frontiers | Research Topics

Distributed learning, optimization, and control methods for future power grids

Edited by
Zhi-Wei Liu, Xiao-Kang Liu, Jun Yan and Chao Deng

Published in
Frontiers in Energy Research

ISSN 1664-8714
ISBN 978-2-8325-4152-4
DOI 10.3389/978-2-8325-4152-4

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

Distributed learning, optimization, and control methods for future power grids

Topic editors

Zhi-Wei Liu — Huazhong University of Science and Technology, China
Xiao-Kang Liu — Huazhong University of Science and Technology, China
Jun Yan — Concordia University, Canada
Chao Deng — Nanjing University of Posts and Telecommunications, China

Citation

Liu, Z.-W., Liu, X.-K., Yan, J., Deng, C., eds. (2024). *Distributed learning, optimization, and control methods for future power grids*.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4152-4

Table of contents

frontiers | Frontiers in Energy Research

TYPE Editorial
PUBLISHED 04 December 2023
DOI 10.3389/fenrg.2023.1337268

OPEN ACCESS

EDITED AND REVIEWED BY
ZhaoYang Dong,
Nanyang Technological University,
Singapore

*CORRESPONDENCE
Zhi-Wei Liu,
zwliu@hust.edu.cn

RECEIVED 12 November 2023
ACCEPTED 20 November 2023
PUBLISHED 04 December 2023

CITATION
Liu Z-W, Liu X-K, Yan J and Deng C (2023), Editorial: Distributed learning, optimization, and control methods for future power grids.
Front. Energy Res. 11:1337268.
doi: 10.3389/fenrg.2023.1337268

Editorial: Distributed learning, optimization, and control methods for future power grids

Zhi-Wei Liu[1]*, Xiao-Kang Liu[1], Jun Yan[2] and Chao Deng[3]

[1]The Key Laboratory of Image Processing and Intelligent Control, Ministry of Education, The School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China, [2]Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC, Canada, [3]Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China

KEYWORDS

distributed learning, distributed optimization, distributed control, data analytics, microgrids, demand response

Editorial on the Research Topic
Distributed learning, optimization, and control methods for future power grids

The development of renewable energy sources makes an important contribution to low-carbon sustainable development. However, renewable energy such as wind and photovoltaic energy are highly random and intermittent, which brings great challenges to the operation and control of power networks. The traditional method is to build a large number of peak-shaving power stations and arrange a large number of energy storage devices to stabilize the fluctuations of renewable energy, and the cost may be prohibitive to be sustainable. In the future power network, advanced sensors, actuators, and communication equipment will be deployed on various systems such as generators, substations, transformers, distributed energy resources, air conditioners, and electric vehicles. A challengeable issue arises on how to manage these hundreds of millions of active endpoints. Recently, distributed learning, optimization, and control methods for networked systems have received growing attention. Distributed learning, optimization, and control methods for managing hundreds of millions of active endpoints in the future power network are expected to enable a stable and economic operation of the power network with a high proportion of new energy, as illustrated in Figure 1.

This issue contains nine research articles focusing on distributed control and optimization for power systems and smart power electronics, multi-agent reinforcement learning in power systems, advanced energy management and economic dispatch, etc.

The paper by Ma et al. proposes a multi-time-scale control method based on a deep Q network-deep deterministic policy gradient (DQN-DDPG) algorithm to maintain optimal voltage in distribution networks. It optimizes voltage regulation for longer times using the DQN algorithm and shorter times using the DDPG algorithm. Then, based on the Markov decision process transformation, the design strategy of the DQN-DDPG algorithm is proposed to extend the energy storage capacity with respect to the stated objectives and constraints, taking into account the state of charge of the energy storage. A simulation platform backs up the proposed strategy, showing its effectiveness compared to a particle swarm optimization algorithm.

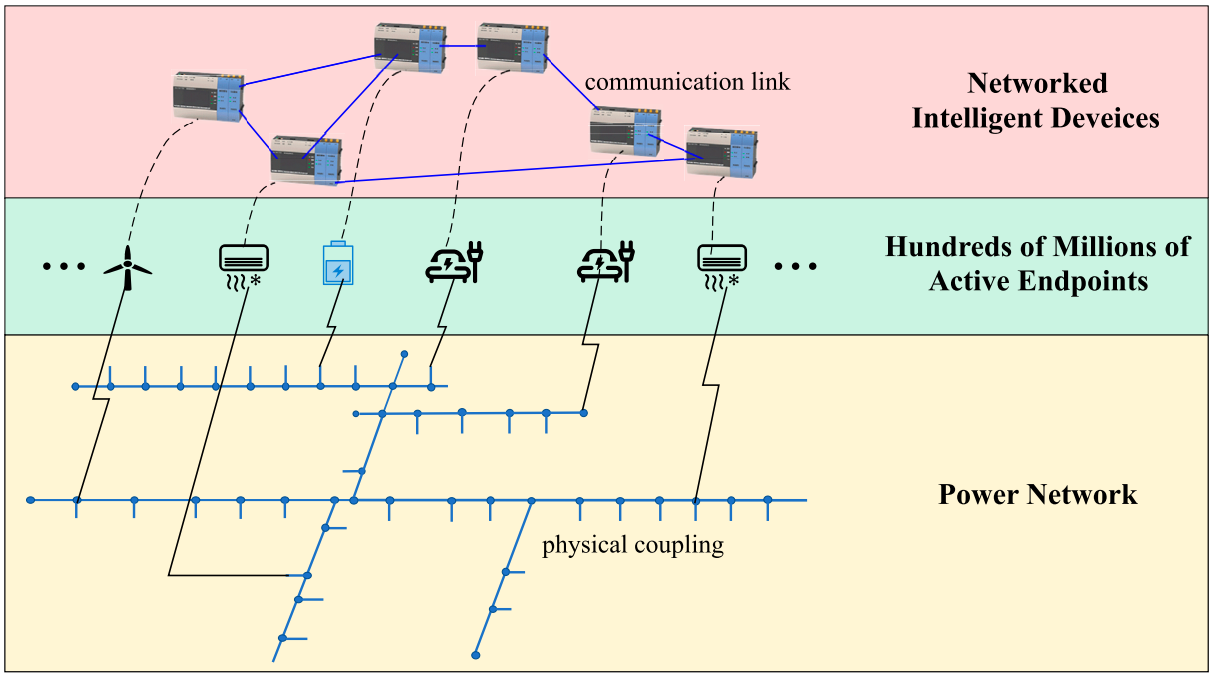


FIGURE 1
Simplified diagram of distributed learning, optimization, and control methods for future power grids.

The stable operation of microgrid clusters is the focus and difficulty of current research. Shi et al. propose a ring-based multi-agent microgrid cluster energy management strategy, which realizes the centerless coordinated autonomous operation of microgrid clusters with high stability. They also offer switchable control strategies for the microgrid cluster's varying control objectives, allowing for seamless grid connectivity changes. Simulation results prove that the proposed method enhances the stability, autonomy, and efficiency of energy utilization in microgrid clusters.

The paper by Xiao et al. proposes a joint sensor secure rate and energy efficiency optimization algorithm to enhance Wireless Sensor Network (WSN) performance in intelligent management of a photovoltaic power system, obtaining the optimal topology and transmission power design. The proposed algorithm comprehensively considers the factors those affect the security and efficiency of WSN. Moreover, Block Coordinate Descent technology is utilized to minimize the energy consumption of WSN and maximize the secure rate of sensor networks, thus ensuring the security and reliability of monitoring. Finally, the sensor association, modeled as a 0–1 multi-knapsack optimization problem, is compared using different methods.

With the increase of devices in power grids, a critical challenge emerges on how to collect information from massive devices, as well as how to manage these devices. Zhao et al. explore device scheduling, where a portion of mobile devices are selected at each time slot to collect more valuable sensing data. However, the device scheduling task is difficult due to the lack of *a priori* knowledge and the large amount of data. Therefore, the device scheduling problem is reformulated as a multi-armed bandit program, which is solved by a device scheduling algorithm based on upper confidence bound policy and virtual queue theory. Simulation results verify the effectiveness of the proposed algorithm, in terms of performance regret and convergence rate.

The trading of virtual power plants (VPPs) should not only consider the economy but also its degree of low carbon. The paper by Chu et al. construct a unified bidding strategy for multi-VPPs that considers carbon–electricity integration trading. After each VPP determines its internal trading strategy, multi-game trading strategy between multiple VPPs is designed to achieve unified trading. Simulations verify that this approach can enhance the efficiency and trading income of VPPs, promoting new energy consumption.

The paper by Han proposes a fault-tolerant control scheme using a wavelet analysis and consensus algorithm to tackle voltage and frequency regulation issues in smart grids impacted by faults. Faults are identified through a wavelet analysis, followed by a distributed fault estimator to capture attack signals. The simulation of a smart grid with four distributed generations in MATLAB/Simulink illustrates the method's effectiveness in achieving voltage and frequency regulation objectives.

The paper by Shen et al. proposes a magnetic induction positioning and communication system for underwater use, including an energy-efficient distributed control algorithm for base stations. This technology uses a network of base stations and an Autonomous Underwater Vehicle with three-axis source coils for precise location and communication. The algorithm controls base

stations to meet operational needs while minimizing energy use, achieving stable, long-term function.

The paper by Yang et al. proposes a Lyapunov-based power sharing control scheme and a fixed-time-based distributed optimization algorithm to achieve optimal power sharing of sources in a DC microgrid. The controller uses a ratio consensus protocol to achieve proportional power sharing by modifying the microgrid's voltage profile. The optimizer, integrating a finite-time weighted consensus algorithm with an iterative algebraic operation, calculates optimal power dispatch to minimize generation costs. Both units function in a wholly distributed manner. Stability and convergence analyses of the control scheme and optimization algorithm are provided.

