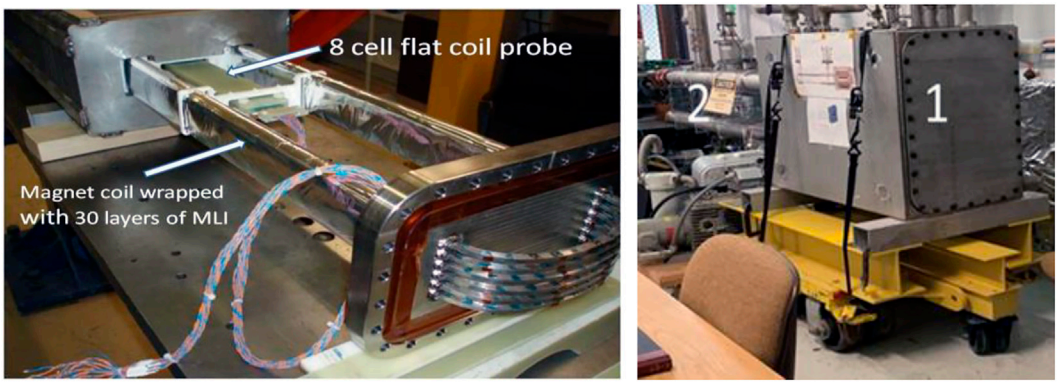


FIGURE 18
Fast cycling dipole model based on REBCO tape.

The use of HTS for the fast pulsed magnets is an option considered, not so much for the potential energy saving, but rather towards higher fields and faster acceleration, with end benefits on the collider integrated luminosity. Pioneering work is being performed at FNAL, where configurations are devised to achieve exceptional performance, shown in Figure 18. The most recent results with a bore field of approximately 0.3 T at field ramp-rate of 300 T/s have shown that the ramp-rate target for the last RCS in the acceleration chain on the level of 400 T/s is within reach, albeit at reduced field [82].

While it is obvious that a strong focus is necessary on the control of AC loss at the exceedingly high ramp-rate, a particular advantage of HTS would be to increase the field swing beyond the reach of resistive magnets, which would reduce the accelerator length and increase the muon survival through the acceleration process. Further priority development topics are quench detection and protection in high voltage conditions, as well as fatigue and ageing in cyclic operation. R&D on magnets and powering systems for a muon accelerator will profit RCS for nuclear physics, as well as accelerator driven systems for radioisotopes transmutation and fission. We also expect that fast and precise management of large electric power at the GW level is an R&D where connections to several technologies can be found beyond accelerator magnets.

The challenges for dipole and quadrupole magnet technology associated with the demands stemming from the Storage Ring arc and Interaction Region are driven by the need of a high field, to keep the collider as compact as possible, and a large aperture, to accommodate shielding and large beam size in the IR magnets to achieve high luminosity. Indeed, the high level and distribution of radiation in the collider ring calls for dipole magnets with aperture much larger than a typical hadron collider. This is needed to accommodate a thick high-Z absorbers to protect the SC coils. As an alternative, open midplane configurations can be devised to provide a path for the decay electrons to high-Z external absorbers, thus sparing the coils and reducing the direct heat load. However, studies of open midplane dipoles revealed several challenges. In addition to the management of the large vertical forces, heat deposition spreads beyond the midplane due to the vertical deflection of the decay electrons. Further studies are planned to see what the most adapted magnet configuration is. For MC with c. o.m. Energies up to ~3 TeV magnet requirements are marginally within the reach of Nb_3Sn technology. For a higher energy MC, the combination of high magnetic field, large aperture and high heat flux may be resolved by devising the storage ring and IR magnets as hybrid of LTS and HTS coils. A large benefit would be obtained also by operating the HTS at higher temperature, absorbing the relatively large heat and radiation load with improved cryogenic efficiency. An important issue to be addressed is the stress management in high-field large-aperture magnet coils to avoid degradation and damage of the brittle Nb_3Sn and HTS superconductor. To control the mechanical stresses in brittle Nb_3Sn and HTS coils, a stress management approach has been proposed and needs to be developed and demonstrated. It is being experimentally studied by the US-MDP (see section 6.2.1).

6.3.4 High field accelerator magnet R&D in China

The Institute of High Energy Physics of Chinese Academy of Science (IHEP-CAS) is carrying out an R&D on high field accelerator magnets in partnership with research organizations working on the superconductivity and on the advanced HTS materials. This R&D program addresses the following key issues related to high-field superconducting magnet technology:

1. Development and test of new HTS materials with superior performance for application in accelerator magnets; investigation of key factors influencing material current-carrying capabilities on the microstructure and vortex dynamics level; development of advanced technologies of

HTS wires with high critical current density for high field applications and high mechanical strength.
2. Development of high-current HTS cables and significantly reduce their costs; exploration of innovative structures and fabrication processes of high-field superconducting magnets based on advanced superconducting materials and helium-free cooling procedures.
3. Exploration of novel methods of stress management and quench protection for high-field superconducting accelerator magnets, particularly for high-field HTS insert coils.
4. Development of a prototype of accelerator magnet with high operation field and accelerator-quality field, building the foundation for using advanced HTS technology in high-energy particle accelerators.

The R&D of the high-field magnet technology and related advanced superconducting materials has been funded by the Chinese Academy of Sciences and the Ministry of Science and Technology on the level of 60M$ for 2018-2024.

As the first step, a 12 T subscale common-coil dipole magnet LPF1 with two apertures and graded coil was designed, fabricated, and tested. With two Nb_3Sn racetrack coils inside and four Nb-Ti racetrack coils outside, the design field of this hybrid two-aperture dipole magnet is 12 T at 6100 A current at 4.2 K with the margin of 17% for the design current and field. LPF1 was tested at 4.2 K. The magnet reached quench plateau around 10.2 T after the thirteen quenches. The parameters of the design, the process of the fabrication and the test performance of LPF1 are reported in [83].

A series of 35-mm diameter IBS (Iron Based Superconductor) coils including single pancake coils (SPC) and double pancake coils (DPC) were designed and wound using the seven-filament Ba122 ($Ba_{1-x}K_xFe_2As_2$) tape. The tape was produced by the Institute of Electrical Engineering of the Chinese Academy of Sciences (IEE-CAS). The tests of the IBS coils were carried out first at 4.2 K in the 10 T background field. The highest quench current of the coils at 10 T was 68.4 A, which is about 79% of the quench current at self-field and about 90% of the short sample limit. Then two SPCs were selected and tested at 4.2 K in 24 T background field [84]. The highest quench current of the coils at 24 T was 25.6 A, which is about 39% of the coil quench current at self-field. The details of coil tests will be published in SUST. These results suggest that the iron-based superconductors are promising materials for applications in high-field accelerator magnets. These studies continue.

6.3.5 Superconducting accelerator magnet R&D in Japan

R&D on superconducting accelerator magnets has been conducted at KEK in Japan for more than 40 years. One of the major accomplishments is the development and construction of the IR quadrupoles MQXA for the LHC at CERN. KEK has also developed superconducting magnets for beam line of the T2K neutrino experiment and solenoids for muon experiments COMET and g-2/EDM. At the present time KEK is responsible for the development of the large-aperture beam separation dipole D1 for HL-LHC (see section 3.2 above). All these magnets are using Nb-Ti superconductor. The R&D works on A15 (Nb_3Sn and Nb_3Al) and HTS advanced superconductors are also performed. Nb_3Sn conductor development for the Future Circular Collider (FCC-hh) is conducted in collaboration with CERN, Tohoku University, Tokai University, NIMS, and two Japanese industrial partners [85]. The development of HTS materials for high-field and high-radiation environment applications is performed by the US-Japan collaboration formed by KEK, Kyoto University, LBNL, and BNL [86].

As an extension of the abovementioned collaborative works, the next KEK mid-term goal is the development of radiation-hard high-field magnet technologies for future energy Frontier hadron colliders and high-intensity proton drivers. One of the R&D targets could be the 12 T 100-mm aperture beam separation dipole D1 for FCC-hh. Due to large aperture and high operation field of this magnet, mechanical stress management in the magnet structure and reduction of stress-strain sensitivity of the Nb_3Sn conductor are important R&D tasks. R&D of the mechanically strengthened Nb_3Sn conductor will be performed in collaboration with Tohoku University. Since the magnet will work in a relatively high radiation area, radiation hardness of the magnet structural materials, such as epoxy and insulations, have to be also studied. The ultimate goal of this R&D is to construct and test a Nb_3Sn dipole model.

For HTS development, a near term goal for KEK is the development of radiation hard HTS magnet technologies for a high-intensity muon production solenoid. At J-PARC Material and Life Science Facility (MLF), the construction of the second target station is being discussed. In this target station a muon production solenoid of about 1 T central field is directly attached to the 1 MW target that produces both muons and neutrons. The facility aims to produce 50 to 100 times more muons than the current MLF muon source that results in world leading intense muon source. The muon production solenoid requires high radiation hardness and high reliability on quench protection. The current US-Japan collaboration is planned to be involved in this project. The ultimate goal of this R&D is the realization of the second target station.

In the long term, the above 2 R&D programs can be extended to the R&D of 16–20 T accelerator magnets, by combining the 12 T Nb_3Sn large-aperture dipole with 4–8 T HTS or Nb_3Sn insert coils. For the HTS part, the development of high current cable conductor is needed. The current US-Japan collaboration already included the study on the high-current cable development and the effort will continue. Extensive studies have been performed in the collaboration with Kyoto University of shielding currents in HTS tape conductors to achieve the accelerator field quality in HTS magnets. Since the

current study on both Nb_3Sn and HTS magnet technologies include the radiation hardness studies, the results of the R&D will lead to high-field, high-radiation hard accelerator magnet technology that is required for FCC insertion quadrupoles.

7 Summary and next steps

Superconducting accelerator magnets are the key enabling technology for present and future particle accelerators in modern high energy physics. All the present accelerator magnets have used for decades the Nb-Ti superconductor. The practical performance limit of this technology in accelerator magnets is limited by 8–9 T. This field level was realized in the Tevatron, HERA, RHIC and now in various LHC magnets.

This review shows that the present of superconducting accelerator magnets for high energy physics is highly dynamic, and they will remain an exciting field in the future thanks to the challenges and perspective developments driven by next generation colliders presently under study. The development of a new generation of accelerator magnets based on Nb_3Sn superconductor with operation fields up to 15–16 T has shown in the last 2 decades a good progress in the U.S., Europe and Asia. On the short term, within the next 5 years, this technology will be implemented in the HL-LHC by using large-aperture high-gradient final-focus quadrupoles in ATLAS and CMS experiments. On the longer term, dipole and quadrupole magnets with nominal operation fields up to 16 T are planned for FCC-hh and 3 TeV MC. Several key technological issues important for the performance of these magnets need to be resolved. They are in the focus of national magnet R&D programs in the U.S. and Europe.

The HL-LHC promises a technology breakthrough with the introduction of Nb_3Sn in the palette of superconducting materials for accelerators. Having surpassed Nb-Ti in terms of performance with characteristics that broadly match the strict beam requirements, the present focus is on the production of the first Nb_3Sn magnets to be installed and operated in a running accelerator. Demonstrating performance retention throughout the whole magnet life, mastering electro- and thermo-mechanical loads, and achieving a considerable simplification of manufacturing will be not only necessary to the exploitation of Nb_3Sn on large scale, but also a useful contribution for the next step in magnet technology based on HTS materials.

More ambitious R&D work towards 20 T and even higher-field magnets, which are considered for SPPC and high-energy MC and based on cost-effective HTS/LTS coils has also been started recently. Indeed, we see HTS as the upcoming technology breakpoint for future colliders. This is obviously because of the extraordinary field reach, which will be necessary to achieve manageable infrastructure scale and cost, but not only. New experiments at the scale of the colliders described here will have to produce a business case that includes environmental impact and sustainability. HTS materials promise higher fields and improved energy efficiency through operation at higher cryogenic temperature than liquid helium, or helium free configurations. Much work needs to be done to exploit the potential of HTS materials, starting from their basic conductor characteristics, through magnet design and technology, finally including considerations of operation and interaction with beam optics. The field is rich and bears connections and implications to other fields of research and societal applications, which makes it even more exciting.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. O Bruning, editor. *The LHC design report*. Geneva, Switzerland: CERN (2004).

2. ATLAS Collaboration. Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC. *Phys Lett B* (2012) 716:1. doi:10.1016/j.physletb.2012.08.020

3. CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys Lett* (2012) B716:30. doi:10.1016/j.physletb.2012.08.021

4. The European Strategy Group. *Deliberation document on the 2020 update of the European strategy for particle physics*. Geneva, Switzerland: CERN-ESU (2020).

5. Adolphsen M, Angalakalini D, Arndt M. In: N Mounet, editor. *European strategy for particle physics - accelerator R&D roadmap*. CERN (2022).

6. Sowmass (2021). Available from: https://snowmass21.org/.

7. Building for Discovery (2018). Available from: https://www.usparticlephysics.org/wp-content/uploads/2018/03/FINAL_P5_Report_053014.pdf.

8. CEEN. Future circular collider (2020). Available from: https://home.cern/science/accelerators/future-circular-collider.

9. CEPC. Circular electron positro collider (2012). Available from: http://cepc.ihep.ac.cn/intro.html.

10. Palmer RB. Muon colliders. *Rev Accl Sci Tech* (2014) 7:137–59. doi:10.1142/s1793626814300072

11. Wang M-H, Noschkov J-P, Cai Y. *Design of a 6 TeV muon collider*. IOP Publishing for Sissa Medialab (2016). JINST 11 P09003.

12. Boscolo M, Delahaye Y, Palmer M. The future prospects of muon colliders and neutrino factories. *Rev Accl Sci Tech* (2019) 10(1):189–214. doi:10.1142/s179362681930010x

13. Adolphsen M, Angalakalini D, Arndt M. *European strategy for particle physics – accelerator R&D roadmap* (2022). arXiv:2201.07895 [physics.acc-ph].

14. B Oliver L Rossi, editors. The high luminosity large hadron collider. In: *Advanced series on direction in high energy physics*, 24. World Scientific publisher (2015).

15. Rossi L, Brüning O. *Progress with the high luminosity LHC project at CERN"*. Melbourne: IPAC (2019). p. 17. doi:10.18429/JACoW-IPAC2019-MOYPLM3

16. Brüning O, Gray H, Klein K, Lamont M, Narain M, Polifka R, et al. The scientific potential and technological challenges of the High-Luminosity Large Hadron Collider program. *Rep Prog Phys* 85:046201. doi:10.1088/1361-6633/ac5106,:

17. Kephart R, Lamm MJ, Limon P. The U.S. LHC accelerator research program: A proposal (2003). Available from: https://www.uslarp.org/LARP_Proposal.pdf.