The paper by Qi et al. proposes a novel visual-admittance-based model predictive control scheme to cope with the problem of vision/force control and several constraints of a nuclear collaborative robotic visual servoing system. The scheme considers the desired image features and force commands in the image feature space, and then uses the desired force commands as constraints for model predictive control, thus eliminating overshooting in interactive force control in most cases. Simulation results with a two-degree-of-freedom robot manipulator confirm this method's effectiveness.

Author contributions

Z-WL: Writing–original draft. X-KL: Writing–review and editing. JY: Writing–review and editing. CD: Writing–review and editing.

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

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 11 January 2023
DOI 10.3389/fenrg.2022.1097319

OPEN ACCESS

EDITED BY
Xiao-Kang Liu,
Huazhong University of Science and Technology, China

REVIEWED BY
Qi-Fan Yuan,
Huazhong University of Science and Technology, China
Xinyao Li,
Nanyang Technological University, Singapore

*CORRESPONDENCE
Cangbi Ding,
dcb19960926@163.com

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 13 November 2022
ACCEPTED 22 December 2022
PUBLISHED 11 January 2023

CITATION
Ma M, Du W, Wang L, Ding C and Liu S (2023), Research on the multi-timescale optimal voltage control method for distribution network based on a DQN-DDPG algorithm.
Front. Energy Res. 10:1097319.
doi: 10.3389/fenrg.2022.1097319

Research on the multi-timescale optimal voltage control method for distribution network based on a DQN-DDPG algorithm

Ming Ma[1,2], Wanlin Du[2], Ling Wang[2], Cangbi Ding[3]* and Siqi Liu[4]

[1]School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, [2]Key Laboratory of Power Quality of Guangdong Power Grid Co., Ltd., Electric Power Research Institute of Guangdong Power Grid Co., Ltd., Guangzhou, China, [3]College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China, [4]College of Automation and College of Artificial Intelligence, Nanjing University of Post and Telecommunications, Nanjing, China

A large number of distributed generators (GDs) such as photovoltaic panels (PVs) and energy storage (ES) systems are connected to distribution networks (DNs), and these high permeability GDs can cause voltage over-limit problems. Utilizing new developments in deep reinforcement learning, this paper proposes a multi-timescale control method for maintaining optimal voltage of a DN based on a DQN-DDPG algorithm. Here, we first analyzed the output characteristics of the devices with voltage regulation function in the DN and then used the deep Q network (DQN) algorithm to optimize the voltage regulation over longer times and the deep deterministic policy gradient (DDPG) algorithm to optimize the voltage regulation mode over short time periods. Second, the design strategy of the DQN-DDPG algorithm as based on the Markov decision process transformation was presented for the stated objectives and constraints considering the state of ES charge for prolonging the energy storage capacity. Lastly, the proposed strategy was verified on a simulation platform, and the results obtained were compared to those from a particle swarm optimization algorithm, demonstrating the method's effectiveness.

KEYWORDS

distributed photovoltaic, deep reinforcement learning, voltage control, multi-timescale, distribution network

1 Introduction

Because of fluctuations in the output and the intermittent nature of DGs, connecting them to light load DNs such as in mountainous areas will cause periodic overvoltage problems in the whole feeder (Impram et al., 2020; Dai et al., 2022). Similarly, the problem of periodic undervoltages will occur when DGs are connected to a heavy-duty DN in an area with major industry production. Traditional voltage control devices, such as on-load tap changers (OLTCs), distributed static synchronous compensators, and switch capacitors can mitigate the overvoltage problem to a certain extent (Kekatos et al., 2015; Zeraati et al., 2019). However, because of the mechanical losses and slow response times, traditional voltage regulation devices cannot prevent voltage problems quickly in real time. At the same time, the frequent regulation may greatly shorten the service life of equipment and affect the voltage quality of the whole DN.

For a DN connected to DGs with strong coupling between active and reactive power, it is obviously not possible for a regulation system to only consider reactive power (Hu et al., 2021; Wang et al., 2021). To ensure safe and stable operation of the DN, both active and reactive power should be taken into account in the control link. Le et al. (2020) adopted different

operational modes for DGs, based on the exchange power between the DN and the external power grid. They considered the capacity utilization rate and power factors as consistent controlling variables to adjust the parameters of the control algorithm, so as to achieve cooperative optimization of voltage and power of the DN. Li et al. (2020) and Zhang et al. (2020) aimed at reducing deviations in reactive power distribution by decreasing the dependence on the transmission of voltage information and adopting an event-triggered consistency control method for distributed voltage control with multiple DG units. Based on active voltage sensitivity, Gerdroodbari et al. (2021) changed the parameters of the active voltage control method of PV inverters in the DN, which improved the regulation of each active PV power reduction (Feng et al., 2018).

In recent years, reinforcement learning, as a type of artificial intelligence technology, has been widely used in smart grids. It has the advantage of not relying on any analytical formula, and it uses a large number of existing data points to produce a mathematical model and generate approximate solutions for grid control. Shuang et al. (2021) used Deep Q network agents and actor-critic agents simultaneously to coordinately control different reactive devices and optimize reactive power online. This method has good robustness and does not depend on communication technology. In contrast to the method of Shuang et al. (2021), Zhang et al. (2021) adopted the DQN algorithm and DDPG algorithm. The DQN-DDPG algorithm was employed in this paper, but we also considered whether the DGs and the reactive voltage regulation equipment were connected as variables for optimizing the active and reactive power. Zhang et al. (2021) did not take into account the effects of active DPV power reduction on voltage regulation of the DN. Liu et al., (2021) and Zhou et al. (2021) proposed a scheduling scheme for an ES system on a DN based on deep reinforcement learning with high permeability DPV access to reduce voltage deviations.

The aforementioned researchers mainly focused on DN regulation using new types of voltage regulation equipment, while ignoring the effects of traditional, stable voltage controllers (SVCs) such as the online tap changer (OLTC) on regulation of the system. Since there have been a large number of traditional voltage regulation devices used in practical DN engineering, this work focused on both the traditional and the new voltage regulation equipment such as DPV and ES in an active DN based on the different response characteristics of each device. We took advantage of the DQN and DDPG algorithms, which can handle discrete variables and continuous variables, respectively, to efficiently and reliably deal with off-limit voltage problems in the DN. At the same time, it is necessary to consider the voltage control method of centralized coordination and distributed cooperation from the perspective of the multi-terminal cooperation of various types of voltage regulation devices.

This paper proposes a multi-timescale method based on the DQN-DDPG algorithms for optimal voltage control in a DN. The DQN algorithm and the DDPG algorithm were used to train the dynamic responses of the different voltage regulators in the framework of the proposed deep reinforcement learning algorithm. Converting the mathematical model of voltage control into a Markov decision process allowed us to decrease the difficulty involved in modeling the several different types of voltage regulation devices. This allowed us to achieve control over long timescales by using OLTC to adjust the average domain voltage of the whole DN; DGs and other devices were used to control local nodes cooperatively over short timescales. Lastly, an IEEE 33-node DN system was constructed on a MATLAB simulation platform and compared with a traditional PSO (particle swarm optimization) algorithm. The proposed control strategy resulted in a faster calculation speed and higher calculation accuracy.

2 Multi-timescale voltage coordination control framework

In order to solve the time period and intermittent voltage overlimit problems caused by high permeability DPVs on the DN, we proposed a voltage control strategy with cooperation among multi-terminal DGs. In a DN with different types of voltage regulation equipment, OLTCs belong to the slower type of discrete regulation devices, while DPVs, ES, and SVCs are continuously active devices, adjusting time to second grade. Therefore, multi-timescale control of the active and reactive power outputs of DPVs, the outputs of ES and SVCs, and the output and network-end OLTC split-regulation were proposed to effectively regulate the voltage of large-scale DN bus nodes.

A centralized coordination controller (CCC) was configured for the OLTC in this paper and was divided into different control regions according to the location of the devices on the branch. Each region was configured with a distributed cooperative controller (DCC), which was regarded as a centralized cooperative agent (CCA) and a distributed cooperative agent (DCA), respectively. The CCC was used to adjust the OLTC splitter and the power and output distribution of the nodes in the region. The DCA and CCA communicated with each other and shared node information in the region. The connection diagram of the centralized coordination-distributed cooperative control method in a DN is shown in Figure 1.

The DCA collects the voltage information of each node in the regional DN and the power information of the incorporated voltage regulation equipment. The voltage unit value of each node can be calculated by Eq. 1:

$$v_i = v_{i\_r} \times v_{n,i}^{-1}, \tag{1}$$

where v_{i_r} is the voltage value of the bus node, i; $v_{n,i}$ is the nominal voltage of the i – th bus node; the superscript "−1" means the bottom form; v_i is the voltage per-unit value of the i – th bus node. If the voltage of some bus nodes exceeds the voltage safety threshold, each DCA sends the value of the voltage standard to the CCA. The safety threshold is set at [.95, 1.05]p.u. After receiving the voltage information of all nodes, the CCA calculates the average unit value of the voltage standard, and if this value exceeds the set feeder threshold range, the OLTC splitter needs to be adjusted, and the feeder threshold range set at [.95, 1.05]p.u. Then, the DQN algorithm is used to obtain the optimal gear position of the OLTC splitter, which ensures that the unit value of the average voltage is kept within the safety threshold.