18. Bottura L, de Rijk G, Rossi L, Todesco E. Advanced accelerator magnets for upgrading the LHC. *IEEE Trans Appl Supercond* (2012) 22:4002008–3. doi:10.1109/TASC.2012.2186109

19. Zlobin AV. N_3Sn 11 T dipole for the high luminosity LHC (FNAL). In: D Schoerling AV Zlobin, editors. *Nb_3Sn accelerator magnets*. Springer (2019). doi:10.1007/978-3-030-16118-7_8

20. Bordini B. Nb_3Sn 11 T dipole for the high luminosity LHC (CERN). In: . V Zlobin, editor. *Nb_3Sn accelerator magnets*. Springer (2019). p. 223–58. doi:10.1007/978-3-030-16118-7_9

21. Barzi E, Wokas T, Zlobin A. Sensitivity of Nb/sub 3/Sn Rutherford-type cables to transverse pressure. *IEEE Trans Appl Supercond* (2005) 15:1541–4. doi:10.1109/tasc.2005.849162

22. Ferradas Troitino J, Bagni T, Barth C, Bordini B, Ferracin P, Gamperle L, et al. Effects of the initial axial strain state on the response to transverse stress of high-performance RRP Nb_3Sn wires. *Supercond Sci Technol* (2021) 34(3):035008. doi:10.1088/1361-6668/abd388

23. Hafalia RR, Bish PA, Caspi S. *A new support structure for high field magnets*. LBNL-47796 SC-MAG 738.

24. Ravaioli E, Ambrosio G, Bajas H, Duda M, Ferracin P, Troitino JF, et al. Quench protection studies for the high luminosity LHC Nb_3Sn quadrupole magnets. *IEEE Trans Appl Supercond* (2021) 31:1–5. doi:10.1109/tasc.2021.3055160

25. Todesco E, Bajas H, Bajko M, Ballarino A, Bermudez SI, Bordini B, et al. The High Luminosity LHC interaction region magnets towards series production. *Supercond Sci Technol* (2021) 34:053001. doi:10.1088/1361-6668/abdba4

26. Ranjbar V. *The RCS design status for the electron ion collider", IPAC'21*. Brazil: Campinas, SP (2021).

27. Parker B, Escallier J, Ghosh A, Jain A, Marone A, et al. BNL direct wind superconducting magnets. *IEEE Trans Appl Supercond* (2012) 22:4101604. doi:10.1109/tasc.2011.2175693

28. Witte H, Parker B, Palmer R. Design of a tapered final focusing magnet for eRHIC. *IEEE Trans Appl Supercond* (2019) 29:1–5. doi:10.1109/tasc.2019.2902982

29. Meyer D, Flasck R. A new configuration for a dipole magnet for use in high energy physics applications. *Nucl Instr Methods* (1970) 80:339–41. doi:10.1016/0029-554x(70)90784-6

30. Caspi S, Arbelaez D, Brouwer L, Gourlay S, Prestemon S, Auchmann B. Design of a canted-cosine-theta superconducting dipole magnet for future colliders. *IEEE Trans Appl Supercond* (2017) 27:1–5. doi:10.1109/tasc.2016.2638458

31. Zlobin AV. *Large-aperture high-field Nb_3Sn dipole magnets*. Vancouver, Canada: IPAC (2018).

32. Schoerling D. The 16 T dipole development program for FCC and HE-LHC. *IEEE TAS* (2019) 29(5):4003109. doi:10.1109/TASC.2019.2900556

33. Lorin C, Fleiter J, Salmi T, Schoerling D. Exploration of a two-layer Nb_3Sn designs of the future circular collider main qaudrupoles. *IEEE Trans Appl Supercond* (2019) 29(5):1. doi:10.1109/tasc.2019.2892814

34. Ballarino A, Bottura L. Targets for R&D on Nb_3Sn conductor for high energy physics. *IEEE Trans Appl Supercond* (2015) 25(3):1–6. doi:10.1109/tasc.2015.2390149

35. Perez JC, Bajko M, Bourcey N, Bordini B, Bottura L, Troitino SF, et al. Construction and test of the enhanced racetrack model coil, first CERN R&D magnet for the FCC. *IEEE Trans Appl Supercond* (2022) 32:1–5. doi:10.1109/tasc.2022.3163064

36. CEPC. *The CEPC conceptual design report*, Vol. I. Accelerator (2018). arXiv: 1809.00285 Available from: http://cepc.ihep.ac.cn/CDR_v6_201808.pdf.

37. CEPC. *CEPC accelerator," submitted to European strategy in 2019* (2019). ArXiv: 1901.03169.

38. Kong ES, Wang C, Wang L, Wang X, Cheng D, Zhang K, et al. Conceptual design study of iron-based superconducting dipole magnets for SPPC. *Int J Mod Phys A* (2019) 34:1940003. doi:10.1142/s0217751x19400037

39. Budker GI. *Proc. Of the 7th int. Conf. Part. Accel. Physics (yerevan, 1969)*, (1970). 33.

40. The U.S. Neutrino factory and muon collider collaboration (NFMCC) (2010). Available from: http://www.cap.bnl.gov/mumu/mu_home_page.html.

41. The U.S. Muon accelerator program proposal (2010). Fermilab-TM-2459-APC Available from: https://map.fnal.gov/.

42. Schulte D. The international muon collider collaboration. *Proc 12th Int Part. Acc Conf Ipac-* (2021) 21:3792–5. Available from: http://muoncollider.web.cern.ch.

43. Weggel RJ. *Design of magnets for the target and decay region of a Muon Collider/Neutrino Factory target*, IPAC2013. Shanghai, China (2013). p. 1514.

44. Stratakis D, Palmer RB. Rectilinear six-dimensional ionization cooling channel for a muon collider: A theoretical and numerical study. *Phys Rev ST Accel Beams* (2015) 18:031003. doi:10.1103/physrevstab.18.031003

45. Kashikhin VV, Barzi E, Lamm M, Sadovskiy Y, Zlobin A. Study of high field superconducting solenoids for muon beam cooling. *IEEE Trans Appl Supercond* (2008) 18:928–32. doi:10.1109/tasc.2008.921946

46. Summers DJ. Hybrid rings of fixed 8T superconducting magnets and iron magnets rapidly cycling between −2 T and +2 T for a muon collider. In: *Proc. Of the 1996 DPF/DPB summer study on new directions in high-energy physics*. Stanford Linear Accelerator Center (1996). p. 238.

47. Kashikhin VV, Alexanin Y, Mokahav Nv. *High-field combined function magnets for a 1.5×1.5 TeV MC storage ring*. New OrleansUSA: IPAC2012Louisiana (2012). 3587.

48. Kashikhin VV, Alexanin Y, Mokahav Nv. *Magnets for interaction regions of a 1.5×1.5 TeV muon collider," IPAC2012*. LouisianaUSA: New Orleans (2012). 3584.

49. Gourlay S. The U.S. Magnet development program plan (2016). Available from: https://escholarship.org/uc/item/5178744.

50. Prestemon S, Amm k, Cooley L. The 2020 updated roadmaps for the us magnet development program. (2020). Available from: https://arxiv.org/abs/2011.09539.

51. Aleksan R (2004). Coordinated accelerator research in Europe. Final report Available from: https://cordis.europa.eu/project/id/506395/reporting.

52. Devred A, Baudouy B, Baynham D, Boutboul T, Canfer S, Chorowski M, et al. Status of the next European dipole (NED) activity of the collaborated accelerator research in Europe (CARE) project. *IEEE Trans Appl Supercond* (2005) 15(2):1106–12. doi:10.1109/TASC.2005.849506

53. de Rijk G. The EuCARD high field magnet project. *IEEE Trans Appl Supercond* (2012) 22:4301204. doi:10.1109/TASC.2011.2178220

54. Rossi L, Chiuchiolo A, Dhalle M, Durante M, Fazilleau P, Fleiter J, et al. The EuCARD2 future magnets program for particle accelerator high-field dipoles: Review of results and next steps. *IEEE Trans Appl Supercond* (2018) 28:1–10. doi:10.1109/TASC.2017.2784357

55. IFAST Project Office. *Innovation fostering in accelerator science and technology (I.FAST)*. H-2020 Proposal 101004730 To Eu Call Infrainnov-04-2020.(2020). Available from: https://ifast-project.eu/(last accessed November 2, 2021).

56. D Schoerling A Zlobin, editors. *Nb_3Sn accelerator magnets. Particle acceleration and detection series*. Springer (2019). doi:10.1007/978-3-030-16118-7

57. Rochepault E, Ferracin P. CEA–CERN block-type dipole magnet for cable testing: FRESCA2. In: *Nb_3Sn accelerator magnets. Particle acceleration and detection series*. Springer (2019). doi:10.1007/978-3-030-16118-7_12

58. Zlobin AV, Novitski I, Barzi E, Kashikhin VV, Carmichael J, Caspi S, et al. Development and first test of the 15 T Nb_3Sn dipole demonstrator MDPCT1. *IEEE Trans Appl Supercond* (2020) 30(4):1–5. doi:10.1109/TASC.2020.2967686

59. Devred A. *Status of the 11 T dipole and CERN magnet programs beyond HiLumi*. In: Presented at the 9th HL-LHC Collaboration Meeting (2019). Fermilab Available from: https://indico.cern.ch/event/806637/contributions/3487461/.

60. Tollestrup A, Larbalestier D. Very high field superconducting magnet collaboration. In: *Presented at EUCARD HTS magnet program meeting*. CERN (2011). Available from: https://indico.cern.ch/event/148320/contributions/1386701.

61. Prestemon S, Amm k, Cooley L. The 2020 updated roadmaps for the US magnet development program. Available from: https://arxiv.org/abs/2011.09539.

62. Eucard. EuCARD-2: Enhanced European coordination for accelerator research & development (2017). Available from: https://eucard2.web.cern.ch/eucard2/welcome.html.

63. ARIES (2017). Available from: https://aries.web.cern.ch/.

64. Bottura L, Rossi L. Technical and economical comparison YBCO/Bi-2212 magnets. Milestone MS68 report, EC-FP7-Eucard2 program Available from: https://edms.cern.ch/ui/#!master/navigator/document?D:1779412637:1779412637:subDocs.

65. Goldacker W, Nast R, Kotzyba G, Schlachter SI, Frank A, Ringsdorf B, et al. High current DyBCO-roebel assembled coated conductor (RACC). *J Phys : Conf Ser* (2006) 43:901–4. doi:10.1088/1742-6596/43/1/220

66. Rossi L, Senatore C. HTS accelerator magnet and conductor development in Europe. *Instruments* (2021) 5(8):8. doi:10.3390/instruments5010008

67. Shen T, Li P, Ye L. Heat treatment control of Bi-2212 coils: I. Unravelling the complex dependence of the critical current density of Bi-2212 wires on heat treatment. *Cryogenics* (2018) 89:95–101. doi:10.1016/j.cryogenics.2017.11.006

68. Wang X, Caspi S, Dietderich DR, Ghiorso WB, Gourlay SA, Higley HC, et al. A viable dipole magnet concept with REBCO CORC® wires and further development needs for high-field magnet applications. *Supercond Sci Technol* (2018) 31:045007–4. doi:10.1088/1361-6668/aaad8f

69. Kashikhin VV, Lombarado V, Velev G. *Magnet design optimization for future hadron colliders*, Melbourne, Australia: IPAC19(2019). paper THPTS084.

70. Zlobin AV, Novitski I, Barzi E. Conceptual design of a HTS dipole insert based on Bi2212 Rutherford cable. *Instruments* (2020) 4(29). doi:10.3390/instruments4040029

71. Zlobin AV, Novitski I, Barzi E. *Reassembly and test of high-field Nb_3Sn dipole demonstrator MDPCT1*, 31. IEEE TAS (2021)

72. Caspi S, Borgnolutti F, Brouwer L, Cheng D, Dietderich DR, Felice H, et al. Canted–cosine–theta magnet (CCT)—a concept for high field accelerator magnets. *IEEE Trans Appl Supercond* (2014) 24(3):1–4. doi:10.1109/tasc.2013.2284722

73. Balachandran S, Tarantini C, Lee PJ, Kametani F, Su YF, Walker B, et al. Beneficial influence of Hf and Zr additions to Nb4at%Ta on the vortex pinning of Nb_3Sn with and without an O source. *Supercond Sci Technol* (2019) 32.4:044006. doi:10.1088/1361-6668/aaff02

74. Xu X, Zlobin AV, Peng X, Li P. Development and study of Nb_3Sn wires with high specific heat. *IEEE Trans Appl Supercond* (2019) 29:1–4. doi:10.1109/tasc.2019.2892325

75. Barzi E, Novitsky I, Rusy A, Turrioni D, Zlobin AV, Peng X, et al. Test of superconducting wires and Rutherford cables with high specific heat. *IEEE Trans Appl Supercond* (2021) 31:1–8. doi:10.1109/tasc.2021.3069047

76. The European Strategy Group. *Update of the European strategy for particle physics"*. CERN-ESU-013 (2020).

77. The European Strategy Group. *Deliberation document on the 2020 update of the European strategy for particle physics"*. CERN-ESU-014 (2020).

78. Abada A. FCC-hh: The hadron collider. *Eur Phys J Spec Top* (2019) 228:755.

79. Benedikt M. Future circular collider the integrated programme (FCC-int) (2018). Available from: https://fcc-cdr.web.cern.ch/reports/EPPSU18_FCCint.pdf.

80. Adolphsen M, Angalakalini D, Arndt M. European strategy for particle physics – accelerator R&D roadmap. *CERN Yellow Rep Monogr* (2022) 1:1–270. doi:10.23731/CYRM-2022-001

81. Delahaye JP. *Muon colliders* (2019). arXiv:1901.06150.

82. Piekarz H, Claypool B, Hays S. *Fast cycling HTS based superconducting accelerator magnets: Feasibility study and readiness demonstration program driven by neutrino physics and muon collider needs* (2022). arXiv:2203.06253.

83. Wang CT, Cheng D, Zhang K, Wang Y, Kong E, Zhang Z, et al. Electromagnetic design, fabrication, and test of LPF1: A 10.2-T common-coil dipole magnet with graded coil configuration. *IEEE Trans Appl Supercond* (2019) 29(7):1. doi:10.1109/tasc.2019.2902983

84. Wang DL, Zhang Z, Zhang X, Jiang D, Dong C, Huang H, et al. First performance test of a 30 mm iron-based superconductor single pancake coil under a 24 T background field. *Supercond Sci Technol* (2019) 32:04LT0. doi:10.1088/1361-6668/ab09a4

85. Ballarino A, Hopkins SC, Bordini B, Richter D, Tommasini D, Bottura L, et al. The CERN FCC conductor development program: A worldwide effort for the future generation of high-field magnets. *IEEE Trans Appl Supercond* (2019) 29:1–9. doi:10.1109/tasc.2019.2896469

86. Tengming S, *A collaboration framework to advance high-temperature superconducting magnets for accelerator facilities*. US-Japan Science Technology Collaboration. 0000250069.

frontiers | Frontiers in Physics

TYPE Review
PUBLISHED 22 June 2023
DOI 10.3389/fphy.2022.916078

Check for updates

OPEN ACCESS

EDITED BY
Alessandro Tricoli,
Brookhaven National Laboratory (DOE),
United States

REVIEWED BY
Stefania Gori,
University of California, Santa Cruz,
United States
Tania Robens,
Rudjer Boskovic Institute, Croatia

*CORRESPONDENCE
Anadi Canepa,
acanepa@fnal.gov
Monica D'Onofrio,
onofrio@liverpool.ac.uk

RECEIVED 08 April 2022
ACCEPTED 15 December 2022
PUBLISHED 22 June 2023

CITATION
Canepa A and D'Onofrio M (2023), Future accelerator projects: new physics at the energy frontier.
Front. Phys. 10:916078.
doi: 10.3389/fphy.2022.916078

COPYRIGHT
© 2023 Canepa and D'Onofrio. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Future accelerator projects: new physics at the energy frontier

Anadi Canepa[1]* and Monica D'Onofrio[2]*

[1]Fermi National Laboratory, Batavia, IL, United States, [2]Department of Physics, University of Liverpool, Liverpool, United Kingdom

High-energy colliders provide direct access to the energy frontier, allowing to search for new physics at scales as high as the machine's center-of-mass energy, perform precision measurements of the Standard Model (SM) parameters, including those related to the flavor sector, and determine the Higgs boson properties and their connection to electroweak symmetry breaking. Each proposed future collider option has its own specific science goals and capabilities, depending on the designed running energy (energies) amongst other parameters. In this paper, an overview of the discovery potential of future circular and linear colliders is presented. Results from searches for beyond the Standard Model (BSM) phenomena at proton–proton, proton–electron, electron–positron, and muon–antimuon colliders are summarized.

KEYWORDS

future, colliders, beyond the Standard Model, resonances, supersymmetry, hidden sector, dark matter, extended Higgs models

1 Introduction

Particle physics advances the fundamental description of "nature" at the smallest scales, leading and influencing global scientific efforts. The Large Hadron Collider (LHC) endeavor remains by far the major focus of the efforts and engagement of the particle physics community. During Run 1 and Run 2 of the LHC, the ATLAS and CMS experiments have produced first observations of fundamental processes, including the discovery of the Higgs boson in 2012 and the determination of its properties and couplings, and hundreds of Standard Model (SM) measurements and searches for new physics. Hints of lepton flavor universality violation (LFV) have been reported by the dedicated LHCb experiment, which also provides improved knowledge of quark mixing matrix parameters, such as the CKM angle, γ, and the discovery of many new hadronic states. Heavy ion studies are also pursued at the LHC by the specialized ALICE experiment, as well as by ATLAS and CMS. During the almost 12 years of LHC operation, there have been many experimental and theoretical advancements: it is remarkable that the precision of SM measurements and the sensitivity reach of new physics searches have exceeded, in some cases by far, the pre-LHC era expectations.

The experimental success of the LHC is certainly a result of the excellent performance of the detectors and the accelerator complex, and its high luminosity upgrade, the HL-LHC, will maximize its potential. Scheduled to run until 2038–2040, the HL-LHC program will allow the general purpose detectors to collect an integrated luminosity of 3 ab^{-1} of proton–proton collisions at a center-of-mass energy (or $\sqrt{s}$) of 14 TeV. The determination of the Higgs boson properties, and their connection to electroweak symmetry breaking (EWSB), is one of the primary science goals of ATLAS and CMS [1]. Outstanding opportunities will emerge for measurements of fundamental

importance, such as the first direct constraints on the Higgs trilinear self-coupling and its natural width. SM processes and parameters, including those related to the flavor sector, will be tested by performing measurements with unprecedented precision [2], such as the production of pairs or triplets of EW gauge bosons, the effective weak mixing angle, and the masses of the top-quark and W-boson. On the latter, it is noticeable that the Large Hadron electron Collider (LHeC) [3], designed to collide an intense electron beam produced by an energy recovery linac with the HL-LHC proton beam, would allow to reduce the uncertainty on this fundamental parameter to 2 MeV. The search for physics beyond the SM (BSM) will remain the main driver of the exploration program at the HL-LHC [4]. In most BSM scenarios, the HL-LHC will increase the present reach in mass and coupling by at least 20%–50% and will allow searches for, among others, additional Higgs bosons in EWSB scenarios, new resonances, candidates for dark matter (DM), and hidden sectors to be performed. The LHCb program will continue to focus on heavy-flavor physics throughout the HL-LHC phase [5], possibly confirming the anomalies in the lepton-flavor sector and, as such, establishing LFV and opening a new era of discoveries. LHCb will also search for feebly interacting new particles arising in hidden sectors, complementing ATLAS, CMS, and other beyond-collider experiments in these searches.

The conclusion of the European Strategy for Particle Physics update (ESPPU) and the ongoing Snowmass and Particle Physics Project Prioritization Panel (P5) process presents an opportunity to revisit the landscape of future accelerators beyond the HL-LHC. Several lepton and hadron collider options have been considered, each with merits and challenges, and emphasis has been given on attainable physics targets as well as technological requirements and drivers.

This review focuses on the proposed accelerators' potential for discovering physics beyond the SM, and it is primarily based on feasibility studies reported in the literature at the time of the ESPPU process. If they were available at the time of writing, updates on searches and measurements submitted through the Snowmass process have been included as well. Considerations regarding the status of the proposed machines are also presented.

2 Overview of proposed collider options

Extensive activities have been carried out worldwide to assess the future of collider experiments beyond the HL-LHC. A summary of the accelerator-based projects proposed by the community in recent years and considered in this paper is presented in Table 1.

Electron–positron colliders (linear or circular) with O (100 GeV) center-of-mass energy such as the Future Circular Collider e^+e^-, FCC-ee [6], the International Linear Collider, ILC [64], and the Compact Linear Collider, CLIC [7] represent primarily a Higgs factory providing unprecedented precision measurements of the Higgs boson properties. At the FCC-ee, this would come in combination with improvements in the knowledge of the SM couplings from the planned Tera-Z facility, also enabling an interesting flavor physics program. At linear colliders, a significant advance in precision is also expected thanks to the available beam polarizations. The e^+e^- machines are also discovery machines, with high center-of-mass energy options such as $CLIC_{3000}$, extending the sensitivity to high-mass phenomena and, in some areas, yielding a reach comparable to high-energy pp colliders. For similar center-of-mass energies and integrated luminosities, the CepC [8] physics program and potential are comparable to those of the FCC-ee, chosen here as a representative case.

The future 100 TeV center-of-mass energy proton–proton collider at the FCC, FCC-hh [9], expected to run after the completion of the FCC-ee stages, offers several unique possibilities for a breakthrough in particle physics. Aiming to collect integrated luminosities up to 20 ab^{-1} per experiment in 10 years of operation, FCC-hh will allow measurements of the Higgs self-coupling at the few percent level, measurements of quartic Higgs self-coupling, and direct searches for new physics at the highest energy, with the possibility to extend by an order of magnitude the LHC sensitivity above the EWSB scale. Precise differential cross-section measurements for high-transverse momentum Higgs boson production and measurements of rare processes will be possible due to the large datasets. With the addition of an energy recovery electron linac of 60 GeV, electron–proton

TABLE 1 Collider options whose physics case is discussed in this paper. Configurations and parameters are those proposed at the time of writing (see Section 10 for more details). In case multiple stages are foreseen for a specific project, all center-of-mass energies and benchmark luminosity values are reported. In the text and figures, the ILC stages are indicated as ILC_{250}, ILC_{500}, and ILC_{1000}; CLIC's three phases are referred to as $CLIC_{380}$, $CLIC_{1500}$, and $CLIC_{3000}$. A similar nomenclature is adopted for FCC-ee and the muon collider (MuC) where relevant. The integrated luminosity is per interaction point (IP), except for the FCC-ee and the CepC, where it corresponds to two IPs.

Collider (type)	$\sqrt{s}$ (GeV) [$\mathcal{L}_{int}$ (ab^{-1}), duration (years)]
HE-LHC (circular, pp)	27×10^3 [15, 20]
ILC (linear, e^+e^-)	91 [0.1, 1.5]; 250 [2, 11]; 350 [0.2, 0.75]; 500 [4, 9]
CLIC (linear, e^+e^-)	380 [1.0, 8], 1.5×10^3 [2.5, 7], 3×10^3 [5, 8]
FCC-ee (circular, e^+e^-)	88–94 [150, 4]; s-channel h [20, 3]; 157–163 [10, 2]; 240 [5, 2]; 340–365 [1.7, 5]
FCC-hh (circular, pp)	100×10^3 [20–30, 25]
FCC-eh (circular plus ERL, ep)	3.5×10^3 [3, 25]
MuC (circular, $\mu^+\mu^-$)	3 TeV [1, 5]; 10 TeV [10, 5]; 10 TeV [20, 5]
CepC (circular, e^+e^-)	91 [16, 2]; 160 [2.6, 1]; 240 [5.6, 7]; 360 [-, -]

interactions could be explored [10], providing additional inputs to achieve the ultimate Higgs physics precision at the FCC-hh, QCD precision measurements, and searches for new physics. The FCC hadron complex would also allow for a dedicated heavy ion program, with lead–lead and proton–lead, and electron–lead collisions possible at the FCC-hh and FCC-eh, respectively.[1] An alternative idea for a *pp* collider considers the possibility of increasing the energy of the LHC machine up to 27–30 TeV, turning the current accelerator into a high-energy machine (HE-LHC [12]). This would allow utilizing the current tunnel and the entire CERN infrastructure with future magnet technologies to collect large datasets at $\sqrt{s}$ at least two times the one of the HL-LHC.

A $\mu^+\mu^-$ collider, MuC [13], could give the opportunity to achieve a multi-TeV energy domain beyond the reach of the e^+e^- colliders and within a much shorter circular tunnel than for a *pp* collider. The picture emerging from studies of the past years is that a 10 TeV muon collider could combine the advantages of *pp* and e^+e^- colliders due to the large $\sqrt{s}$ available for direct exploration and to the achievable accuracy for precise measurements of the SM. By exploiting the copious rate of vector–boson fusion and vector–boson scattering processes, a MuC provides the opportunity to probe details of the EWSB mechanism. Muon-philic new physics scenarios, possibly explaining the $g-2$ [14] and *B*-physics anomalies [15], are additional natural targets. Because a muon production and cooling complex could be used at all energies, and muon acceleration proceeds through a sequence of rings, a $\mu^+\mu^-$ collider can be built in stages, with 3 TeV center-of-mass energy foreseen as the first stage.

Future collider concepts [16] not explicitly listed previously are also being considered within the Snowmass/P5 process. Among those, the C^3 linear collider project [17] could fit on the Fermilab site and would have a similar potential to that of the aforementioned *O* (100 GeV) machines, with a starting center-of-mass energy of 250 GeV to be potentially increased to 550 GeV and to 3 TeV by extending the accelerator's length. Other options based on high-gradient superconducting radio frequency (SRF) technology, such as the compact SRF Higgs-Energy LEptoN (HELEN) linear collider, are also being investigated [18]. In terms of circular machines, a 16-km circumference circular e^+e^- collider with center-of-mass energy between 90 GeV and 240 GeV is being examined with Fermilab as a potential site, as well as a possible proton–proton collider with center-of-mass energy between 24 and 27 TeV located in the same tunnel, with a reach similar to that of a HE-LHC. Finally, a Super proton–proton Collider (SppC [19]) is proposed as a machine located in China, running after the CepC, and using the same tunnel complex and infrastructure in a multi-staged approach similar to that envisaged for the CERN FCC. More details are given in Section 10.

The planning spans a 30-year horizon, as major accelerator-based projects require developments on that timescale. Comparing the physics potentials, the required technology and prospects for its availability, and the cost-to-benefit ratio of the proposed machines is extremely challenging. Each collider program, to varying degrees and dependent in part on the center-of-mass energy considered, gives good coverage of almost all fundamental physics questions. They also have unique synergies with the neutrino and precision frontiers, as well as with astrophysics and cosmological investigations ongoing or planned during the next decades. An overview of those complementarities is, however, beyond the scope of this paper.

3 Searches at colliders: physics landscape

The SM has been proven very successful in describing elementary particles and their interactions. It has been validated extensively through precision experiments, and the discovery of the Higgs boson has certainly been a major milestone in this respect. However, there are a number of shortcomings and several open questions that the SM fails to answer. Severe fine-tuned cancellations of large quantum corrections are required to obtain a Higgs boson mass close to the EW scale, leading to the so-called hierarchy problem. The SM does not incorporate gravity as described by general relativity, or account for the accelerating expansion of the universe. It does not contain any viable dark matter particle and fails to explain in full baryon asymmetry, neutrino oscillations, and non-zero neutrino masses. As such, a plethora of theories beyond the SM have been developed in the past decades, and the search for them is at the core of the particle physics community's experimental activities. While formulating an exhaustive and complete classification of all existing BSM models is not possible, it is evident that the exploration of the unknown is one of the main drivers of all future colliders:

- Important goals of future colliders include searches for the existence of new gauge or space-time symmetries and tests of theories containing multi-TeV resonances. Mostly related to the dynamics of EWSB, vector resonances, leptoquarks, and contact interactions are among the BSM theories considered in this paper. Direct searches for heavy new particles can be complemented by precision studies of SM observables, and deviations from predictions would be an indirect but powerful way to provide evidence of new physics.
- Several new physics models focus on the nature of the Higgs boson, either considering the possibility that it is a composite state or that it belongs to an extended sector with new scalar particles, where one closely resembles the SM Higgs boson. For the latter, various models with different Higgs representations have been proposed. Among those receiving the most attention for future collider studies is the extension of the SM scalar potential by a singlet massive scalar field that can change the nature of the EW phase transition. Another common set of extended Higgs sector models searched for is characterized by the addition of a second SU(2) doublet, which naturally appears in supersymmetric extensions of the

1 We note that the nuclear physics community is also pursuing the Electron-Ion Collider (EIC) project, the first-ever collider of polarized electrons with nuclei or polarized protons. The target center-of-mass energy (20–140 GeV) is substantially smaller than that achieved at HERA, but the target luminosity is foreseen to be 1,000 times higher. Considerable synergies with accelerator particle physics projects in terms of detector technology and physics potential are expected. For a detailed report on the EIC project, see Ref. [11].

Higgs sector or in models with a non-minimal pattern of symmetry breaking.

- Supersymmetry (SUSY) certainly remains one of the most plausible beyond the SM scenarios, as it provides the only known dynamical solution to the Higgs naturalness problem that can be extrapolated up to very high energies. SUSY gives an excellent potential candidate for DM as well as a framework for gauge coupling unification and possibly reconciliation of gravity and other forces. As such, it is the focus of multiple studies from various facilities both in the strong and electroweak sectors.
- Cosmological data suggest that DM particles could have masses in the range from multi-keV to approximately 100 TeV and couplings to SM particles of comparable or weaker strength than EW interactions. High-energy colliders could produce DM particles within this mass range in controlled conditions and, as such, complement experiments and observations from astroparticle physics experiments. A typical DM thermal relic studied at colliders is a weakly interacting massive particle, referred to as WIMP. Several DM models predict the presence of mediator particles, whose exchange may be responsible for the annihilation processes that determine the DM particle abundance and can be directly searched at colliders. If the DM particle is lighter than $m_h/2$ and it is coupled to the Higgs, a compelling exploration channel is an invisible Higgs decay. Of particular interest are the cases of spin-1/2 particles transforming as doublets or triplets under SU(2) symmetry.
- An alternative possibility for new physics is that particles responsible for the still unexplained phenomena have not been detected because they interact too feebly with SM particles. These particles could belong to an entirely new sector, the so-called hidden or dark sector. While masses and interactions of particles in the dark sector are largely unknown, the mass range between the keV and tens of GeV appears interesting, both theoretically and experimentally.
- Heavy new physics can induce, through the exchange of virtual particles, processes that are extremely rare in the SM, such as flavor-changing neutral current (FCNC) effects in the top-quark sector. The expected intensity of some of the future lepton collider proposals at critical production thresholds will allow improvements in sensitivity. Hadron colliders at very high luminosities and muon colliders at multi-TeV center-of-mass energy are also complementary when looking for rare processes.
- In the absence of evidence for new physics at low energy and assuming that BSM is realized at a scale Λ much larger than the collider $\sqrt{s}$, the effective field theory (EFT) formalism is adopted as a framework to study BSM physics with a model-independent approach. Several EFT representations exist, and a subset of those is considered in this review.

This program is continuously evolving and broadening in response to results from the current LHC, the HL-LHC, and other ongoing and future non-collider experiments. Beyond-collider projects in construction, planned, or proposed to further exploit the LHC accelerator complex will significantly boost the discovery potential in the next two decades, offering complementarities and synergies in the quest for new physics. Among those, FASER [20] and SND@LHC [21] will start operations during the Run 3 of the LHC; others like MATHUSLA [22], CODEX-b [23], MilliQan [24], and the LHeC [25] are foreseen to operate in parallel to the HL-LHC. An extensive proposal on the Forward Physics Facility at CERN has been presented recently [26], while Ref. [27] presents detectors for fixed-target experiments and beam-dump experiments at the ILC complex.

A subset of representative prospective results is reported in the rest of this review. They originate from hundreds of studies of varying degrees of sophistication carried out over several years. Detailed or fast simulations are used in some cases, whilst simple detector parameterizations, direct extrapolations of results from existing data, or even simple rescaling are performed in other cases. The reader is referred to the original publications for details on the analyses and on the approach and hypotheses made.

4 New resonances

On-shell resonances decaying into visible SM particles are a distinct signature of several BSM theories, ranging from new models of EWSB to extensions of the SM gauge group. Classic scenarios include singly produced resonances with integer spin or pair-produced heavy resonances. Direct access requires the center-of-mass energy of the collider to be large enough to produce them. Performances can then be evaluated considering the reach in mass, or the reach in mass vs. coupling, with the results depending on the assumptions on the couplings of the new particle to quarks and leptons. If the colliders' center-of-mass energies are below the mass of the new hypothetical resonance, indirect access can be achieved by studying deviations in SM observables.