The aforementioned adjustment method can only ensure that the average standard unit value of voltage reaches the safety threshold. If the voltage of some bus nodes still exceeds the safety threshold range after adjusting the OLTC splitter, the power coordination control strategy based on the DPV, SVC, and ES output characteristics is adopted. The generalized node-based partitioning method is used to divide the control region of the DN (Zhang et al., 2014). In the region where the bus nodes are located, deep reinforcement learning is used to train the optimal power regulation sequence by coordinating the active and reactive power outputs of DPVs, the reactive power output of SVCs, and the

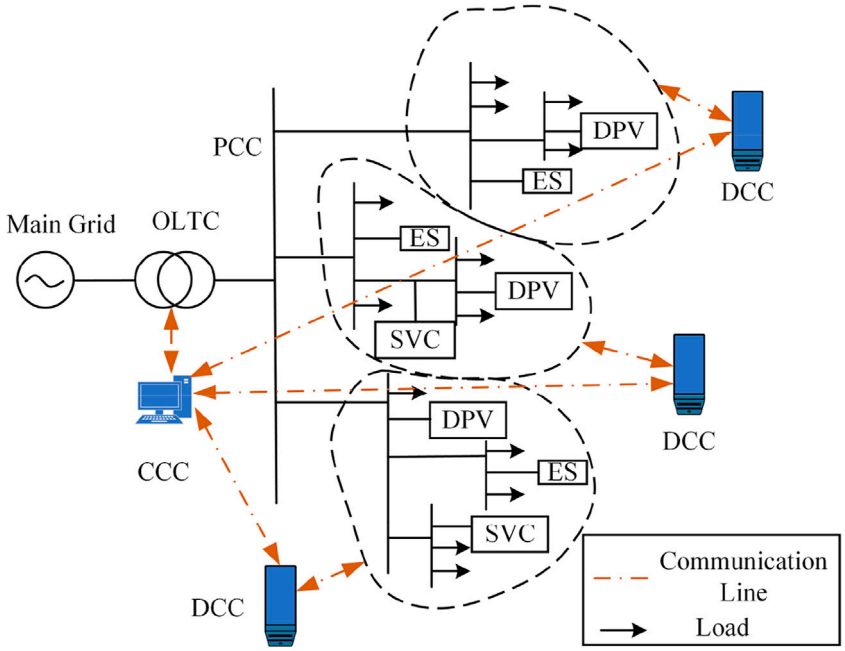


FIGURE 1
Connection diagram of the proposed control method.

ES output (Amir et al., 2022). The CCC determines the optimal control strategy and then issues commands to the CCC communicating with the area to adjust the power output of the inverter of each voltage regulation device. After receiving the instructions for executing the adjustments, the inverters guarantee that the voltage of each node is kept within the safety threshold, reducing the problem of bus node voltage overlimit. In the process of voltage regulation, the DPVs follow the principle that reactive power control voltage is first regulated before active power control voltage is cut, so as to give full play to the absorption capability of the PVs. According to the response times of DPVs, ES, and SVCs over the short timescale (Liu et al., 2022), the specific regulation priority is: OLTC > PV reactive power, ES, and SVC > PV active power. If it is necessary to reduce the active power of the DPVs, the active power reduction of a PV shall not exceed half of the DPV output active power. The overall control process is shown in Figure 2.

3 Modeling of voltage regulation devices on a DN

3.1 DPV inverter

As used in this paper, the DPV inverter adopts PQ control, and its controller is divided into inner and outer rings. The outer ring tracks the DC side of the active power output $p_{dc,ref}$ and the reference value of the reactive power output $q_{dc,ref}$, while the inner ring generates the SPWM modulation signal. The active and reactive power is calculated according to the output current and voltage after dq conversion, and the voltage component of the shaft is obtained by PI control. Lastly, the output voltage of the inverter is obtained by voltage modulation (Atia et al., 2016; Vinnikov et al., 2018). The PQ control block diagram of a DPV inverter is shown in Figure 3.

The active power output and the reactive power output of a model DPV inverter can be calculated as follows:

$$\begin{cases} \Delta i_{od} = \dfrac{1}{T_{in}^{p}s+1}\left(k_p^p + \dfrac{k_i^p}{s}\right)\Delta P_{ref}^{PV} - \Delta P_{PV}, \\ \Delta P_{PV} = \dfrac{3u_{od}}{2}\Delta i_{od}, \end{cases} \tag{2}$$

$$\begin{cases} \Delta i_{oq} = \dfrac{1}{T_{in}^{q}s+1}\left(k_p^q + \dfrac{k_i^q}{s}\right)\Delta Q_{ref}^{PV} - \Delta Q_{PV}, \\ \Delta Q_{PV} = -\dfrac{3u_{od}}{2}\Delta i_{oq}, \end{cases} \tag{3}$$

where Δi_{od} and Δi_{oq} are the differences between the components on axis d and q of the current and the previous time, respectively; ΔP_{PV} and ΔQ_{PV} are the differences between the current active power output and reactive power at the last time, respectively; and ΔP_{ref}^{PV} and ΔQ_{ref}^{PV} are the differences between the current active power output reference and the reactive power reference at the previous time, respectively. Thus, as long as the reference values of the active and reactive power outputs of the DPV inverter are adjusted, any changes in the grid-connected active and reactive components can be controlled.

3.2 OLTC

The OLTC regulates the voltage of the secondary side of the transformer by adjusting the location of the transformer connector, changing the ratio and the distribution of reactive power in the DN line (Wu et al., 2017). In this paper, the regulation of the on-load OLTC by

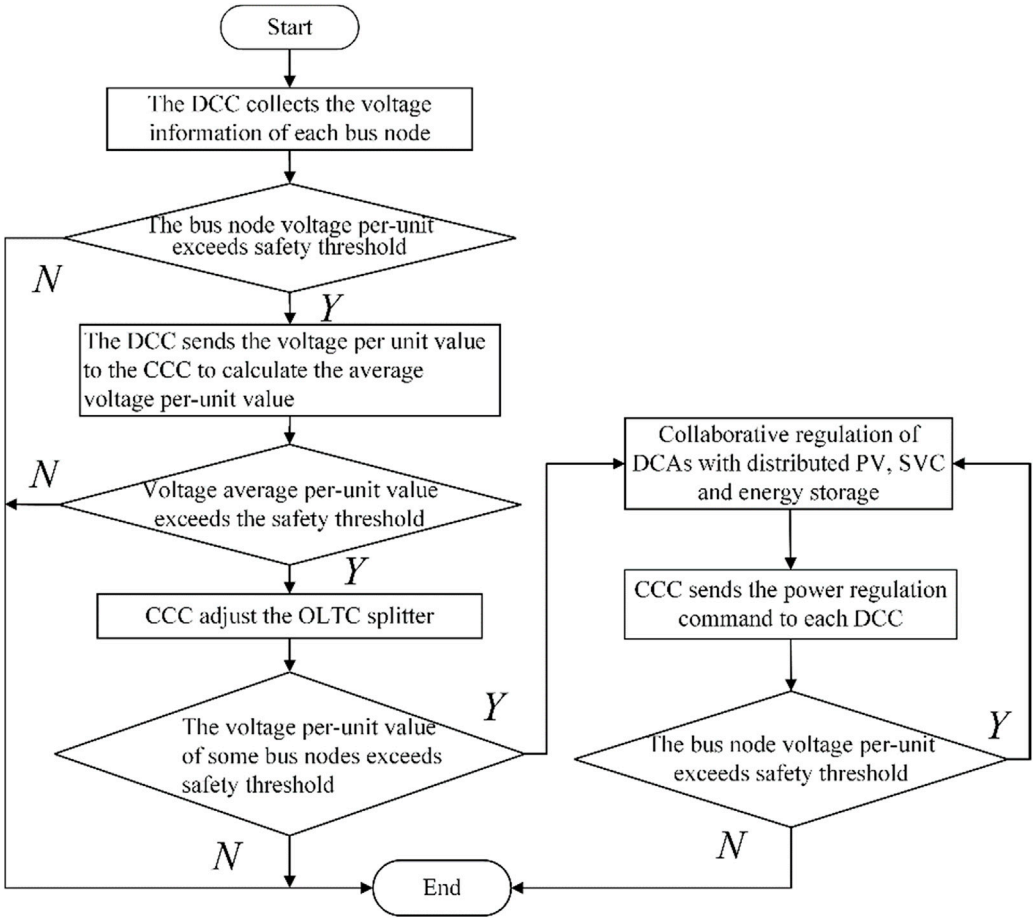


FIGURE 2
Overall framework of the cooperative voltage control strategy.

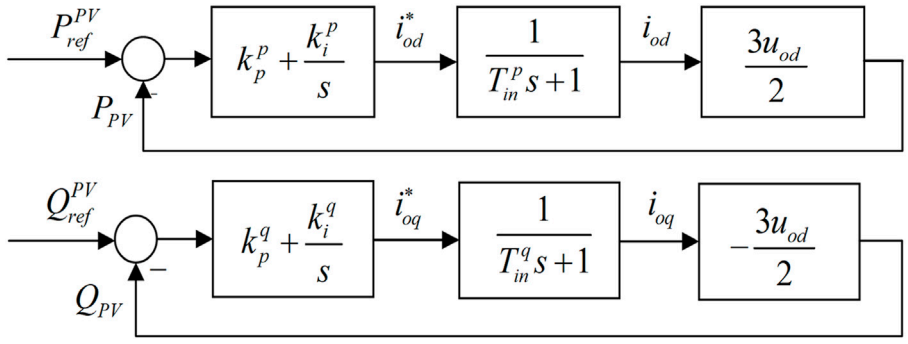


FIGURE 3
PQ control block diagram of the DPV inverter.

the discrete variable ratio was used to control the voltage value of the secondary side of the transformer to keep it within the allowable range during operation. The adjusting process of the splitter is as follows:

$$e = V_1 - V_{ref}, \tag{4}$$

$$t(\tau + 1) = t(\tau) + \Delta T_\tau, \tag{5}$$

$$f(e,t) = \begin{cases} 1 & e > \frac{\varepsilon}{2}, t \geq T_d, \\ -1 & e < -\frac{\varepsilon}{2}, t \geq T_d, \\ 0 & \text{other}, \end{cases} \tag{6}$$

$$n(t+1) = n(t) - d \cdot f(e(t), t), \tag{7}$$

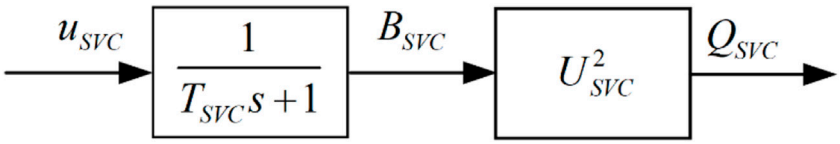


FIGURE 4
Equivalent transfer function of the control loop in the SVC.