One of the most widely used benchmark scenarios predicts a new high-mass vector (spin-1) boson, the Z'. Examples [28] are the sequential SM (SSM), $B-L$, and E_6 Z' models, as well as models of little Higgs or extra dimensions. The primary discovery mode for a Z' at hadron colliders is the Drell–Yan production of a dilepton resonance, but hadronic final states are also widely studied. The mass reach is typically in the $(0.3-0.5)\sqrt{s}$ range, given sufficient statistics. FCC-hh [29] could discover a SSM Z' with a mass of up to 43 TeV if it decays into an electron or muon pair, assuming 30 ab^{-1} of luminosity. Masses between 20 and 30 TeV could be reached if decays to $\tau^+\tau^-$ or to $t\bar{t}$ are instead considered. A multi-TeV MuC could become competitive in accessing directly a Z', especially in the case of muon-philic models [30], where the new vector boson dominantly couples to $\mu^+\mu^-$, *e.g.*, via left-handed currents.

Given the current mass limits from the LHC experiments, a direct observation of these new resonances is not expected at the currently planned linear and circular e^+e^- accelerators for most of the scenarios considered in the literature. Nonetheless, the presence of high-mass resonances can be inferred indirectly using an EFT approach to describe BSM virtual effects. In the EPPSU studies, a benchmark model dubbed "Y-Universal Z'" has been used for a quantitative assessment of the potential of future colliders to search for new gauge bosons, directly and indirectly [31]. Figure 1 (left) shows the 95% confidence level (CL) limits in mass vs. coupling at various colliders. The model assumes the same couplings, $g_{Z'}$, to quarks and leptons, and it was chosen because it allows for a fair comparison between hadron and lepton colliders. The direct

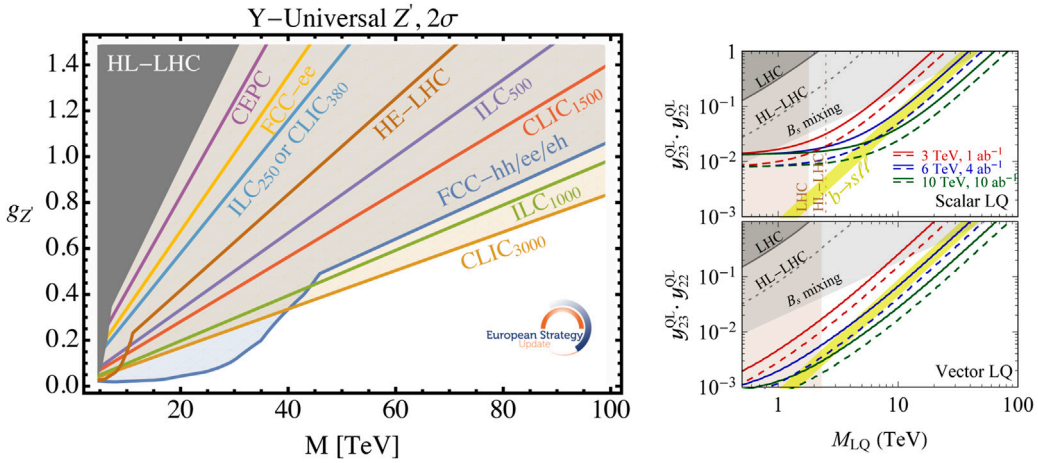


FIGURE 1
Left: Exclusion reach of different colliders on the *Y*-Universal Z' model parameters [31]. Right: Muon collider sensitivity contours at 95% CL for scalar (upper panel) and vector (lower panel) LQ models via the process $\mu^+\mu^- \to bs$, where $y^{QL_{ij}}$ indicates the coupling between the *i*-generation lepton and the *j*-generation quark. For the various setups considered, see Ref. [32].

constraints from FCC-hh are the most stringent at low $g_{Z'}$, while indirect reaches of both e^+e^- and *pp* colliders become superior at high $g_{Z'}$. At lepton colliders, an EFT framework allows to achieve sensitivity thanks to the EW precision measurements of the oblique parameter *S* [33]. At hadron colliders, Drell–Yan predictions are also sensitive to the ratio between $g_{Z'}$ and the Z' mass. As such, very precise parton distribution function (PDF) fits obtained using *ep* collisions would further improve the sensitivity [3]. Finally, the muon collider reach, not reported in this figure, is estimated to be similar to that of CLIC_{3000} for $\sqrt{s} = 3$ TeV and exceeding it proportionally to the increase in center-of-mass energy [34].

Expected sensitivity on the production and decay of spin-0 and -2 particles decaying into several different SM final states has also been studied at high-energy lepton and hadron colliders. Models considered include, among others, resonant double-Higgs production and heavy scalar singlets that could mix with the Higgs boson, i.e., see Refs [35–37].

Leptoquark (LQ) models, alongside Z' models, have gained considerable renewed interest in recent years as they can give rise to lepton universality violating decays of heavy mesons at the tree level, provided that couplings are generation-dependent and they couple to the second and third generations of quarks. LQs are hypothetical particles that carry both baryon and lepton quantum numbers. They are color-triplets and carry fractional electric charge. The spin of a LQ state is either 0 (scalar) or 1 (vector). Models predicting a rather light LQ coupled predominantly to the third generation are a natural target for hadron colliders where scalar or vector LQs are pair-produced via strong interaction and results [31] are independent of the coupling to the lepton quark current. If discovered, FCC-eh could contribute to their characterization, assuming that the coupling to the first-generation quark is non-negligible and can be produced as an *s*-channel resonance [3]. Muon colliders have the best sensitivity for a LQ model via $\mu^+\mu^- \to bs$. With a few to 10 TeV center-of-mass energy and predicted luminosities of 1–10 ab^{-1}, a MuC could cover the entire parameter space that explains the flavor anomalies for both scalar and vector LQ. Results are shown in Figure 1 (right), from Ref. [32].

If new particles arising in BSM theories are much heavier than the energy reach for on-shell production even at future colliders, their existence can still be formalized through contact interactions (CIs). An effective four-fermion CI could represent the exchange of a virtual heavy particle, such as an LQ, a Z', or elementary constituents of quarks and leptons in composite models. The effective CI scale represents the typical mass scale of the new particles, and the experimental sensitivity increases significantly with $\sqrt{s}$. Lepton colliders are powerful in testing the neutral-current case, owing to the precision that can be achieved in analyses of di-fermion final states with suitable statistics. Linear colliders can also exploit different longitudinal polarizations of the two beams. Hadron colliders have excellent sensitivity up to their $\sqrt{s}$ via Drell–Yan production for both neutral and charged currents. The highest reach as reported in the EPSSU studies [31] is up to 120 TeV (CLIC_{3000}). The so-called two-fermion/two-boson CIs are also phenomenologically relevant for BSM theories of EWSB because they describe new physics effects in the interaction between the gauge and Higgs sectors. In this case, estimated reaches [31] are, at best, 30–35 TeV. Precision differential measurements of the *ZH* production provide the lead sensitivity for lepton colliders. Hadron colliders' sensitivity mostly comes from precision measurements of SM diboson production observables, as used in the FCC-hh studies. Additional studies on CIs related to new physics models possibly contributing to the muon $g-2$ and to high-energy scattering processes have also been carried out at the muon collider [30].

5 Composite Higgs and extended sectors

The role of the Higgs boson could be even more complex than that known so far in the SM formulation, and hence it is logical to also question its nature and whether or not it is a point-like particle. Composite Higgs models (CHMs) predict that the Higgs is not an

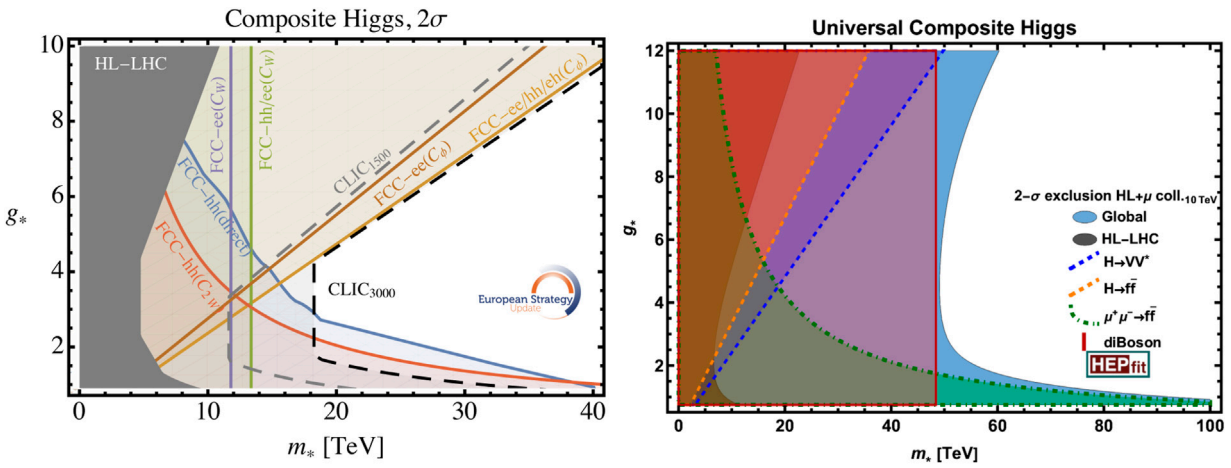


FIGURE 2
Left: Exclusion reach on the composite Higgs model parameters of FCC-hh, FCC-ee, and of the high-energy stages of CLIC [31]. Right: Comparison of the global reach for universal composite Higgs models at the HL-LHC and at a 10 TeV muon collider. For the latter, contributions to the constraints from different processes are also shown [30].

elementary particle and that new particles might arise as excitations of the composite Higgs, with mass possibly at the O(TeV) scale. The foundation of composite Higgs models is that the Higgs emerges as a bound state of a new strongly interacting confining composite sector, analogous to QCD, but with a much higher confinement scale. The Higgs, similarly to the pions in QCD, emerges as a Goldstone boson associated with a spontaneously broken global symmetry of the composite sector. The phenomenology of CHMs is mainly controlled by two parameters: the mass scale $m_\star$, which controls the mass of the new resonances, and the coupling $g_\star$, representing the interaction strength among particles originating from the composite sector. EFT operators that describe the indirect effects of Higgs compositeness at low energy are then defined, and their scale is set by $g_\star$.

Figure 2 (left) shows the exclusion reach on $m_\star$ and $g_\star$ for FCC-hh, FCC-ee, and the high-energy stages of CLIC. Contours for the reach of HE-LHC, ILC, CepC, and CLIC_{380} are also available in Ref. [31]. The 95% CL exclusion contours of each collider project arise from effects on coupling measurements [38] of the Higgs boson related to its possible composite nature and, for the FCC-hh and the HL-LHC, also from direct searches for an EW triplet ρ vector resonance in dilepton and diboson final states [43]. Figure 2 (right) presents the exclusion reach for the 10 TeV stage of the muon collider. The reach of HL-LHC is also reported and statistically combined in the global result [30]. Other curves denote the contributions to the constraints from different processes, including that of searches for composite Higgs fermionic top partners.

Theories predicting an extended Higgs sector acquired significant prominence in the experimental programs of collider experiments, with searches targeting a broad spectrum of models. In minimal scenarios, the Higgs sector is augmented by a singlet massive scalar field which, e.g., can mix with the SM Higgs boson with a mixing parameter γ. The presence of the singlet can either modify the SM Higgs boson properties or be detected as single production of the massive particle associated with the field, S, which subsequently decays into SM particles. Figure 3 (left) summarizes the reach [31], in the mass-$\sin^2\gamma$ space, of direct searches and indirect constraints derived from the Higgs boson couplings measurements (horizontal lines). Among the indirect searches, those performed at CLIC_{3000} are the most sensitive searches and allow to probe mixing angles for values as low as $\sin^2\gamma \sim 10^{-3}$ for any value of m_S. In contrast, the reach of direct searches depends on the singlet's mass. The muon collider at $\sqrt{s}$ = 14 TeV explores masses as high as 9 TeV while extending the sensitivity to $\sin^2\gamma$ by almost one order of magnitude with respect to the best indirect constraint from e^+e^- colliders. Thanks to the larger center-of-mass energy, the FCC-hh is sensitive to higher masses but yields a more limited reach in the mixing values.

Under the no-mixing assumption, the singlet-associated particle S would be stable and thus searched for in events with significant missing transverse momentum. The best sensitivity is yielded by indirect searches based on the precision measurement of the SM Higgs couplings at CLIC_{3000}, probing masses between 50 and 350 GeV and λ_{HS} between 0.1 and 1, where λ_{HS} is the coupling term in the potential $V \sim \lambda_{HS}|H|^2S^2$ [31]. Experiments at the FCC-hh achieve a similar sensitivity through direct searches for the pair production of S. It is interesting to note that this region of phase space is compatible with a strong first-order EW phase transition, demonstrating that colliders have the potential to test models predicting the baryon asymmetry in the universe and gravitational waves. As such, the energy frontier complements the program at cosmology experiments, like the future gravitational wave experiment LISA. Another example is presented in Ref. [30], where the reach of a 3 TeV muon collider is compared to that of LISA.

More complex scenarios extending the Higgs sector by a new SU [2] doublet, e.g., supersymmetry or more generically type-II two-Higgs doublet models, predict the existence of two CP-even scalars, h and H, one CP-odd scalar, A, and a charged scalar, $H^\pm$. This rich phenomenology leads to a variety of probes at future machines. As an example, Figure 3 (right) shows the constraints on m_A as a

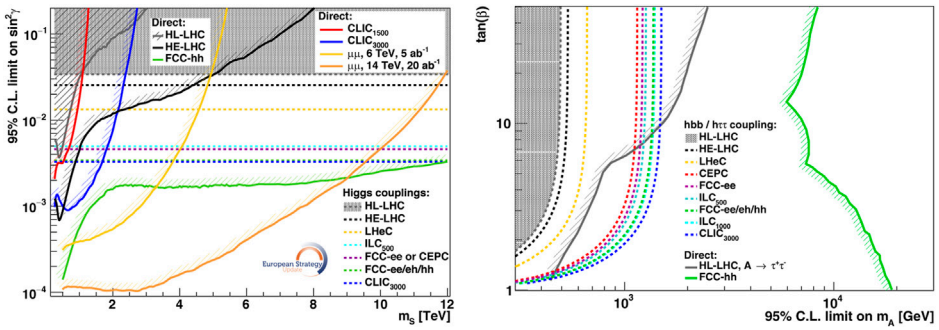


FIGURE 3
Exclusion reach obtained from the precision measurements of the Higgs boson couplings and from direct searches for new states. Left: Sensitivity in the plane $\sin^2\gamma - m_S$, where γ and m_S are the mixing angle and the mass of the singlet mixing with the SM Higgs boson, respectively. Right: Sensitivity in the plane $\tan\beta - m_A$, where $\tan\beta$ is the ratio of the vacuum expectation value for the two Higgs doublets and m_A is the mass of the CP-odd scalar Higgs boson, in a type-II two-Higgs doublet model [31].

function of $\tan\beta$, the ratio of the vacuum expectation value for the two Higgs doublets. While precision measurements of the Higgs couplings to third-generation fermions offer sensitivity to models at low values of m_A (~1 TeV), the FCC-hh probes the existence of the new states for masses as high as 10–20 TeV for any values of $\tan\beta$. Assuming efficient signal detection and background suppression, the sensitivity to the pair production of new states at lepton colliders goes up to $m \leq \sqrt{s}/2$. Measurements of flavor physics observables can also lead to constraints on the type-II two-Higgs doublet models (i.e., see Ref. [39]).

6 Supersymmetry

The phenomenology of SUSY is mostly driven by its breaking mechanism and breaking scale, which define the SUSY particle masses, the mass hierarchy, the field contents of physical particles, and thus their cross sections and decay modes. In addition, signal topologies crucially depend on whether R-parity, defined as $R = (-1)^{3(B-L)+2S}$, where B and L are baryon and lepton numbers, respectively, and S is the spin, is conserved or violated.

Indirect constraints from flavor physics experiments, high-precision electroweak observables, including the discovery of the 125 GeV Higgs boson, and astrophysical data impose strong constraints on the allowed SUSY parameter space. Still, SUSY can be the key to understand Higgs naturalness, and in R-parity conserving scenarios, the lightest supersymmetric particle (LSP) is an excellent candidate for DM. These are certainly strong motivations to search for colored SUSY particles, squarks, and gluinos, for EW gauginos and Higgsinos that mix into neutralino and chargino mass states, collectively referred to as electroweakinos (EWkinos, χ), and for the superpartner of charged and neutral leptons, the sleptons.

Squarks and gluinos are produced via the strong interaction and have the highest cross sections at hadron colliders. Scalar partners of the left-handed and right-handed chiral components of the bottom-quark and top-quark mix to form mass eigenstates for which the bottom and top squarks are defined as the lighter of the two ($\tilde{b}_1$ and $\tilde{t}_1$, respectively) and might be significantly lighter than the other

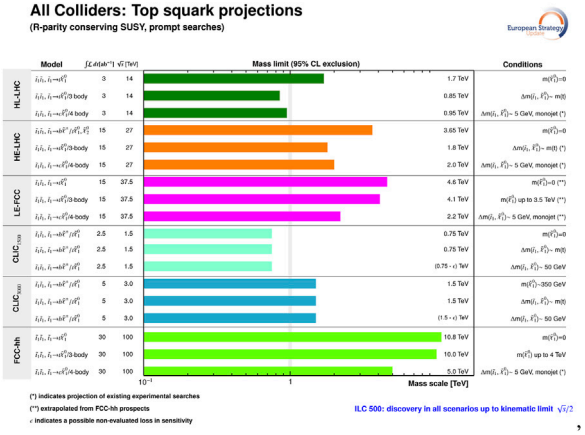


FIGURE 4
Exclusion reach top squark at different hadron and lepton colliders [31].

squarks and the gluinos. EWkinos cross sections depend on mixing parameters and are typically much smaller than those of colored superpartners at hadron colliders. For this reason, the EW sector remains more difficult to test at hadron machines, and searches at e^+e^- colliders would complement the SUSY parameter space coverage. Similar considerations can be made for slepton pair production, as cross sections can be up to two orders of magnitude smaller than those for EWino pair production.

Prospects for SUSY searches are presented in terms of mass exclusion limits at 95% CL. The corresponding definitive observation with a significance of 5σ would be 5%–10% lower depending on the process. High-energy pp colliders provide the most stringent bounds on first- and second-generation squarks and gluinos. In R-parity conserving scenarios, gluino (squark) masses up to 17 [9] TeV could be reached by the FCC-hh exploiting the typical multijet plus missing transverse momentum SUSY signature for a massless LSP, while monojet-like analyses, where the SUSY particles recoil against an initial state radiation (ISR) jet, are most effective for

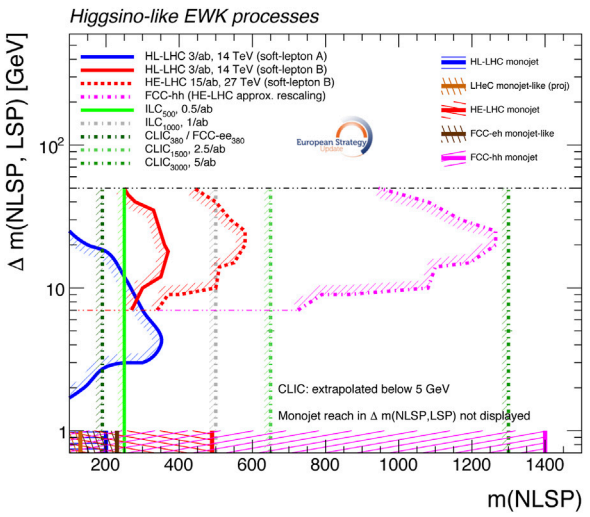


FIGURE 5
Exclusion reach at 95% CL for Higgsino-like charginos and next-to-lightest neutralinos with equal mass m (NLSP) as a function of the mass difference Δm between NLSP and LSP. Exclusion reaches using monojet searches at pp and ep colliders are also superimposed [31].

compressed scenarios. Lepton colliders are ineffective in the searches for gluinos, which are neutral with respect to the EW interaction, while current limits on first- and second-generation squark masses make the results not competitive. Similar conclusions can be drawn for top-squark pair-production searches if the preferred decay is $\tilde{t}_1 \rightarrow t\tilde{\chi}_1^0$ and $\Delta m(\tilde{t}, \tilde{\chi}_1^0) \gg m_t$, where $\tilde{\chi}_1^0$ is the lightest neutralino and m_t is the top-quark mass. On the other hand, for small mass splitting, the sensitivity of pp colliders significantly degrades so that high-energy lepton colliders, e.g., CLIC_{3000} and MuC at 3–10 TeV, become competitive [34]. Their stop mass reach is close to $\sqrt{s}/2$ even for low $\Delta m(\tilde{t}, \tilde{\chi}_1^0)$, although a loss in acceptance and efficiency could be expected for mass differences of the order of 50 GeV. The exclusion limits are summarized in Figure 4; see [31] and references therein for details on the assumptions.

At hadron colliders, the largest production rates for EWkinos are obtained when the lightest chargino ($\tilde{\chi}_1^{\pm}$) and next-to-lightest neutralino ($\tilde{\chi}_2^0$) are Wino-like, forming an approximately mass degenerate SU(2) triplet referred to as the next-LSP (NLSP). Exclusion reaches for R-parity conserving Wino-like scenarios show that NLSP masses up to 3.3 TeV can be excluded at FCC-hh for a massless LSP, to be compared to a sensitivity of up to EWkino masses equal to $\sqrt{s}/2$ for high-energy e^+e^- and $\mu^+\mu^-$ colliders even for $\Delta m(\tilde{\chi}_1^{\pm}, \chi_1^0)$ as low as 1 GeV, with no loss in acceptance. If the Higgsino mass is much smaller than the gaugino masses, the production rates are smaller, $\tilde{\chi}_{1,2}^0$ and $\tilde{\chi}_1^{\pm}$ form an approximately mass degenerate SU(2) doublet, and the EWkino spectrum is compressed. Feasibility studies and projections are summarized in Figure 5 (for details, see [31] and references therein). The sensitivity of lepton colliders depends only weakly on the nature of the LSP as cross sections are less dependent on the choice of mixing parameters. The high-energy lepton colliders allow a reach close to the pair production threshold, approximately 1.3 TeV for CLIC_{3000}, with the mass splitting down to approximately 0.5 GeV, and it is similar (not shown) for a MuC of 3 TeV center-of-mass energy. Beam polarization effects might also play a crucial role in searches for EWK SUSY at linear colliders [40, 41]. Analyses exploiting ISR jets and/or soft-momentum leptons show good prospects at hadron colliders in the case of Higgsino-like scenarios: $\tilde{\chi}_1^{\pm}, \tilde{\chi}_2^0$ masses up to approximately 600 GeV can be probed at the HE-LHC for mass splittings $\Delta m \equiv \Delta m(\tilde{\chi}_2^0, \tilde{\chi}_1^0) \approx \Delta m(\tilde{\chi}_1^{\pm}, \tilde{\chi}_1^0)$ between 7 and 50 GeV. FCC-hh projections show expected 95% CL limits up to 1.3 TeV, also depending on Δm, with monojet searches possibly complementing the reach for very compressed scenarios. Prospects for ep colliders (LHeC and FCC-eh) performed using monojet-like signatures are also shown. Finally, if the lightest neutralino is either pure Higgsino or Wino, EWinos' mass splittings are theoretically calculated to be approximately 340 MeV and 160 MeV, respectively. In these cases, taking advantage of the long lifetime of the charginos, searches for disappearing charged tracks can be performed at hadron and electron–hadron colliders. Analyses exploiting displaced decays of the charged SUSY state have also been studied for lepton colliders. Results can be interpreted in the context of generic DM models and are reported in Section 7.

Significant sensitivity to sleptons is expected at future accelerators. High-mass selectrons, smuons, and staus are best accessed by hadron colliders for large mass splitting between the slepton and the LSP masses, with limits up to or in excess of 5 TeV for the FCC-hh [42], significantly depending on the assumptions of slepton handedness and mass degeneracy. Dedicated searches for staus, on the other hand, might be particularly challenging at pp colliders because of the potentially high rate of misidentified tau leptons. The HE-LHC would provide sensitivity of up to 1.1 TeV, and an additional three-fold increase is expected for the FCC-hh [31]. Prospect studies at linear lepton colliders [43–45] show excellent expected sensitivity to slepton masses up to close to $\sqrt{s}/2$ and good potential for characterizing the nature of the new particles in case of discovery by exploiting beam polarization. The SUSY EW sector, comprising sleptons, can also account for the long-standing discrepancy of $(g-2)\mu$. Feasibility studies focusing on the relevant parameter region have been reported in the past year [46], showing good complementarity between HL-LHC and high-energy electron–positron colliders. Sensitivity to staus at lepton colliders would again be complementary to pp colliders in case of compressed scenarios, with substantial dependency on the assumptions on $\tilde{\tau}$ handedness and the beam polarization conditions. A multi-TeV muon collider would push the sensitivity up to half the center-of mass energy [34].

A systematic study of the potential of lepton and hadron colliders for R-parity violating (RPV) SUSY scenarios has not been attempted. RPV models might lead to very diverse signatures depending on which couplings are different from 0 and on their strength. The lightest neutralino, as LSP, would decay into SM particles so that final state events present high lepton and/or jet multiplicities and modest or no missing transverse momentum. If RPV couplings are small, particles might travel macroscopic distances before decaying and be long-lived. Searches for high-mass long-lived particles (e.g., gluinos and top squarks) at high-energy pp colliders can exploit the capability of reconstructing unconventional signatures such as massive displaced

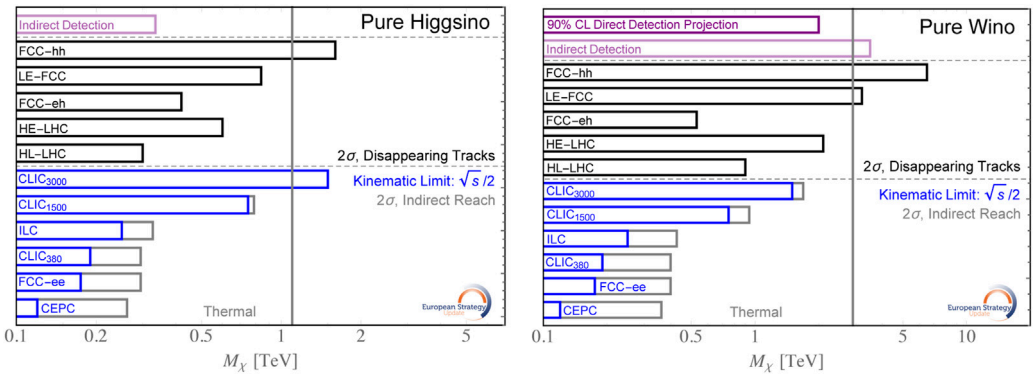


FIGURE 6
Exclusion reach for pure Higgsino- and Wino-like DM candidates at future colliders. Current indirect DM detection constraints and projections for future direct DM detection are also indicated. The vertical line shows the mass corresponding to the DM thermal relic [31].

vertices or displaced leptons similarly to current and planned searches at the LHC and HL-LHC, respectively. Similar techniques can be adopted at lepton colliders for EWinos. To illustrate the potential to discover low-mass SUSY particles decaying non-promptly and as such eluding LHC constraints, interesting prospect studies have been made for O (100 GeV) center-of-mass energy e^+e^- colliders [47] and at *ep* colliders [48].

7 Dark matter

Collider experiments could be sensitive to many possible thermal freeze-out scenarios which assume that DM was produced through interactions of unknown nature with SM particles in the early universe. If DM is part of an extended BSM sector and carries SM charges, a mechanism should exist, as in SUSY, to make it stable or very weakly interactive (WIMP). In case of DM being part of a richer hidden sector, several new massive particles might arise, and one or more could mix with SM particles. A hidden sector that contains DM is generically called a dark sector and can be connected to the SM by small but non-zero couplings through a mediator. The operator interacting or mixing with it is referred to as a portal.

Depending on the model assumptions, the nature of DM and the new-physics phenomenology could be profoundly different. For instance, dark sectors might be characterized by an abundance of feebly interacting particles. Feasibility studies on these scenarios are reported in Section 8.

WIMP DM is invisible to detectors due to the weak strength of its interaction with SM particles, and hence the main signature at colliders is the missing transverse momentum carried by the DM particle. Consequently, searches focus on the associated production of the undetectable DM with visible SM particle(s) like one (or more) jet(s), a *Z* boson, a photon, or a Higgs boson. Additional BSM mediators can lead to a variety of even more complex collider signatures in visible channels, i.e., involving heavy-flavor quarks.

A straightforward model of DM thermal relic is that of a massive particle with EW gauge interactions only. The case of spin-1/2 particles transforming as doublets or triplets under SU(2) symmetry is considered an excellent benchmark model for future colliders. The production rate of the charged state in the DM EW multiplet is high, but it decays into the invisible DM plus a soft undetectable pion. The sensitivity to these models, usually referred to as Higgsino and Wino, respectively, is summarized in Figure 6.

The direct reaches through the so-called disappearing track analyses are compared with indirect reaches at lepton colliders, derived from the sensitivity to the EW parameters *W* and *Y*. FCC-hh can conclusively test the hypothesis of thermal DM for both the Higgsino and Wino scenarios, while CLIC_{3000} could cover in full the Higgsino case. A 3 TeV muon collider would reach masses slightly lower than CLIC_{3000} for the Wino case, while a 10 TeV machine would be competitive with the FCC-hh [34]. As usual, several caveats must be considered when comparing these projections. For instance, projections for future direct DM detection might suffer from uncertainties on the Wino-nucleon cross section, whilst indirect constraints might suffer from unknown halo-modeling uncertainties. More details can be found in Ref. [31] and references therein.

If DM belongs to a richer BSM sector, the phenomenology might be very diverse. Simplified models are therefore used as benchmarks for collider searches to minimize the number of unknown parameters: a single mediator is introduced, which is either a new BSM particle or a SM particle such as the Higgs boson or the Z boson. In the models considered by the EPSSU studies, based on widely accepted benchmark proposals [49], the DM particle is a massive Dirac fermion (χ), and the mediator is either a spin-1 (axial-vector) or a spin-0 (scalar) BSM particle. Figure 7 (left) reports the 2σ sensitivity on the mediator mass of collider experiments for axial-vector models. Results are strongly dependent on the choice of couplings (indicated in the figure), and hence it is difficult to compare among accelerator projects.

The sensitivity at *pp* colliders is driven by dijet and monojet searches, which decreases if couplings to quark decrease. Lepton colliders might reach reasonably high mediator masses through mono-photon analyses, so the achievable sensitivity depends on the strength of the mediator coupling to leptons. Similar results to those of CLIC_{3000} can be achieved by a MuC of the same center-of-mass energy, with sensitivity provided by mono-photon and mono-*W*-boson analyses [30].

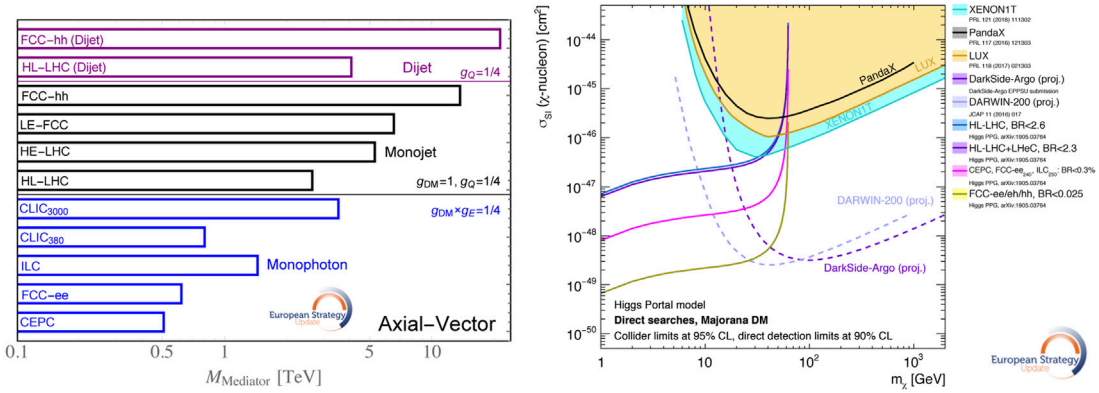


FIGURE 7
Left: Exclusion reach at 95% CL for axial-vector simplified models at future colliders assuming a DM mass of M_{DM} =1 GeV. Right: Results from searches for invisible decays of the Higgs boson, assuming a Majorana-type DM, compared to constraints of current and future direct detection experiments on the spin-independent WIMP–nucleon scattering cross section [31]. The region where the irreducible background from neutrinos may hide a DM signal, usually labeled as the "neutrino floor," is not reported. For further reading, a recent paper on the subject is given in Ref. [50].