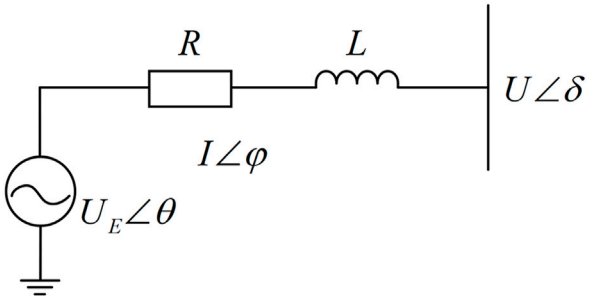


FIGURE 5
Charging and discharging schematic diagram of the ES battery.

where e is the difference between the voltage value V_1 of the OLTC secondary side and the reference value V_{ref}; t is a discrete moment during OLTC operation; τ is a counter; and ΔT_τ is a constant determined by OLTC characteristics and voltage drop. After the OLTC has started to run from time t, if the time is $\geq\Delta T_\tau$, the counter will add one forward. ε is the voltage dead zone to avoid unnecessary actions when OLTC is operating within the permissible voltage range; n is the location of the transformer splitter; d is the number of gears changed by the OLTC splitter; and T_d is the action delay time during OLTC operation.

3.3 SVC

The SVC used in this paper is a thyristor-controlled reactor model, and the control diagram is shown in Figure 4. The SVC is connected to the DN through an inverter, and the equivalent transfer function of the control loop of the inverter in reactive power control mode is given by the following formula (Chen et al., 2018):

$$\begin{cases} \Delta B_{SVC} = \dfrac{1}{sT_{SVC}+1}\Delta u_{SVC}, \\ \Delta Q_{SVC} = \Delta B_{SVC}U_{SVC}^2, \end{cases} \tag{8}$$

where ΔB_{SVC} is the difference between the current and the equivalent susceptance at the previous time; T_{SVC} is the time constant of the control loop; Δu_{SVC} is the difference between the current and the control variable at the previous time; U_{SVC} is the output voltage of the SVC inverter; and ΔQ_{SVC} is the difference between the current and the reactive power output of the previous time. It can be seen that SVC changes the SVC equivalent susceptance in the access system by controlling the trigger angle of the thyristor so as to adjust the reactive power.

3.4 ES battery

The ES system can be equivalent to a voltage source. Figure 5 shows the charging and discharging schematic diagram of the ES system (Yang et al., 2014; Gush et al., 2021), where U_E and U are, respectively, the voltage amplitude of the ES and the voltage amplitude of the connection point, and the voltage phase angles are θ and δ, respectively. I is the amplitude of the current injected into the grid by the ES, and the current phase angle is φ; R and L are the resistance and induction in series in the line, respectively.

ES uses SOC to represent the actual capacity of the battery at a certain time, and the ratio of the residual current C_r to the stable capacity C_{ba} is expressed as follows:

$$SOC_t = \left(\frac{C_R}{C_{ba}}\right) \times 100\%. \tag{9}$$

SOC changes during the charging and discharging process as follows:

Battery discharging:

$$SOC(t) = SOC(t-1) - \frac{P(t)\cdot\Delta t}{C_{ba}\cdot\eta_{dis}}. \tag{10}$$

Battery charging:

$$SOC(t) = SOC(t-1) + \frac{P(t)\cdot\Delta t\cdot\eta_c}{C_{ba}}, \tag{11}$$

where $P(t)$ is the ES output power at the $t-$ th time; η_{dis} and η_c are, respectively, the discharge efficiency and charging efficiency of ES; $SOC(t-1)$ is the ES SOC at the last moment; Δt is the time interval. The power generated by ES and DPV needs to be converted from DC to AC through the inverter before injection into the power grid and is subject to PQ control. Therefore, as long as the reference values of the active and reactive power outputs of the ES inverter are adjusted, the changes in grid-connected active and reactive components can be controlled.

4 DQN-DDPG algorithm

4.1 Principles of the DQN algorithm

The DQN algorithm uses an experience playback mechanism, which stores the experience data (s_t, a_t, r_t, s_{t+1}) obtained at each time

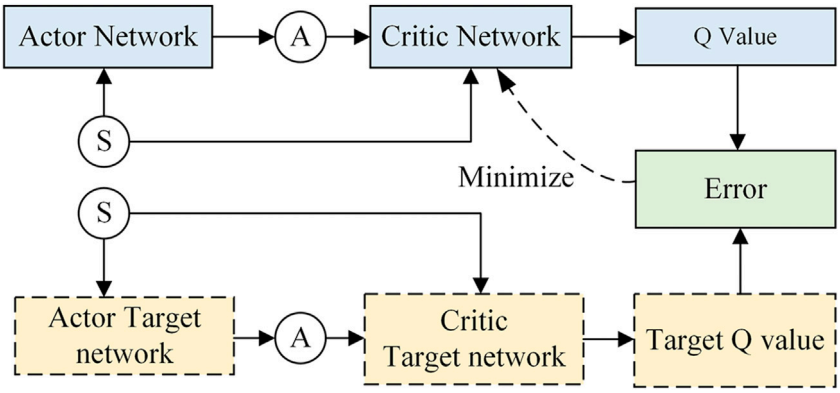


FIGURE 6
Diagram of the relationship between the actor and critic networks.

point in the interaction process between the agent and the environment into the experience pool, and then randomly samples from the experience pool to reduce the correlation of the training data (Labash et al., 2020) so that it is easier to converge the agent training. The DQN algorithm establishes a separate target Q network and updates the weight parameter, θ, of the Q network by constantly approximating the output $r + \gamma \max_{a'} Q_\pi(s', a'; \theta')$ of the target Q network and the output $Q_\pi(s, a; \theta)$ of the Q network. The loss function is defined and the weight parameter, θ, is used to represent the mean squared error loss:

$$L(\theta) = \underset{i \in M}{E}\left[(r_i + \gamma \max_{a'} Q_\pi(s_i', a'; \theta') - Q_\pi(s_i, a_i; \theta))^2\right], \tag{12}$$

where $E(\cdot)$ is the expected value; r_i is the immediate reward of the $i-$th group of data randomly sampled from the experience pool; γ is the discount factor; and θ' is the weight parameter of the target Q-network. The parameter, θ, in the loss function is updated by the gradient descent method, and the algorithm expression is as follows:

$$\theta_{t+1} = \theta_t + \alpha\left[r + \gamma \max_{a'} Q_\pi(s', a'; \theta') - Q(s, a; \theta)\right]\nabla Q_\pi(s, a; \theta), \tag{13}$$

where ∇ is the gradient; θ_t and θ_{t+1} are the Q-network parameters at the $t-$th time and the $(t+1)-$th time, respectively; α is the step length; and r is the value of the reward obtained. To ensure that the agent can simultaneously explore the unknown environment and the obtained environment information, an $\varepsilon - greedy$ strategy is adopted to select the actions of the Q-network:

$$\begin{cases} \text{select at random from A} & \beta < \varepsilon, \\ \arg\max_{a \in \mathbf{A}} Q_\pi(s, a; \theta) & \beta \geq \varepsilon, \end{cases} \tag{14}$$

where $\varepsilon \in [0, 1]$ is a constant and $\beta \in [0, 1]$ is a random number.

4.2 Principles of the DDPG algorithm

The DDPG algorithm was based on and developed from the DQN algorithm. It mainly uses an actor network to make up for the shortcoming that DQN cannot deal with continuous control problems. The DDPG is an algorithm based on the 'actor-critic' architecture to obtain the optimal control sequence. In the actor-critic architecture, the actor network takes the state vector as the input, and the action vector as the output. The critic network takes the state and the action vector as the input, and the estimated Q value is the output (Sutton and Barto, 2018; Qin et al., 2022). The output of the network obtains the maximum value of Q. At the same time, the actor target network and the critic target network are established to output the target Q value, and the optimization training is completed by minimizing the difference between the target Q values and the target Q value. The relationship between actor and critic networks is shown in Figure 6.

The actor network obtains the current state from the environment and outputs a definite action $\mu(s_t \mid \theta^\mu)$ through the deterministic policy gradient method, where θ^μ is the weight coefficient of the actor network. However, one disadvantage is that the environment cannot be fully explored because of the small amount of sampled data, so random noise N_t is introduced in the output action. The *OU* random process in this algorithm is selected to stimulate the ability of the agent to explore the optimal policy in the environment. The DDPG algorithm also uses the experience playback mechanism to store the experience data (s_t, a_t, r_t, s_{t+1}) into the experience pool, so that the W groups of experience data sampled randomly are trained.

The algorithm updates the parameters of the critic network by minimizing the loss function $L = \frac{1}{W}\sum_i (y_i - Q(s_i, a_i \mid \theta^Q)^2)$, where θ^Q is the weight parameter of the critic network, $Q(s_i, a_i \mid \theta^Q)$ is the output Q value of the critic network, $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_i + 1 \mid \theta^{\mu'}); \theta^{Q'})$ is the long-term cumulative reward, $\theta^{Q'}$ is the weight parameter of the critic target network, and $\theta^{\mu'}$ is the weight coefficient of the actor target network. $Q'(s_{i+1}, \mu'(s_i + 1) \mid \theta^{Q'})$ is the target Q value of the output of the critic target network, and r_i is the immediate reward obtained by the critic target network. The algorithm updates the parameters of the actor network through the policy gradient method:

$$\nabla_{\theta^\mu} J \mid_{s_i} \approx \frac{1}{W}\sum_i \nabla_a Q\left(s, a \mid \theta^Q\right)\Big|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu}\mu(s \mid \theta^\mu)\Big|_{s_i}. \tag{15}$$

Finally, by a $soft\ update$, which updates the parameters of the critic target network and the actor target network,

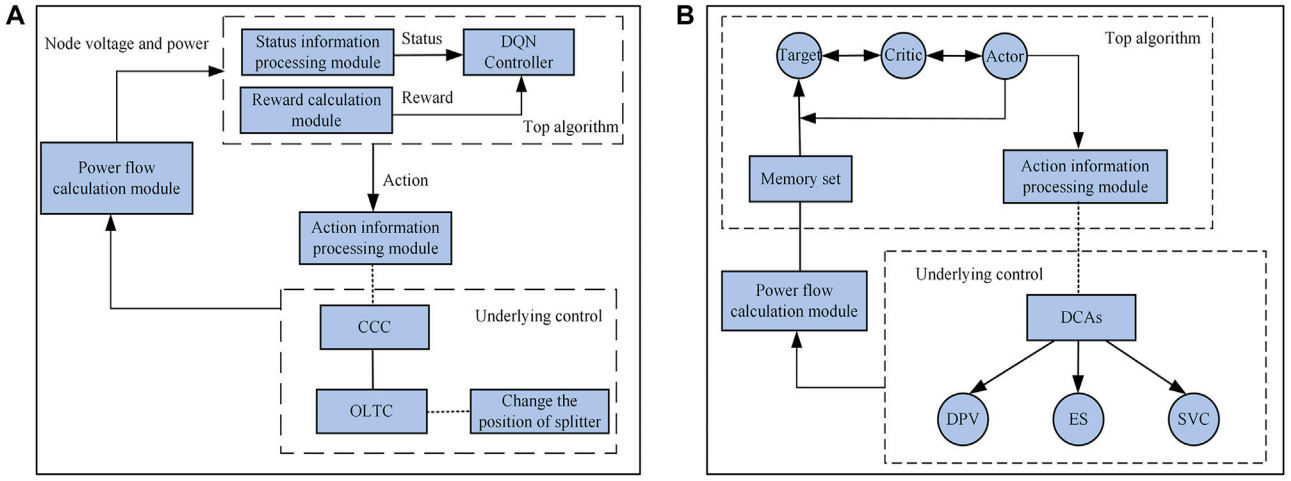


FIGURE 7
Voltage control architecture based on DQN-DDPG; **(A)** DQN algorithm; **(B)** DDPG algorithm.

$$\begin{cases} \theta_t^{Q'} = \eta\theta_t^{Q} + (1-\eta)\theta_{t-1}^{Q}, \\ \theta_t^{\mu'} = \eta\theta_t^{\mu} + (1-\eta)\theta_{t-1}^{\mu}, \end{cases} \tag{16}$$

where η is the divergence factor, $0 < \eta < 1$; θ_t^Q and $\theta_t^{Q'}$ are the critic network parameters and the critic target network parameters at the $t-$ th time, respectively; and θ_t^{μ} and $\theta_t^{\mu'}$ are the actor network parameters and the actor target network parameters at the $t-$ th time, respectively.

5 Voltage cooperative control method for a DN based on the DQN-DDPG algorithm

Combined with the multi-time scale voltage coordination control framework described in the second part, the voltage cooperative control architecture based on the DQN-DDPG algorithm can be divided into a top algorithm layer and a bottom control layer, as shown in Figure 7.

5.1 Mathematical model of voltage control

In adjusting the OLTC splitter, the control objective is to minimize the average voltage exceedances of the bus nodes in the whole DN:

$$\min F(x) = \begin{cases} \bar{U} - 1.05\bar{U}_N, & \bar{U} \geq 1.05\bar{U}_N, \\ 0.95\bar{U}_N - \bar{U}, & \bar{U} \leq 0.95\bar{U}_N, \end{cases} \tag{17}$$

where $\bar{U}$ is the average voltage of bus nodes in the whole DN; $\bar{U}_N$ is the average voltage rating of the DN; N is the number of bus nodes in the whole DN; and ± 5% is the safety threshold range of average bus node voltage per unit value set. When adjusting the voltage of some nodes in the control area, the control objective is to minimize the over-limit values of the bus node voltage in the control area:

$$\min F(x) = \begin{cases} \sum_{i=1}^{M} (U_i - 1.05U_N), U_i \geq 1.05U_N, \\ \sum_{i=1}^{M} (0.95U_N - U_i), U_i \leq 0.95U_N, \end{cases} \tag{18}$$

where U_i is the node voltage of the $i-$ th node; U_N is the rated voltage of the DN; M is the number of DN bus nodes in the control region; and ± 5% is the maximum safe voltage range.

The constraints are as follows:

1) Power flow constraints

$$\begin{cases} \sum_{i=1}^{M} P_{i,l}(t) + P_{loss}(t) = P_M(t) + P_{i,PV}(t) + P_{i,ES}(t), \\ \sum_{i=1}^{M} Q_{i,l}(t) + Q_{loss}(t) = Q_M(t) + Q_{i,PV}(t) + Q_{i,SVC}(t), \end{cases} \tag{19}$$

where $P_{i,l}(t)$ and $Q_{i,l}(t)$ are the active power and reactive power consumed by the load at the $t-$ th time of the $i-$ th node, respectively; $P_{loss}(t)$ and $Q_{loss}(t)$ are the active and reactive power losses of the DN line at the $t-$ th time, respectively; $P_M(t)$ and $Q_M(t)$ are the active power and reactive power emitted by the main network at the $t-$ th time, respectively; $P_{i,PV}(t)$ and $Q_{i,PV}(t)$ are the active power output and the reactive power output of the DPV at the $t-$ th time, respectively; $P_{i,ES}(t)$ is the ES active power output at the $t-$ th time of the $i-$ th node; and $Q_{i,SVC}(t)$ is the reactive power output of SVC at the $t-$ th time of the $i-$ th node.

2) ES output constraints

$$\begin{cases} P_{i,\min}^{ES}(t) \leq P_i^{ES}(t) \leq P_{i,\max}^{ES}(t), \\ \Delta P_{i,\min}^{ES}(t) \leq \Delta P_i^{ES}(t) \leq \Delta P_{i,\max}^{ES}(t), \end{cases} \tag{20}$$

where $P_i^{ES}(t)$ and $\Delta P_i^{ES}(t)$ are the ES active power outputs and active power output increases at the $t-$ th time of the $i-$ th node, respectively; $P_{i,\min}^{ES}(t)$ and $P_{i,\max}^{ES}(t)$ are the upper and lower limits of the ES active power outputs at the $t-$ th time of the $i-$ th node, respectively; $\Delta P_{i,\min}^{ES}(t)$ and $\Delta P_{i,\max}^{ES}(t)$ are the active power output increases at the upper and lower limit at the $t-$ th time of the $i-$ th node, respectively.

3) SOC constraints of ES

$$\left| SOC_i - SOC_{ref} \right| \leq \varepsilon_{ES}, \tag{21}$$

where SOC_{ref} and SOC_i are the reference values for the ES SOC and the SOC at the i – th node, respectively, and ε_{ES} is the convergence error threshold of the ES SOC consistency protocol.

4) SVC output constraints

$$\begin{cases} Q_{i,\min}^{SVC}(t) \le Q_i^{SVC}(t) \le Q_{i,\max}^{SVC}(t), \\ \Delta Q_{i,\min}^{SVC}(t) \le \Delta Q_i^{SVC}(t) \le \Delta Q_{i,\max}^{SVC}(t), \end{cases} \tag{22}$$

where $Q_i^{SVC}(t)$ and $\Delta Q_i^{SVC}(t)$ are the SVC active power outputs and active power output increases at the i – th node, respectively; $Q_{i,\min}^{SVC}(t)$ and $Q_{i,\max}^{SVC}(t)$ are the lower and upper limits of SVC reactive output at the i – th node, respectively; $\Delta Q_{i,\min}^{SVC}(t)$ and $\Delta Q_{i,\max}^{SVC}(t)$ are the SVC active power output increases at the lower and upper limits at the t – th time of the i – th node, respectively.

5) Regulation constraints of the OLTC splitter

$$D_{\min} \le D \le D_{\max}, \tag{23}$$

where D is the splitter position of the OLTC, and $D_{\min}$ and $D_{\max}$ are the lower and upper limits of the OLTC splitter tap, respectively.

5.2 Design of the DQN-DDPG algorithm

5.2.1 Long timescale DQN algorithm

5.2.1.1 State space

The objective of the voltage control strategy in this paper was to modulate the voltage of the bus nodes. The power input of each node needs to be monitored in real time, and the OLTC splitter needs to be adjusted accordingly. Therefore, the state space of the DQN algorithm was defined as the voltage of each bus node in the DN:

$$S_{DQN} = \{v_1, v_2, \cdots, v_i, \cdots v_N\}, \tag{24}$$

where v_i is the per-unit value of the i – th node voltage, $1 \le i \le N$.

5.2.1.2 Action space

The voltage amplitude is changed by changing the tap position of the OLTC splitter, so the action space A_{DQN} of the DQN algorithm is defined from this tap position, and it is assumed that the OLTC splitter has n tap positions and the adjustment range is $\pm\, n \times 1\%$p.u., and each tap position is adjusted as $i \times 1\%$p.u., $-n \le i \le n$. The expression is as follows:

$$A_{DQN} = \{-n \times 1\%\text{p.u.}, -(n-1) \times 1\%\text{p.u.}, \cdots, 0\%\text{p.u.}, \cdots, (n-1) \times 1\%\text{p.u.}, n \times 1\%\text{p.u.}\}. \tag{25}$$

5.2.1.3 Reward function

The centralized cooperative controller calculates the average per-unit value of the voltage after receiving voltage information from the bus nodes in the whole DN. If the average per-unit value of voltage exceeds the safety threshold, that is, [.95,1.05]p.u., it is adjusted by the OLTC splitter. Therefore, the immediate reward function is as follows:

$$r_{DQN,i} = -\alpha_w \Delta \bar{v}_i^{\,2}, \tag{26}$$

where α_w is the weight coefficient and $\Delta\bar{v}$ is the value where the average per-unit value of the voltage of the bus nodes in the DN exceeds the safety threshold, specifically:

$$\Delta \bar{v}_i = \begin{cases} \bar{v}_i - 0.95, \bar{v}_i > 0.95, \\ 1.05 - \bar{v}_i, \bar{v}_i < 1.05, \end{cases} \tag{27}$$

where $\bar{v}_i$ is the average per-unit value of the voltage at the bus nodes in the whole DN.

5.2.2 Short timescale DDPG algorithm

5.2.2.1 State space

The DDPG algorithm is mainly used to prevent voltage control problems in the DN area where the voltage overlimit nodes are located. Therefore, the DDPG state space is different from that of DQN algorithm, and the power output of each node in this area needs to be collected in real time to follow up on the regulation of power output by the voltage regulation equipment. Therefore, the state space of the DDPG algorithm is defined as follows:

$$S_{DDPG,i} = \{v_1, \cdots, v_i, \cdots, v_M, p_1, \cdots, p_i, \cdots, p_M, q_1, \cdots, q_i, \cdots, q_M\}, \tag{28}$$

where v_i is the per-unit value of the i – th node voltage; p_i is the active power of the i – th node; q_i is the reactive power of the i – th node, $1 \le i \le M$; and M is the number of bus nodes in the DN region.

5.2.2.2 Action space

If the voltage at some nodes still exceeds the safety threshold after adjusting the OLTC splitter according to the DQN algorithm, then the active and reactive power outputs of the DPVs, the ES output, and the reactive power output of the SVCs in the control region where the node is located can be adjusted to control the voltage. The action space of the DDPG algorithm is defined as follows:

$$A_{DDPG,i} = \{A_i^{PV}, A_i^{ES}, A_i^{SVC}\}. \tag{29}$$

When only DPV reactive power regulation is used, $A_i^{PV} = \Delta Q_i^{PV}$. If regulation of the reactive power cannot achieve the desired voltage control, then DPV active power reduction is added and $A_i^{PV} = \{\Delta P_i^{PV}, \Delta Q_i^{PV}\}$; similarly, $A_i^{ES} = \Delta P_i^{ES}$, $A_i^{SVC} = \Delta Q_i^{SVC}$; where ΔP_i^{PV} and ΔQ_i^{PV} are the variations in the active and reactive power outputs of DPVs at the i – th node; ΔP_i^{ES} is the variation in the active power output of ES at the i – th node; and ΔQ_i^{SVC} is the variation in SVC reactive power output at the i – th node. The power output of each voltage regulation device must comply with the corresponding following constraints.

5.2.2.3 Reward function

The DDPG algorithm controls the voltage of each node by adjusting the outputs of the DPVs, ES, and SVCs connected to each node. The control objective is to stabilize the voltage and keep it from exceeding the safety threshold; thus, the immediate reward function is set as the sum of the quadratic form of the higher limit voltage of each node, the active power and reactive power regulation of the DPV output, the active power regulation of the ES output, and the reactive power regulation of the SVC output:

$$\begin{aligned} r_{DDPG,i} = & -\Delta v_{i_v} \cdot B \cdot \Delta v_{i_v}^{\,T} - \eta_s \Delta P_{i_{pv}}^{PV} \cdot C \cdot \Delta P_{i_{pv}}^{PV^T} - \Delta Q_{i_{pv}}^{PV} \cdot C \cdot \Delta Q_{i_{pv}}^{PV^T} \\ & -\Delta P_{i_{es}}^{ES} \cdot D \cdot \Delta P_{i_{es}}^{ES^T} - \Delta Q_{i_{svc}}^{SVC} \cdot E \cdot \Delta Q_{i_{svc}}^{SVC^T} \\ & \forall i_v \in M, \forall i_{pv} \in \mathbb{N}_{PV}, \forall i_{es} \in \mathbb{N}_{ES}, \forall i_{svc} \in \mathbb{N}_{SVC}, \end{aligned} \tag{30}$$

where Δv_i is the voltage limit of the i – th bus node; M is the number of bus nodes in the control region; and $\mathbb{N}_{PV}$, $\mathbb{N}_{ES}$, and $\mathbb{N}_{SVC}$ are the number of DPVs, ES, and SVCs in the control region,

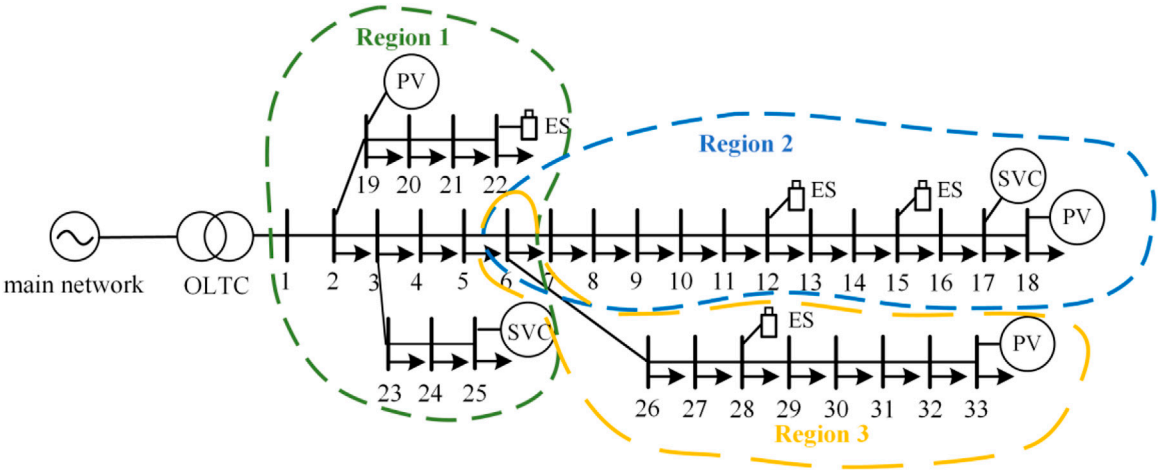


FIGURE 8
DN topology of the IEEE-33 node.

respectively. The matrices B, C, D, and E are all weight matrices; η_s is the selection coefficient, and when DPV active power reduction is not used, $\eta = 0$; when PV active power reduction is added, $\eta = 1$.

5.3 Overall flow of the algorithm

1) The iterative process of OLTC tap adjustment based on DQN is as follows:
   S1: Calculate the average standard per-unit value of voltage according to DN environment. If the average standard value of voltage exceeds the safety threshold, the DQN agent will be trained at the initial position of the OLTC splitter.
   S2: The standard unit value of voltage at each bus node in the DN obtained by the CCC is taken as the set, $s_t = \{v_1, v_2, \cdots, v_N\}$, of the initial states of the DQN agent.
   S3: Select the corresponding action a_t of the OLTC splitter tap from the action set $A_{DQN,i}$ according to the $\varepsilon - greedy$ strategy, and execute the action a_t to obtain the immediate reward according to the reward function, $r_{DQN,i}$. The updated per-unit value of voltage of each bus node obtained by the MATPOWER power flow calculation is the next state set s_{t+1}, and the experience data set, (s_t, a_t, r_t, s_{t+1}), is stored in the experience pool (Zimmerman et al., 2011).
   S4: Sampling the empirical data set randomly from the experience pool according to the sampled data set, (s_j, a_j, r_j, s_{j+1}).
   S5: After updating the state set, the per-unit value of the bus node changes, and it must be recalculated to determine if the average per-unit value of the voltage meets the conditions of the safety threshold [.95,1.05]p.u. If the conditions are met, the iteration will be terminated. The target Q value y_j is substituted into the current immediate reward, r_j. If the condition is not met, the objective Q value y_j is substituted into the long-term cumulative reward value, $r_j + \gamma \max_{a'} Q_\pi(s_{j+1}, a'; \theta')$.
   S6: Use the gradient descent method to update the parameter θ in the loss function.
   S7: Update the parameters, $\theta' = \theta$, of the target Q network at every other iteration.
   S8: Continue to perform S3 until the set of states meets the termination iteration condition.
2) If the standard voltage value of some bus nodes is still beyond the safety threshold [.95,1.05] after adjusting the OLTC splitter by DQN algorithm, the output power of DPVs, ES, and SVCs should be optimized and adjusted by the DDPG algorithm. The specific control process is as follows:
   S1: The per-unit value of the voltage and the power information from all bus nodes in the control region where the overlimit bus node is located are used as the initial state set, $s_t = \{v_1, v_2, \cdots, v_M, p_1, p_2, \cdots, p_M, q_1, q_2, \cdots, q_M\}$, of the DDPG algorithm.
   S2: Select the actions, $a_t = \mu(s_t \mid \theta^\mu) + N_t$, corresponding to the DPV, ES, and SVC output power according to the current strategy and random noise, $A_{DDPG,i}$, and execute the actions a_t to obtain immediate rewards r_t according to the reward function $r_{DDPG,i}$. After MATPOWER power flow calculations, the updated voltage per-unit value, the active power, and the reactive power at each bus node are taken as the next state set, s_{t+1}, and the empirical data set, (s_t, a_t, r_t, s_{t+1}), is stored in the experience pool.
   S3: W groups of empirical data sets are randomly sampled from the experience pool, and the target Q value, $y_i = r_i + \gamma Q'(s_{k+1}, \mu'(s_{k+1} \mid \theta^{\mu'}) \mid \theta^{Q'})$, of the output of the critic objective function is calculated according to the sampled data, (s_k, a_k, r_k, s_{k+1}).
   S4: The critic network parameters are updated by minimizing the loss function, $L = \frac{1}{W}\sum_i (y_i - Q(s_i, a_i \mid \theta^Q)^2)$, and the actor network parameters are updated by the policy gradient method.
   S5: Update the critic target network parameters $\theta^{Q'}$ and the actor target network parameters $\theta^{\mu'}$ according to the $soft\,update$ method.
   S6: If the standard value of the bus node voltage in the control region lies within the safety threshold [.95, 1.05]p.u., the iteration ends. If the conditions are not met, perform Step 3.

TABLE 1 Load and generator parameters of the standard IEEE 33-node DN system.

Node load (MVA)			Generator (MVA)	
Bus number	Active power (MW)	Reactive power (MVar)	Active power (MW)	Reactive power (MVar)
1	0	0	1.5	2.3
2	0.1	.06	0	0
3	.09	.04	0	0
4	.12	.08	0	0
5	.06	.03	0	0
6	.06	.02	0	0
7	0.2	0.1	0	0
8	0.2	0.1	0	0
9	.06	.02	0	0
10	.06	.02	0	0
11	.045	.03	0	0
12	.06	.035	0	0
13	.06	.035	0	0
14	.12	.08	0	0
15	.06	.01	0	0
16	.06	.02	0	0
17	.06	.02	0	0
18	.09	.04	0	0
19	.09	.04	0	0
20	.09	.04	0	0
21	.09	.04	0	0
22	.09	.04	0	0
23	.09	.05	0	0
24	.42	0.2	0	0
25	.42	0.2	0	0
26	.06	.025	0	0
27	.06	.025	0	0
28	.06	.02	0	0
29	.12	.07	0	0
30	0.2	0.6	0	0
31	.15	.07	0	0
32	.21	0.1	0	0
33	.06	.04	0	0

6 The simulation verification

6.1 Example introduction

To verify the effectiveness of the cooperative voltage optimization control method considering the load and storage of the source network proposed in this paper, the improved IEEE 33-node DN system was adopted to perform a simulation analysis on a MATLAB r2019a platform. The improved topology diagram of the DN is shown in Figure 8. The total load was 3715.0 kW + j2300.0kVar, and the rated voltage was 10 kV in the DN. The node loads and generator parameters in the DN

TABLE 2 Branch parameters of the standard IEEE 33-node DN system.

Branch number	Starting node	Ending node	Branch resistance (Ω)	Branch reactance (Ω)
1	1	2	.0922	.0407
2	2	3	.493	.2511
3	3	4	.366	.1864
4	4	5	.3811	.1941
5	5	6	.819	.707
6	6	7	.1872	.6188
7	7	8	.7114	.2351
8	8	9	1.03	.74
9	9	10	1.044	.74
10	10	11	.1966	.065
11	11	12	.3744	.1238
12	12	13	1.468	1.155
13	13	14	.5416	.7129
14	14	15	.591	.526
15	15	16	.7463	.545
16	16	17	1.289	1.721
17	17	18	.732	.574
18	2	19	.164	.1565
19	19	20	1.5042	1.3554
20	20	21	.4095	.4784
21	21	22	.7089	.9373
22	3	23	.4512	.3083
23	23	24	.898	.7091
24	24	25	.896	.7011
25	6	26	.203	.1034
26	26	27	.2842	.1447
27	27	28	1.059	.9337
28	28	29	.8042	.7006
29	29	30	.5075	.2585
30	30	31	.9744	.963
31	31	32	.3105	.3619
32	32	33	.341	.5302

system are shown in Table 1, and the DN branch parameters are shown in Table 2. The OLTC of the DN is located at node 1, its rated capacity is 100MVA, and the adjustment range is $\pm 5 \times 1\%$p.u. The DPVs are located at nodes 18, 19, and 33, and the rated PV capacity is 2.2 MW. The ES are located at nodes 12, 15, 22, and 28 with a rated capacity of 800 kW · h, a maximum charge–discharge power of 400 kW, and a convergence error threshold of .05. The SVCs are located at nodes 17 and 25, and the maximum capacity is 2MVar. The safety threshold of the bus node voltage per-unit value is set as [.95,1.05]p.u. According to the DN control region division method, the DN is partitioned, as shown in Figure 6.

6.2 Analysis of simulation results

The power disturbance was introduced into the DN at a certain time, and the voltage amplitude of each node is shown in Figure 7.

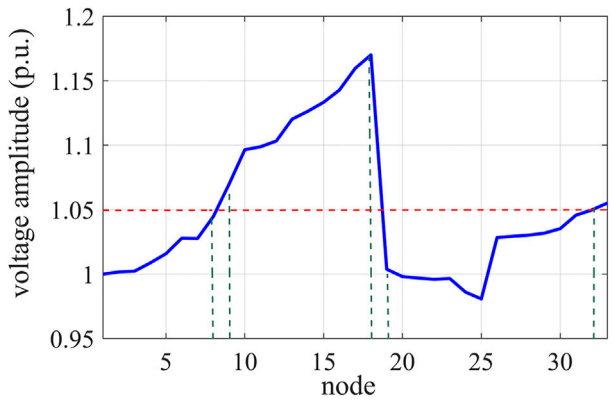


FIGURE 9
Voltage amplitude of each bus node without control.

Overvoltage events occurred at nodes 9–18, 32, and 33 (Figure 9). The DQN algorithm was first designed according to the regulation range of the OLTC:

- State Space. There are 33 bus nodes in the whole-domain DN, and the state space of the DQN algorithm can be obtained by using Eq. 24, $S_{DQN} = \{v_1, v_2, \cdots, v_i, \cdots v_{33}\}$.
- Action Space. The regulation range of the OLTC is ± 4 × 1%p.u. There are 9 splitter taps, the reference per-unit value of bus node is 1, and the action space of the DQN algorithm can be obtained according to Eq. 25:

$$A_{DQN} = \{-4\%\text{p.u.}, -3\%\text{p.u.}, -2\%\text{p.u.}, -1 \times 1\%\text{p.u.}, 0\%\text{p.u.}, 1\%\text{p.u.}, 2\%\text{p.u.}, 3\%\text{p.u.}, 4\%\text{p.u.}\} = \{0.96\text{p.u.}, 0.97\text{p.u.}, 0.98\text{p.u.}, 0.99\text{p.u.}, 1\text{p.u.}, 1.01\text{p.u.}, 1.02\text{p.u.}, 1.03\text{p.u.}, 1.04\text{p.u.}\}. \tag{31}$$

- Reward Function. The purpose of the OLTC splitter is to keep the average per-unit value of the voltage in the DN within the safety threshold range. According to Eq. 26, the reward function of the DQN algorithm can be obtained as follows:

$$r_{DQN} = -\Delta\bar{v}^2 \times 10^5$$
$$\Delta\bar{v} = \begin{cases} \dfrac{v_1 + \cdots + v_i + \cdots + v_{33}}{33} - 0.95 \dfrac{v_1 + \cdots + v_i + \cdots + v_{33}}{33} \geq 0.95, \\ 1.05 - \dfrac{v_1 + \cdots + v_i + \cdots + v_{33}}{33} \dfrac{v_1 + \cdots + v_i + \cdots + v_{33}}{33} < 1.05. \end{cases} \tag{32}$$

The learning rate of the neural network was chosen as .001, the discount coefficient was .99, the capacity of the experience pool was 4,000, and the capacity of the mini-batch was 64. DQN agents were trained with a total of 500 episodes, and each episode was completed after 300 samples were trained. The results of training the agents according to the iterative process of algorithm 4.1 are shown in Figure 10. The episode reward is the cumulative reward value in an episode obtained by the agent during training, and the average reward is the average of the reward values in every four episodes. As can be seen from Figure 8, at the beginning of the training, the reward value was very low due to the limited learning experience. As the training continued, the agents kept exploring and learning, and the reward value kept increasing. After 250 episodes, the reward value of the DQN agent fluctuated within a small range, which indicated that the algorithm gradually converged and the agents' training had developed an optimal strategy for controlling the voltage by adjusting the OLTC splitter.

During the test, the OLTC training data were used as the control. Figure 11 shows the effects of adjusting the OLTC splitter, at which point the OLTC splitter was set at −3 × 1%p.u. It can be seen that the voltage amplitude of each bus decreased, but there were still some bus nodes whose voltage standard value exceeded the safety threshold. Therefore, it was necessary to readjust the voltage through DPV, ES, and SVC regulation in the area where it exceeded the limit.

The DDPG algorithm was then applied according to the DPV, ES, and SVC in the region where the voltage overlimit node is located.

- State Space. Control region 2 contains 13 bus nodes. As can be seen from Figure 11, after adjusting the OLTC splitter, the standard voltage values at nodes 9, 32, and 33 were controlled within the safety threshold, but the voltages at nodes 10–18 were still beyond the upper limit of the safety threshold. According to Eq. 28, the state space of the DDPG algorithm can be obtained as follows:

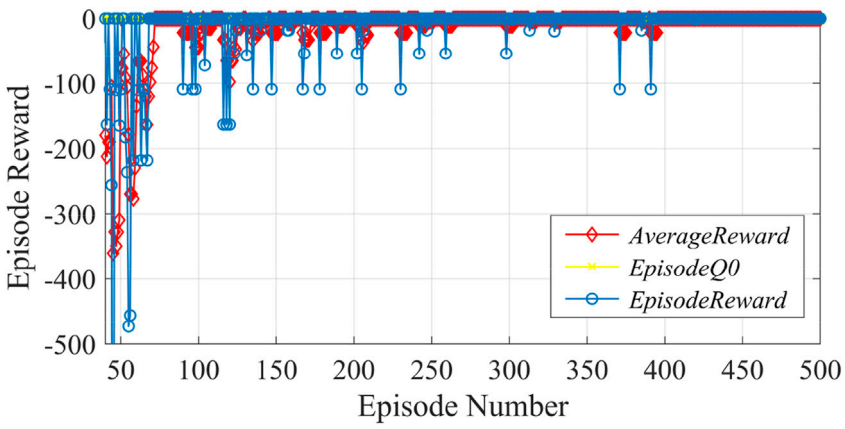


FIGURE 10
DQN agent training process.

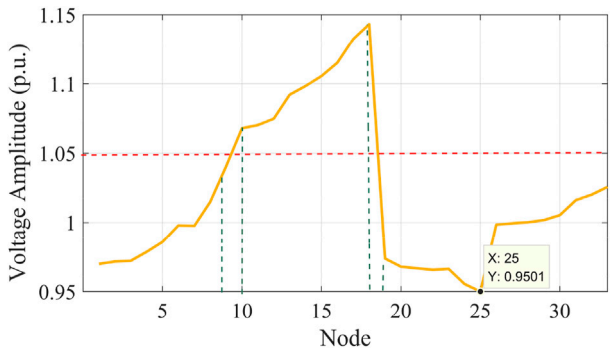


FIGURE 11
Voltage amplitude of each bus node after adjusting OLTC.

$$S_{DDPG} = \{v_1, \cdots, v_i, \cdots, v_{13}, p_1, \cdots, p_i, \cdots, p_{13}, q_1, \cdots, q_i, \cdots, q_{13}\} 1 \le i \le 13. \tag{33}$$

- Action Space. Node 18 incorporates DPV, nodes 12 and 15 are ES, and node 17 is SVC. The DDPG action space can be obtained according to Eq. 29:

$$\begin{cases} A_{DDPG} = \{A_i^{PV}, A^{ES}, A^{SVC}\}, \\ A_i^{PV} = \{\Delta P^{PV}, \Delta Q^{PV}\}, \\ A^{ES} = \{\Delta P_i^{ES}\} 1 \le i \le 2, \\ A^{ES} = \{\Delta Q^{SVC}\}. \end{cases} \tag{34}$$

According to variations in the active power output of DPVs, the adjustable range of the DPV reactive power can be obtained as follows:

$$\Delta Q = \left[-\sqrt{S^2 - P_{PV}^2}, \sqrt{S^2 - P_{PV}^2} \right]. \tag{35}$$

According to the SOC consistency protocol for ES, the SOC reference value is .5, the convergence error threshold is .02, and the charge–discharge efficiency is 80%.

- Reward Function. The DDPG algorithm controls node voltage by regulating the DPVs, ES, and SVCs connected in Region 2. The reward function can be expressed by Eq. 30:

$$r_{DDPG} = -\Delta v_{i_v} \cdot B \cdot \Delta v_{i_v}{}^T - \Delta P^{PV} \cdot C \cdot \Delta P^{PV^T} - \Delta Q^{PV} \cdot C \cdot \Delta Q^{PV^T} - \Delta P_{i_{es}}^{ES} \cdot D \cdot \Delta P_{i_{es}}^{ES^T} - \Delta Q^{SVC} \cdot E \cdot \Delta Q^{SVC^T} 1 \le i_v \le 13, 1 \le i_{es} \le 2, \tag{36}$$

where $B = 100^* \mathbf{I}_{13\times13}$, $D = 10 \times \mathbf{I}_{2\times2}$, and $C = E = 10$ constitute the weight matrix and $\mathbf{I}$ is the identity matrix.

The parameters of the DDPG algorithm for the neural network were set as follows. The learning rate of the actor network was .001, the learning rate of the critic network was .0001, the discount coefficient was .99, the update coefficient was .01, and the capacity of the experience pool was 4,000. The capacity of the mini-batch was selected as 64, and the noise variable was .3. The DDPG agent was trained with 1,000 episodes according to the iterative process of algorithm 4.2 and each episode involved training 300 samples (Figure 12). It can be seen at the beginning of the training that the reward value was very low due to the limited learning experience. As the training continued, the agent kept exploring and learning, and the reward value kept increasing. After 800 episodes, the reward value of the DDPG agent fluctuated very little within a small range, which indicated that the algorithm gradually converged; the agent's exercise training had developed an optimal strategy for voltage control of DPV and ES in the regulation region.

To make full use of the active DPV power consumption capacity, the DPV reactive power and ES were regulated. During the test, the power output training data on the voltage regulation equipment were used as control. Figure 13 shows the effect of controlling nodes 9–18 without reducing the active DPV power. It can be seen that the bus contact voltage in this control region at this time was not regulated within the safety threshold. Thus, DPV reactive power and ES cannot achieve the desired voltage control, so the active power must be further reduced. Figure 14 shows the voltage control effect of nodes 9–18 with active power reduction, and Figure 15 shows the level of power adjustment of each voltage regulation device in this control region.

The DDPG algorithm was used for voltage control for 0.04 s. According to Figures 14, 15, the ES was discharged, and the output active power was about 250 KW. The SVC reactive power output was about 537.3 kVar, the DPV reactive power output was 372.8 kVar, and the active power reduction was less than 200 KW. The consumption capacity of the DPVs in the DN can be improved by reducing the active power reduction of the DPVs as much as possible. By adjusting

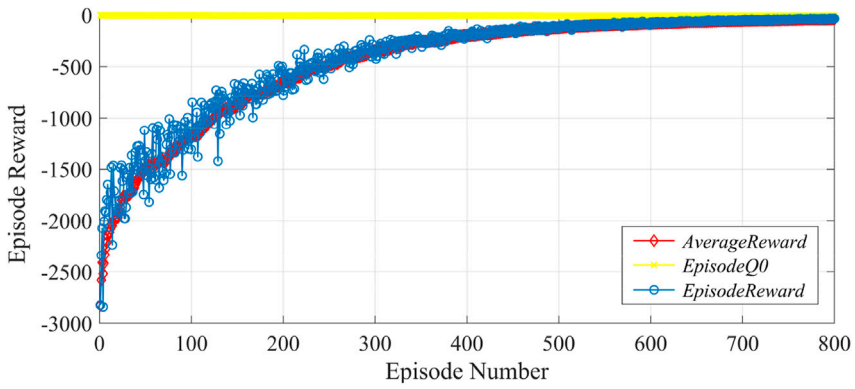


FIGURE 12
DDPG agent training process.

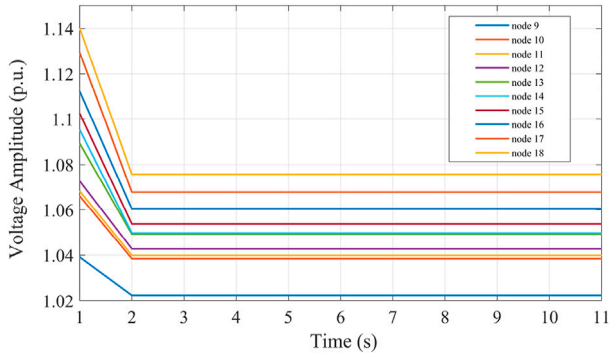


FIGURE 13
Voltage control effect in the case of no active power reduction.

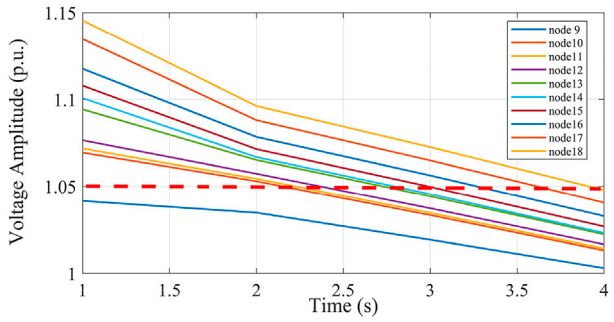


FIGURE 14
Voltage control effect when active power reduction is added.

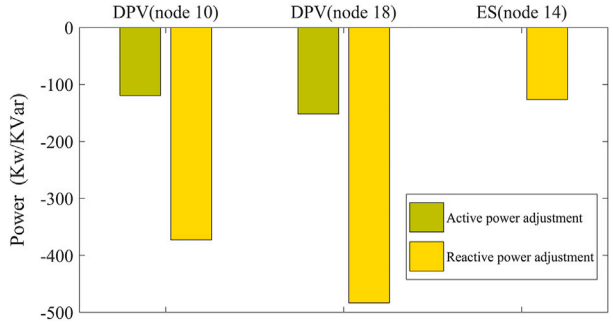


FIGURE 15
Power adjustments of each voltage regulation device.