Mediators could also be SM particles, and a remarkable example is given by models where the Higgs boson acts as a mediator (or portal). If the DM mass is below half of the mass of the Higgs boson, the latter could decay into a DM pair. As such, precision measurements of the branching ratio (BR) of the Higgs boson decaying into invisible particles can be turned into exclusion limits on the spin-independent WIMP–nucleon scattering cross section. This is illustrated in Figure 7 (right): 90% CL limits for a simplified model with the Higgs boson decaying to Majorana DM particles are compared to current and future DM direct detection experiments. Low-energy e^+e^- colliders are particularly competitive in this scenario, thanks to unprecedented precision expected in measuring Higgs couplings, whilst hadron colliders remain competitive thanks to the large datasets and high production rates.

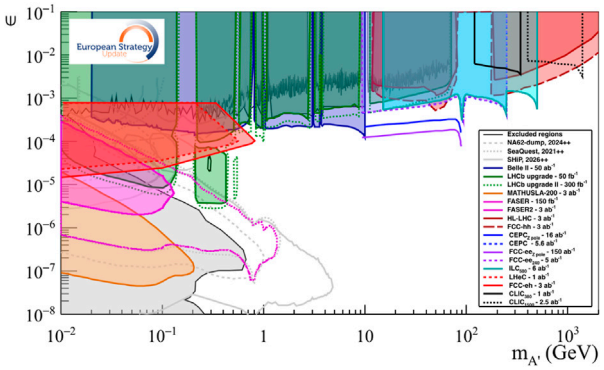


FIGURE 8
Exclusion reach for dark photons at various experiments as a function of the dark photon mass m_A and the mixing parameter between the dark and SM photon, ϵ. Exclusion limits are computed at 95% CL in the case of FCC-ee, FCC-hh, and CepC, while the contour for the FC-eh corresponds to the observation of 10 signal events. All other limits are computed at 90% CL [31].

8 Feebly interacting particles

BSM theories extending the SM with a hidden sector populated by feebly interacting particles (or FIPs) are gaining significant attention as they can provide, depending on the model's implementation, an explanation for the origin of neutrino masses, matter–antimatter asymmetry in the universe, and cosmological inflation, as well as insights into the EWK hierarchy and the strong CP problem. A comprehensive overview of the vast program at both current and future collider-based, fixed-target, and beam-dump experiments can be found in Refs [31, 51]. In this review, the focus is on the minimal portal framework introduced in the aforementioned references. In these models, the FIPs, which are not charged under the SM gauge groups, interact with the SM through portals that can be classified based on the type and dimension of the mediator. The most studied cases, listed as follows according to the operator's spin, are the vector, Higgs, axion, and neutrino portals:

where $F'_{\mu\nu}$ is the field strength for the dark vector, which mixes with the hypercharge field strength $B^{\mu\nu}$; S (sometimes referred to as ϕ) is the *dark Higgs*, a new scalar singlet that couples to the SM Higgs doublet; a is the axion or axion-like particle that couples to gauge and fermion fields (pseudoscalar portal, where f_a is the axion decay constant); and N is a *heavy neutral lepton* (HNL) that couples to the SM left-handed leptons.

In the minimal vector portal, the interaction between the SM and the hidden sector takes the form of a kinetic mixing between one dark and one visible Abelian gauge boson. In selected realizations of the portal, the new U (1) gauge boson in the hidden sector is a dark photon, A', either massive or massless, with ϵ being the mixing coupling parameter between the dark and ordinary photon. Figure 8 presents the sensitivity of various experiments, demonstrating that future colliders can probe the MeV to TeV mass region, compatible with the hypothesis of DM as a thermal relic. Through searches for Drell–Yan production, $pp \rightarrow A' \rightarrow \ell^+\ell^-$, high-energy hadron colliders explore scenarios with large couplings and heavy dark

Portal	Coupling
Vector (dark vector, A_μ)	$-\frac{\epsilon}{2\cos\theta_W} F'_{\mu\nu} B^{\mu\nu}$
Scalar (dark Higgs, S)	$(\mu S + \lambda_{HS} S^2) H^\dagger H$
Pseudo-scalar (axion, a)	$\frac{a}{f_a} F_{\mu\nu} \tilde{F}^{\mu\nu}$, $\frac{a}{f_a} G_{i,\mu\nu} \tilde{G}_i^{\mu\nu}$, $\frac{\partial_\mu a}{f_a} \bar{\psi} \gamma^\mu \gamma^5 \psi$
Fermion (sterile neutrino, N)	$y_N LHN$

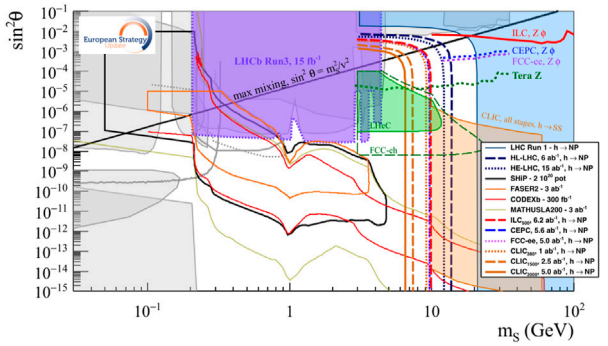


FIGURE 9
Exclusion reach in the m_S – $\sin^2\theta$ plane at various experiments. The symbols S and ϕ are used interchangeably to identify the singlet. Exclusion limits are computed at 95% CL in the case of FCC-ee, FCC-hh, ILC, CLIC, and CepC, while all other limits are computed at 90% CL [31].

photons, with the FCC-hh yielding a sensitivity to ϵ as small as 10^{-3} for masses in the 10–100 GeV range [52] and thus complementing the reach of future LHCb upgrades [4]. An integrated program of precision measurements of the Z boson properties and of direct searches exploiting the radiative return processes ($e^+e^- \rightarrow \gamma A'$) enables electron–positron colliders to significantly extend the sensitivity to lower couplings and lower masses, with FCC-ee probing couplings close to $\epsilon \sim 10^{-4}$ for dark photon masses below the Z mass. For higher masses, the sensitivity of circular and linear e^+e^- colliders with similar luminosities is comparable. Searches for long-lived dark photons produced in $ep \rightarrow eA'$ and decaying into two charged fermions enable the FCC-eh to probe for masses below 1 GeV and couplings in between 10^{-5} and 10^{-3} [53], filling the gap between LHCb, future e^+e^- and pp colliders, and low-energy experiments. Recent results from searches at a muon collider are presented in Ref. [30], while Ref. [54] offers a comprehensive review on more general dark-photon models and corresponding searches. It is noted that hadron and lepton colliders could offer significant sensitivity to non-minimal models where dark photons are produced through BSM Higgs decays, as shown for the HL-LHC in Ref. [4]. A detailed discussion of the sensitivity to non-minimal scenarios is, however, outside the scope of this review.

The minimal dark Higgs model originates from the extension of the SM Higgs sector by the addition of a scalar singlet which mediates the interaction between the SM Higgs boson and the dark sector. In the context of general extended Higgs models, the Lagrangian contains a term proportional to $\sin\theta$ (referred to as $\sin\gamma$ in Section 5), enabling the mixing between the SM Higgs and the new particle S associated with the singlet field, with mixing angles θ. The Lagrangian also contains a term proportional to $\lambda_{HS}S^2$, leading to the coupling between the h and two S particles. If either of the couplings $\sin\theta$ or λ_{HS} is non-zero, a rich phenomenology is expected. If the new scalar mixes with the SM Higgs boson, S can either be produced like a SM Higgs boson or originate from exotic decays of the SM Higgs boson. The corresponding cross sections and branching fractions would depend on the mixing angle. Once produced, the new scalar could decay like a SM Higgs boson, with probabilities reduced by $\sin\theta$, and into the SM Higgs boson itself if $m_S > 2m_h$. In the no-mixing scenarios, S can only be pair-produced through an off-shell or on-shell Higgs boson. The new scalar is stable in the no-mixing minimal models, leading to signatures with missing transverse momentum. The region of parameter space with larger couplings, $\sin^2\theta \geq 10^{-4}$, is explored by searches for the associated production of ZS conducted using the recoil technique at e^+e^- colliders. The Tera-Z configuration of the FCC-ee extends the reach in couplings by one order of magnitude for masses in between a few GeV and half of the Z boson mass by exploring the exotic decays $Z \rightarrow \ell^+\ell^- S$. Precision measurements of the Higgs couplings place constraints on the mass over a large range of $\sin^2\theta$ values: for a fixed luminosity, e^+e^- colliders yield a better sensitivity than those proton–proton machines included in this study, with CLIC_{3000} covering masses as low as 6 GeV for $\sin^2\theta \leq 10^{-5}$. Searches for $h \rightarrow SS$ in visible final states at the FCC-eh allow the experiments to probe intermediate values of $\sin^2\theta$ for masses m_S between 3 and 30 GeV, while similar analyses at CLIC extend the sensitivity to lower values of the couplings for masses between 10 and 60 GeV. These searches assume $\lambda_{HS} \sim 10^{-3}$, corresponding to the level of precision on the measurements on the SM Higgs coupling expected at future e^+e^- colliders. A summary is presented in Figure 9, which also includes the relation between the relevant parameters under the maximal mixing assumption in relaxion models as they exhibit a similar phenomenology via relaxion-Higgs mixing.

Axion-like particles (ALPs, a) are gauge-singlet pseudoscalar particles with derivative couplings to the SM. ALPs can mediate the interactions between the SM and the hidden sector by coupling to photons, gluons, W and Z bosons, and fermions. The interactions with the Higgs boson are suppressed since there is no dimension-5 operator at the tree level in the models considered here. At high-energy colliders, ALPs emerge from either resonant production or from exotic decays of the Z or Higgs bosons ($Z \rightarrow a\gamma, h \rightarrow aZ, aa$). In addition, they can be produced via vector–boson fusion at pp colliders and in association with a gauge or Higgs boson at lepton colliders ($e^+e^- \rightarrow aX$ with $X = \gamma, Z, h$). In ep machines, the incoming electron interacts with a photon from the proton, leading to $e^-\gamma \rightarrow e^-a$. For ALP masses, m_a, below the Z mass, the dominant decay modes are into gluons and photons, where the latter has received the most attention to date. Results from recent searches are therefore presented as a function of the ALP mass and coupling to photons (Figure 10). Thanks to excellent sensitivity to the process $e^+e^- \rightarrow Z \rightarrow a\gamma(\gamma\gamma)$, the Tera-Z configuration of the FCC-ee reaches the best sensitivity for ALP masses between the ~1 GeV and the Z mass, probing couplings $g_{a\gamma\gamma}$ as small as 10^{-8}. Searches for the same rare decay at the FCC-hh probe have a similar mass range but with somewhat limited coverage in couplings, as expected. On the other hand, hadron colliders offer excellent sensitivity to scenarios where,

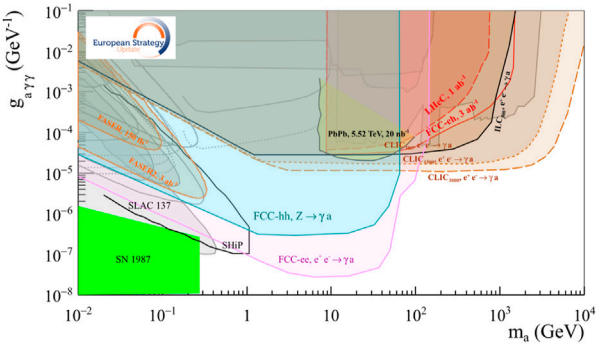


FIGURE 10
Exclusion reach for axion-like particle portal models in the m_a – $g_{a\gamma\gamma}$ plane at various experiments. Exclusion limits are computed at 90% CL, except for the FCC-eh, computed at 95% CL, the FCC-ee, reporting the observation of four signal events, and FCC-hh, reporting that of 100 signal events [31].

e.g., the ALP originates from rare Higgs decays [55]. The e^+e^- linear colliders extend the reach at larger masses thanks to their higher center-of-mass energy and probe couplings as small as 10^{-5}. Experiments at ep colliders have a reach similar to those of low-energy e^+e^- linear colliders by searching for the existence of $e\gamma \to ea$. A detailed overview of the subject, comparing the reach at various machines, is provided in Ref. [55], while Ref. [56] goes into the details of searches at the FCC-ee, exploring all combinations of ALP production modes with visible and invisible decay modes, including those associated with long-lived ALPs. Recent studies at $\sqrt{s} = 10$ TeV muon collider, performed using a modified nomenclature, indicate a discovery potential up to an effective energy scale Λ of 238 TeV, where Λ controls the strength of the interactions [30].

Heavy neutral leptons (also referred to as heavy neutrinos or sterile neutrinos) are one of the most promising extensions of the SM to generate the light neutrino masses observed in neutrino oscillation experiments. At colliders, HNL can emerge from leptonic decays of the W, Z, and Higgs bosons with a probability proportional to the mixing with the SM neutrinos, where the mixing angles and their magnitude can be expressed as $\theta_i = \frac{y^*_{\nu_i}}{\sqrt{2}} \frac{VEV}{M}$ and $|\Theta^2| = \sum_i \theta^2$ ($y^*_{\nu_i}$ are the Yukawa couplings, while M is the sterile neutrino degenerate mass, and VEV is the vacuum expectation value). Additional production mechanisms include the W-exchange t-channel at both e^+e^- and ep colliders ($eq \to Nq$; $e^+e^- \to N\nu$) and γW fusion at e^+e^- machines ($e\gamma \to NW$). Subsequent decays of N occur via emission of a W, Z, h boson, if kinematically allowed. Depending on the value of the couplings and masses, the decay may be prompted or delayed. This rich phenomenology [57] offers opportunities for both direct searches for these new states, e.g., $Z \to N\nu$ or $W \to N\ell$, as well as indirect searches through precision measurements of the gauge and Higgs bosons' properties. The $h \to WW$ channel, e.g., is used to place indirect constraints on $h \to N\nu$. Figure 11 (left) shows that the best sensitivity to the mixing parameter between the electron neutrino and HNL in the region between a few GeVs and the Z mass is yielded by a combination of the conventional and of the displaced-vertex searches performed at the FCC-ee at the Z-pole. For larger masses, FCC-eh provides the best sensitivity for couplings as small as 10^{-6} through searches for lepton-flavor-violating decays. Further studies in Ref. [57] are presented in Figure 11 (right), where $\theta_e = \theta_\mu = \theta_i$, $\theta_i \neq 0$, and $\theta_\tau = 0$. In these models, the sensitivity of indirect searches pushes the sensitivity of the FCC-ee to the TeV scale. Muon colliders could complement the FCC-ee capability in hierarchical scenarios where the mixing to the second generation is dominant. Finally, the recent work in Ref. [56] provides in-depth considerations about the reach of searches for long-lived HNL, the potential to discriminate between the Dirac and Majorana hypotheses, to measure the mass, and to probe regions of parameter space consistent with leptogenesis. The experimental sensitivity to heavy neutrinos embedded in UV complete theories, like supersymmetry or type III 2HDM, is discussed, e.g., for the ILC, in Ref. [27].

9 Rare processes and indirect BSM physics searches

The presence of new phenomena at a high-energy scale can impact the production rate of processes that are otherwise very rare in the SM. Examples already mentioned in this review are the anomalies in measurements such as R_K and R^*_K at LHCb [15] that can be explained by the presence of LQ or Z'. Flavor-changing neutral current effects in the heavy-quarks and gauge boson sectors are another case extensively investigated at future colliders. Prospect studies can be found in (32) and references therein, with the expected sensitivity of future lepton collider proposals at critical production thresholds complementary to that of hadron colliders at very high luminosities.

In the absence of evidence for new physics, the formalism of EFT can be adopted as a global framework to perform model-independent searches. Two effective field theory approaches are considered here. The first one, the Standard Model EFT or SMEFT, extends the SM with operators ($\mathcal{O}_i$) at higher canonical dimension "d," constructed as combinations of SM fields, invariant under the Lorentz and SM gauge symmetries. If lepton and baryon number conservation is imposed to reduce the otherwise very large number of possible new operators, the first corrections to the SM are provided by operators of dimension six. BSM physics at energies below Λ is then described by a Lagrangian $\mathcal{L} = \mathcal{L}_{SM} + \mathcal{L}_{BSM}$, where $\mathcal{L}_{BSM} = \sum_{d>4} \frac{1}{\Lambda^{d-4}} \mathcal{L}_d$ and $\mathcal{L}_d = \sum_i c_i^{(d)} O_i^{(d)}$. The Wilson coefficients $c_i^{(d)}$ depend on the structure of new physics. Furthermore, the first corrections to the SM are provided by operators of dimension six if lepton and baryon number conservation is imposed. Since BSM-induced corrections to the SM parameters can be grouped into sets of models, any deviations of the SM parameters from their expectations could provide an indication about $\frac{c_i}{\Lambda^2}$. Thanks to its linearized Lagrangian, SMEFT is an excellent tool to probe for weakly coupled theories. Recent studies based on global fits to SMEFT operators are documented in Refs [30, 38] and shown in Figure 12: these selected results, shown as a relative improvement compared to the HL-LHC results, indicate that BSM scales between 1 and several tens of TeVs can be probed at future colliders under the assumption of $c_i \sim 1$. Precision measurements at future lepton colliders, in particular where Z-pole runs are planned, contribute substantially to the extraction of EW but also triple-gauge coupling parameters. Diboson and $Zh, h \to b\bar{b}$ measurements in the boosted regime are among the probes most relevant for high-energy pp

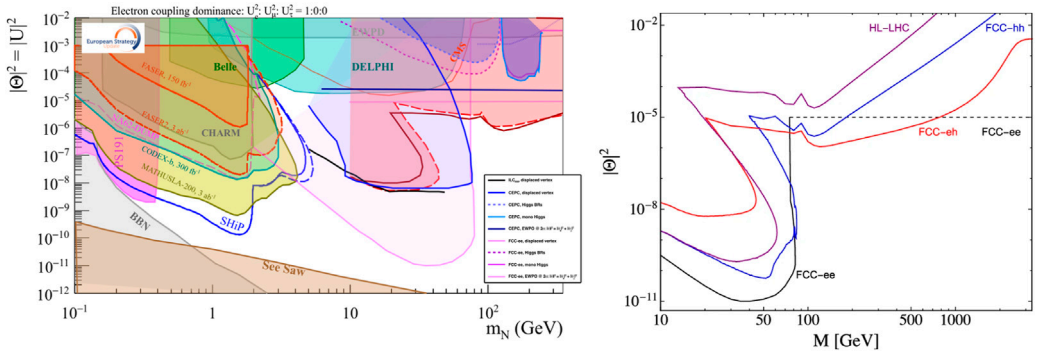


FIGURE 11
Left: Exclusion reach for minimal heavy lepton portal models assuming electron-coupling dominance in the $|\Theta|^2$– m_N plane at various experiments. Exclusion limits are computed at 90% CL [31]. Right: Comparison of exclusion reaches at 90% CL from searches at the HL-LHC, FCC-hh, and FCC-eh and precision measurements at the FCC-ee [57].

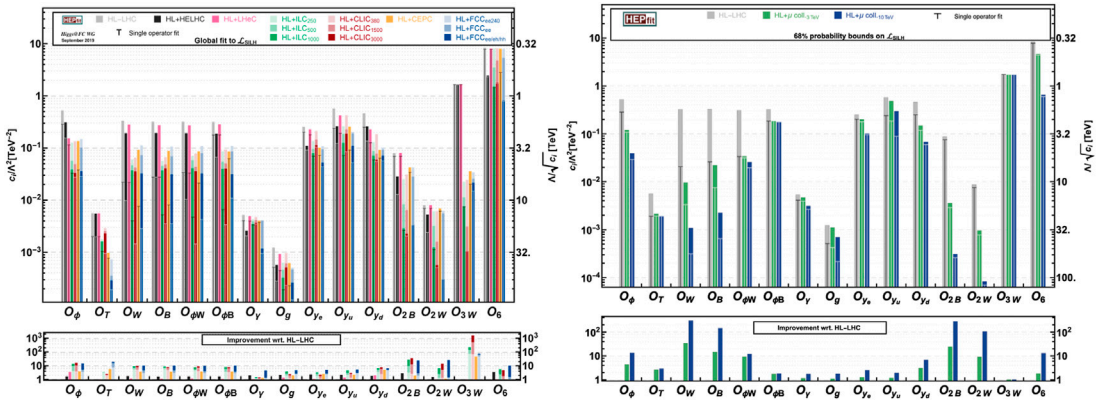


FIGURE 12
Probability reach for the Wilson coefficients computed at 68% CL from the global fit (solid bars). The vertical "T"-shaped lines report the results obtained if only one operator is generated by the UV dynamics. Left: Reach of all options considered in Ref. [38]. Right: Sensitivity of the muon collider at both $\sqrt{s}$ = 3 and 10 TeV, compared to that of the HL-LHC [30].

colliders. For muon colliders, high-energy measurements in two-to-two fermion processes as well as single-Higgs and di-Higgs precision measurements have been considered. Even though the SMEFT provides a consistent framework to describe the impact of BSM physics, it is important to highlight that the results of the global fit depend on the choice of operators, basis, selected observables and their correlations, experimental and theoretical systematic uncertainties, etc. A detailed discussion on the subject can be found in Ref. [58]. For instance, off-diagonal flavor structures are not considered. If the imposed benchmark flavor symmetry is relaxed, top-quark FCNC interactions can be incorporated [59] into the SMEFT framework.

The second EFT approach is the so-called Higgs EFT (HEFT). In this framework, the Higgs boson is not required to belong to an exact $SU(2)_L$ doublet, and the physical Higgs and the three EW Goldstone bosons are treated as independent objects. The physical Higgs is typically assigned to a singlet representation of the SM gauge groups. The HEFT, with a non-linear realization of the EWSB, offers the most general description of the Higgs couplings, and it is suitable to investigate a large set of distinct theories, including composite Higgs, and scenarios with modified trilinear Higgs couplings. Even though HEFT is outside the scope of this review, the reader is invited to find detailed comparative overviews of SMEFT and HEFT in Refs [60, 61].

10 State of the art of the proposed colliders

The broad and ambitious science program presented in the previous sections, and references therein, depends critically upon the performance of the accelerators and experiments, including both the instrumentation and software and computing elements. Advances in theoretical methods are also essential for the full exploitation of

these machines' discovery potential. The technology needed to realize these projects might not exist yet, and cutting-edge and vigorous R&D is therefore being pursued by the global HEP experimental community. Extensive studies on detector concepts are ongoing within the broader worldwide detector R&D programs, as recently presented in Ref. [62]. A succinct summary of the status of the machines considered in this paper, as well as recently proposed modifications and upgrades, is presented in the following paragraphs. The reader is invited to consult dedicated reviews to learn about the latest developments in advanced accelerator techniques, instrumentation, software and computing, and theoretical calculations and methods.

The technical design report (TDR) of the ILC, released in 2013 [64], focused on the 250–500 GeV option (with a possible energy upgrade to 1 TeV). The recent comprehensive report submitted to Snowmass [27] reviews in detail the accelerator design, proposes new luminosity and energy running conditions (including technology options for multi-TeV upgrades), and presents robust solutions to deliver electron and positron beams in the energy region of the Higgs boson. In addition, it updates the proposal made in the detector TDR [63] for two detectors at the interaction region, the SiD and the ILD detector, and outlines that further R&D is needed. These new detector designs have been carried out at the level of a conceptual design report (CDR). In 2020, ICFA approved the formation of the ILC International Development Team as the first step toward the preparatory phase of the ILC project, with a mandate to make preparations for the ILC Pre-Lab in Japan [64] while waiting for a decision by Japan about hosting the facility. If approved, the machine is expected to deliver physics data in the mid-2030s.

As a direct response to a high-priority request from the 2013 update of the European Strategy for Particle Physics, CERN developed the design for the Future Circular Collider. The CDR for the FCC housed in a 100-km-long tunnel at CERN was delivered in 2018 with preliminary cost estimates and feasibility assessments [6, 9, 42]. Updates were presented within the 2018 process for the European Strategy and, more recently, in the context of the Snowmass community planning process. The latest proposals [65–67] include a phased approach with an e^+e^- machine at various center-of-mass energies (including at the Higgs mass), followed by the 100 TeV pp, the ep, and the heavy ion programs. Under the auspices of CERN, the FCC Collaboration is also considering a tunnel of 91 km. Assuming a timely completion of the R&D for the FCC-ee, start of operations is expected in the 2040s, with data taking lasting till 2060. The FCC-hh is scheduled to run between 2070 and 2090. The program may be modified to focus on the pp, the ep, and the heavy ion programs if an e^+e^- collider other than the FCC-ee is approved for construction. Two to four experiments could be hosted by the FCC complex at each stage.

After the release of the CepC CDR and subsequent studies documented in Ref. [68], the CepC accelerator study group entered the TDR phase, expected to be completed by the end of 2022. Meanwhile, an update to the design of the CepC and SppC, including a target center-of-mass energy of 125 TeV for pp collisions, is discussed in Ref. [69] and Ref. [70], respectively. According to the currently envisioned schedules, the CEPC (SppC) starts operations in the mid-2030s (2050s).

The CLIC CDR [7], documenting the 3 TeV machine, dates back to 2012, while a project implementation plan, PIP [71], was finalized in 2018 and included the option for the 380 GeV running. The recent Snowmass report [72], building on both the CDR and the PIP, describes recent achievements in accelerator design, technology development, system tests, and beam tests for CLIC, demonstrating that the performance goals are realistic. In addition, results from ongoing R&D are likely to allow for further upgrades, both in $\sqrt{s}$ and instantaneous luminosity. Assuming project approval in 2028 (after the next ESPP), construction can start in ~2030 and operations ~7 years later.

The option of a muon collider has gained substantial interest in the past two years. Documented in Ref. [73] is the latest proposal for a muon collider with three tentative target center-of-mass energies: 3, 10, and 14 TeV. Other energy conditions are also explored, e.g., operations at $\sqrt{s}$ =125 GeV. It is recognized that the muon collider is not as mature as the other high-energy lepton collider options listed previously. However, no major technical limitations are identified to date, and the outlined R&D path to address the remaining challenges makes the 3 TeV viable, with start of data taking in the mid-2040s.

As mentioned in Section 2, the newly proposed C^3 linear accelerator [17] benefits from recent advances that increase the efficiency and operating gradient of a normal conducting accelerator and provides a solution to a multi-TeV machine. The current proposal is for a compact 8-km-long cold copper-distributed coupling complex that could fit on the Fermilab site. The Snowmass 2021-contributed paper [18] instead discusses the Higgs-Energy LEptoN (HELEN) e^+e^- linear collider, based on advances in superconducting radio frequency technology, but with potential cost and AC power savings and smaller footprint (relative to the ILC). If the ILC cannot be realized in Japan in a timely fashion, the HELEN collider would be another viable option for a Higgs factory in the U.S.

11 Conclusion

Several collider projects have been proposed and discussed in recent years. Each proposal offers compelling opportunities for precision measurements and searches for new physics, albeit carrying challenges in accelerator, detector, and computing technologies. The ESPPU and the Snowmass/P5 processes have outlined future prospects at linear and circular e^+e^-, high-energy pp, ep, and high-energy $\mu^+\mu^-$ colliders, along with their capability to solve long-standing problems, such as the understanding the EWSB mechanism, the origin and nature of dark matter, the flavor problem, the origin of neutrino masses, the strong CP problem, and baryogenesis. This review has briefly summarized the outcomes of those processes, resulting from the huge combined effort of the theory and experimental particle physics communities during the last 5 years, in the context of direct and indirect searches for new physics. Well-motivated BSM scenarios have been considered to provide quantitative comparisons between the reach of different proposed projects. The emerging picture shows that, while there are excellent chances for fundamental discoveries at the HL-LHC, the datasets might not be sufficient to fully characterize new physics if deviations from the SM are found, and future colliders will make this possible. Each future collider offers exciting prospects to enable the exploration of the unknown beyond the HL-LHC, and the realization of one or more of the proposed accelerator projects in the next decades should be strongly pursued by the scientific community to guarantee unique advancements in the understanding of the laws of nature.

Author contributions

AC and MD reviewed the existing published material on searches for new physics at future colliders and summarized it in this review paper. All authors contributed to the article and approved the submitted version.

Funding

MD was partially funded by STFC (PP(E)CG 2022-2025, ST/W000466/1), United Kingdom, and by the CHIST-ERA grant, CHIST-ERA-19-XAI-009, under Horizon 2020. AC was supported by Fermi Research Alliance, LLC, under Contract No. DE-AC02- 07CH11359, with the U.S. Department of Energy, Office of Science, Office of High-Energy Physics.

Acknowledgments

The authors wish to thank the thousands of scientists that in the past years have worked strenuously on the proposals considered in this report for their dedication and vision. The authors thank Carl Gwilliam for useful discussions and the thorough review of this manuscript and Oliver Fischer for useful discussions about future colliders' reach for heavy neutral leptons.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Cepeda M, Gori S, Ilten P, Kado M, Riva F, Abdul Khalek R, et al. Report from working group 2 Higgs physics at the HL-LHC and HE-LHC. *CERN Yellow Rep Monogr* (2019) 7:221. arXiv 1902.00134. doi:10.23731/CYRM-2019-007.221

2. Azzi P, Farry S, Nason P, Tricoli A, Zeppenfeld D, Abdul Khalek R, et al. Report from working group 1 standard model physics at the HL-LHC and HE-LHC. *CERN Yellow Rep Monogr* (2019) 7:1. arXiv 1902.04070. doi:10.23731/CYRM-2019-007.1

3. Agostini P, Aksakal H, Alekhin S, Allport PP, Andari N, Andre K, et al. The large Hadron\textendash{}Electron collider at the HL-LHC. *J Phys G* (2021) 48:110501. arXiv 2007.14491.

4. Vidal X, D'Onofrio M, Fox P, Torre R, Ulmer KA, Aboubrahim A, et al. Report from working group 3 beyond the standard model physics at the HL-LHC and HE-LHC. *CERN Yellow Rep Monogr* (2019) 7:585. arXiv 1812.07831.

5. Cerri A, Gligorov VV, Malvezzi S, Martin Camalich J, Zupan J, Akar S, et al. Report from working group 4 opportunities in flavour physics at the HL-LHC and HE-LHC. *CERN Yellow Rep Monogr* (2019) 7:867. arXiv 1812.07638. doi:10.23731/CYRM-2019-007.867

6. Abada A, Abbrescia M, AbdusSalam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. HE-LHC: The high-energy large hadron collider. *Eur Phys J ST* (2019) 228:1109–382. doi:10.1140/epjst/e2019-900088-6

7. Aicheler M, Burrows P, Draper M, Garvey T, Lebrun P, Peach K, et al. A multi-TeV linear collider based on CLIC technology CLIC conceptual design report. *CERN Yellow Rep Monogr* (2012) 7. doi:10.5170/CERN-2012-007

8. CEPC Study Group, (2018), CEPC conceptual design report: Volume 1 - accelerator, arXiv 1809.00285.

9. Abada A, Abbrescia M, AbdusSalam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. FCC-hh: The hadron collider future circular collider conceptual design report volume 3. *Eur Phys J ST* (2019) 228:755–1107. doi:10.1140/epjst/e2019-900087-0

10. Arduini G, Brüning O, Klein M Energy frontier DIS at CERN: The LHeC and the FCC-eh, PERLE. *Pos DIS* (2018). 183:8. doi:10.22323/1.316.0183

11. Willeke F *BNL-221006-2021-FORE* (2021).

12. Abada A, Abbrescia M, AbdusSalam SS, Abdyukhanov I, Abelleira Fernandez J, Abramov A, et al. HE-LHC: The high-energy large hadron collider: Future circular collider conceptual design report volume 4. *Eur Phys J ST* (2019) 228:1109–382. doi:10.1140/epjst/e2019-900088-6

13. Long K, Lucchesi D, Palmer M, Pastrone N, Schulte D, Shiltsev V. Muon colliders: Opening new horizons for particle physics. *Nat Phys* (2021) 17:289–92. arXiv 2007.15684. doi:10.1038/s41567-020-01130-x

14. Muon G, Abi B, Al-Kilani S, Allspach D, Alonzi L, Anastasi A, et al. Measurement of the positive muon anomalous magnetic moment to 0.46 ppm. *Phys Rev Lett* (2021) 126:141801. arXiv 2104.03281. doi:10.1103/PhysRevLett.126.141801

15. Aaij R, Abellán Beteta C, Ackernley T, Adeva B, Adinolfi M, Afsharnia H, et al. Test of lepton universality in beauty-quark decays, *Nat Phys*. 18, 277 (2022), arXiv 2103.11769. doi:10.1038/s41567-021-01478-8

16. Bhat P, Jindariani S, Ambrosio G, Apollinari G, Belomestnykh S, Bross A, et al. (2022), Future collider options for the US, arXiv 2203.08088.

17. Bai M, Barklow T, Bartoldus R, Breidenbach M, Grenier P, Huang Z, et al. (2021), C3: A "Cool" route to the Higgs boson and beyond, arXiv 2110.15800.

18. Belomestnykh S, Bhat PC, Grassellino A, Checchin M, Denisov D, R. L. Geng, et al. (2022), Higgs-Energy LEptoN (HELEN) Collider based on advanced superconducting radio frequency technology, arXiv 2203.08211.

19. Ahmad M, Alves D, An H, An Q, Arhrib A. *CEPC-SPPC preliminary conceptual design report. 1. Physics and detector*. IHEP-CEPC-DR-2015-01 (2015).

20. Ariga A, Ariga T, Boyd J, Cadoux F, Casper DW, Favre Y, et al. (2019), Faser: ForwArd search ExpeRiment at the LHC, arXiv 1901.04468.

21. Ahdida C, Akmete A, Albanese R, Alexandrov A, Andreini M, Anokhina A, et al. (2020), SND@LHC, arXiv 2002.08722.

22. Lubatti H, Alpigiani C, Arteaga-Velázquez JC, Ball A, Barak L, Beacham J, et al. Explore the lifetime frontier with MATHUSLA. *JINST* (2020) 15:C06026. arXiv 1901.04040. doi:10.1088/1748-0221/15/06/c06026

23. Aielli G, Ben-Haim E, Cardarelli R, Charles MJ, Cid Vidal X, Coco V, et al. Expression of interest for the CODEX-b detector. *Eur Phys J C* (2020) 80:1177. arXiv 1911.00481. doi:10.1140/epjc/s10052-020-08711-3

24. Milli Q, Yoo Y. The milliQan experiment: Search for milli-charged particles at the LHC. *PoS* ***ICHEP2018*** (2019) 520. arXiv 1810.06733.

25. Abelleira-Fernandez J, Akay AN, Aksakal H, Albacete JL, Alekhin S, et al. A large hadron electron collider at CERN: Report on the physics and design concepts for machine and detector. *J Phys G* (2012) 39:075001. arXiv 1206.2913. doi:10.1088/0954-3899/39/7/075001

26. Anchordoqui L, Ariga A, Ariga T, Bai W, Balazs K, Batell B, et al. (2021), The forward physics facility: Sites, experiments, and physics potential, arXiv 2109.10905.

27. Aryshev A, Behnke T, Berggren M, Brau J, Craig N, Freitas A, et al. (2022), The international linear collider: Report to Snowmass 2021, arXiv 2203.07622.

28. Langacker P The physics of heavy Z ′ gauge bosons. *Rev Mod Phys* (2009) 81:1199–228. arXiv 0801.1345. doi:10.1103/RevModPhys.81.1199

29. Helsens C, Jamin D, Mangano M, Rizzo T, Selvaggi M. Heavy resonances at energy-frontier hadron colliders. *Eur Phys J C* (2019) 79:569. arXiv 1902.11217. doi:10.1140/epjc/s10052-019-7062-3

30. de Blas J, Buttazzo D, Capdevilla R, Curtin D, Franceschini R, Maltoni F, et al. (2022), The physics case of a 3 TeV muon collider stage, arXiv 2203.07261.

31. Ellis R, Beate H, de Blas J, Maria C, Christophe G, Fabio M, et al. (2019), Physics briefing book input for the European Strategy for particle physics update 2020 arXiv 1910.11775.

32. Huang G, Jana S, Queiroz F, Rodejohann W. Probing the ${R}_{{K}^{(*)}}$ anomaly at a muon collider. *Phys Rev D* (2022) 105:015013. doi:10.1103/PhysRevD.105.015013

33. Giudice G, Grojean C, Pomarol A, Rattazzi R The strongly-interacting light Higgs. *JHEP* (2007) 06:045. 0703164. doi:10.1088/1126-6708/2007/06/045

34. Aimè C, Apyan A, Attia Mahmoud M, Bartosik N, Bertolin A, Bonesini M, et al. (2022), Muon collider physics summary, arXiv 2203.07256.

35. Barducci D, Mimasu K, No J, Vernieri C, Zurita J, Enlarging the scope of resonant di-Higgs searches: Hunting for Higgs-to-Higgs cascades in 4b final states at the LHC and future colliders, *JHEP* 02, 2(2020), arXiv 1910.08574. doi:10.1007/JHEP02(2020)002

36. No J, Spannowsky M Signs of heavy Higgs bosons at CLIC: An $$ê+ ê-$$ road to the electroweak phase transition. *Eur Phys J C* (2019) 79:467. arXiv 1807.04284. doi:10.1140/epjc/s10052-019-6955-5

37. Buttazzo D, Franceschini R, Wulzer A. Two paths towards precision at a very high energy lepton collider. *JHEP* (2021) 05:219. arXiv 2012.11555. doi:10.1007/JHEP05(2021)219

38. de Blas J, Cepeda M, D'Hondt J, Ellis R, Grojean C, Heinemann B, et al. Higgs boson studies at future particle colliders. *JHEP* (2020) 01:139. arXiv 1905.03764. doi:10.1007/JHEP01(2020)139

39. Athron P, Balazs C, Gonzalo TE, Jacob D, Mahmoudi F, Sierra C. Likelihood analysis of the flavour anomalies and **g−2** in the general two Higgs doublet model. *JHEP* (2022) 01:037. arXiv 2111.10464. doi:10.1007/jhep01(2022)037

40. Baer H, Berggren M, Fujii K, List J, Lehtinen SL, Tanabe T, et al. ILC as a natural SUSY discovery machine and precision microscope: From light Higgsinos to tests of unification. *Phys Rev D* (2020) 101:095026. arXiv 1912.06643. doi:10.1103/PhysRevD.101.095026

41. de Blas J, Franceschini R, Riva F, Roloff P, Schnoor U, Spannowsky M, et al. The CLIC potential for new physics, 3/2018 (2018), arXiv 1812.02093. doi:10.23731/CYRM-2018-003

42. Abada A, Abbrescia M, AbdusSalam SS, Abdyukhanov I, Fernandez JA, Abramov A, et al. FCC physics opportunities future circular collider conceptual design report volume 1. *Eur Phys J C* (2019) 79:474. doi:10.1140/epjc/s10052-019-6904-3

43. Bambade P, Barklow T, Behnke T, Berggren M, Brau J, Burrows P, et al. (2019), The international linear collider: A global project arXiv 1903.01629.

44. Battaglia M, Blaising JJ, Marshall JS, Poss S, Sailer A, Thomson M, et al. Physics performance for scalar electron, scalar muon and scalar neutrino searches at $ \sqrt{s} $ = 3 TeV and 1.4 TeV at CLIC. *JHEP* (2013) 09:001. arXiv 1304.2825. doi:10.1007/JHEP09(2013)001

45. Baum S, Sandick P, Stengel P. Hunting for scalar lepton partners at future electron colliders. *Phys Rev D* (2020) 102:015026. arXiv 2004.02834. doi:10.1103/PhysRevD.102.015026

46. Chakraborti M, Heinemeyer S, Saha I. The new "MUON G-2" result and supersymmetry. *Eur Phys J C* (2021) 81:1114. arXiv 2104.03287. doi:10.1140/epjc/s10052-021-09900-4

47. Wang Z-S, Wang K. Long-lived light neutralinos at future Z factories. *Phys Rev D* (2020) 101:115018. doi:10.1103/PhysRevD.101.115018

48. Curtin D, Deshpande K, Fischer O, Zurita J. Closing the light gluino gap with electron-proton colliders. *Phys Rev D* (2019) 99:055011. arXiv 1812.01568. doi:10.1103/physrevd.99.055011

49. Abercrombie D, Akchurin N, Akilli E, Maestre JA, Allen B, Gonzalez BA, et al. Dark matter benchmark models for early LHC run-2 searches: Report of the ATLAS/CMS dark matter forum. *Phys Dark Univ* (2020) 27:100371. arXiv 1507.00966. doi:10.1016/j.dark.2019.100371

50. O'Hare C. New definition of the neutrino floor for direct dark matter searches. *Phys Rev Lett* (2021) 127:251802. doi:10.1103/PhysRevLett.127.251802

51. Agrawal P, Bauer M, Beacham J, Berlin A, Boyarsky A, Cebrian S, et al. Feebly-interacting particles: FIPs 2020 workshop report. *Eur Phys J C* (2021) 81:1015. arXiv 2102.12143. doi:10.1140/epjc/s10052-021-09703-7

52. Curtin D, Essig R, Gori S, Shelton J. Illuminating dark photons with high-energy colliders. *JHEP* (2015) 02:157. arXiv 1412.0018. doi:10.1007/JHEP02(2015)157

53. D'Onofrio M, Fisher O, Wang Z. Searching for dark photons at the LHeC and FCC-he. *Phys Rev D* (2020) 101:015020. arXiv 1909.02312. doi:10.1103/PhysRevD.101.015020

54. Fabbrichesi M, Gabrielli E, Lanfranchi G, (2020), The dark photon arXiv 2005.01515. doi:10.1007/978-3-030-62519-1

55. Bauer M, Heiles M, Neubert M, Thamm A. Axion-like particles at future colliders. *Eur Phys J C* (2019) 79:74. arXiv 1808.10323. doi:10.1140/epjc/s10052-019-6587-9

56. Alimena J, Bauer M, Azzi P, Ruiz R, Neubert M, Rizzi C, et al. Searches for long-lived particles at the future FCC-ee. *Front Phys* (2022) 10:967881. arXiv 2203.05502.

57. Antusch A, Cazzato E, Fischer O. Sterile neutrino searches at future e−e+, pp and e−p colliders. *Int J Mod Phys A* (2017) 32:1750078. arXiv 1612.02728. doi:10.1142/S0217751X17500786

58. de Blas J, Durieux G, Grojean C, Gu J, Paul A On the future of Higgs, electroweak and diboson measurements at lepton colliders. *JHEP* (2019) 12:117. arXiv 1907.04311. doi:10.1007/JHEP12(2019)117

59. Aguilar-Saavedra J, Degrande C, Durieux G, Maltoni F, Vryonidou E, Zhang C, et al. (2018), Interpreting top-quark LHC measurements in the standard-model effective field theory arXiv 1802.07237.

60. Cohen T, Craig N, Lu X, Sutherland D. Is SMEFT enough? *JHEP* (2021) 03:237. arXiv 2008.08597. doi:10.1007/JHEP03(2021)237

61. Brivio I, Gonzalez-Fraile J, Gonzalez-Garcia M, Merlo L. The complete HEFT Lagrangian after the LHC Run I. *Eur Phys J C* (2016) 76:416. arXiv 1604.06801. doi:10.1140/epjc/s10052-016-4211-9

62. ECFA Detector R&D Roadmap Process Group. *ECFA detector R&D roadmap process group* (2020).

63. Behnke T, Brau J. E., Foster J, Fuster J, Harrison M, McEwan Paterson J, et al. (2013), The international linear collider technical design report - volume 1: Executive summary, arXiv 1306.6329.

64. ICFA. *ICFA announces a new phase towards preparation for the International Linear Collider*. ICFA (2020). Available at: https://www.interactions.org/press-release/icfa-announces-new-phase-towards-preparation-international.

65. Bernardi G, Brost E, Denisov D, Landsberg G, Aleksa M, d'Enterria D, et al. (2022), The future circular collider: A summary for the US 2021 Snowmass process, arXiv 2203.06520.

66. Agapov I, Benedikt M, Blondel A, Boscolo M, Brunner O, Chamizo Llatas M, et al. (2022), Future circular lepton collider FCC-ee: Overview and status, arXiv 2203.08310.

67. Benedikt M, Chance A, Dalena B, Denisov D, Giovannozzi M, Gutleber J, et al. (2022), Future circular hadron collider FCC-hh: Overview and status, arXiv 2203.07804.

68. CEPC Accelerator Study Group, (2019), arXiv 1901.03169. The CEPC input for the European Strategy for particle physics - accelerator.

69. CEPC-Accelerator-Study-Group, (2022), Snowmass2021 white paper AF3-CEPC, arXiv 2203.09451.

70. Tang J, Zhang Y, Xu Q, Gao J, Lou X, Wang Y, (2022), arXiv 2203.07987. Study overview for super proton-proton collider.

71. CLIC acceleratorAicheler M, Burrows PN, Catalan N, Corsini R, Draper M, Osborne J, et al. *The compact linear collider (CLIC) - project implementation plan*, 4 (2018). arXiv 1903.08655.

72. Brunner O, Burrows PN, Calatroni S, Catalan Lasheras N, Corsini R, D'Auria G, et al. (2022), The CLIC project, arXiv 2203.09186.

73. Stratakis D, Mokhov N, Palmer M, Pastrone N, Raubenheimer T, Rogers C, et al. (2022), A muon collider facility for physics discovery, arXiv 2203.08033

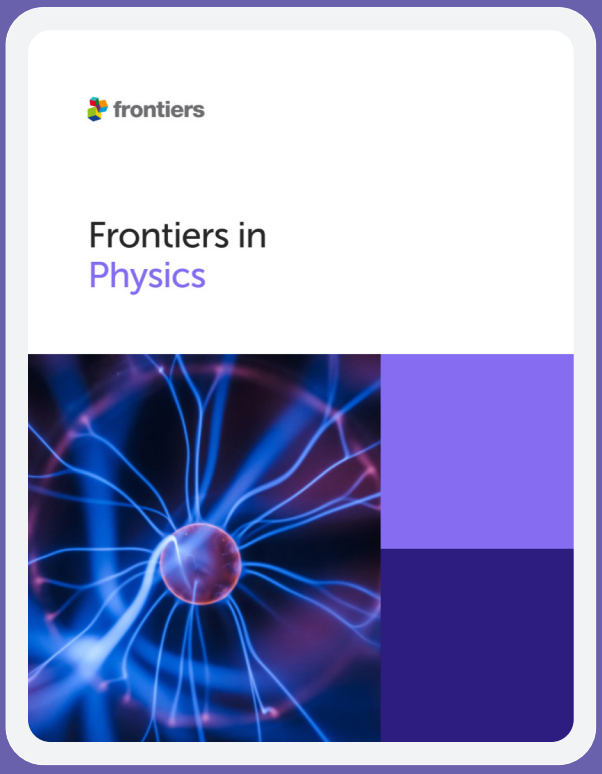
Frontiers in Physics

Investigates complex questions in physics to understand the nature of the physical world

Addresses the biggest questions in physics, from macro to micro, and from theoretical to experimental and applied physics.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact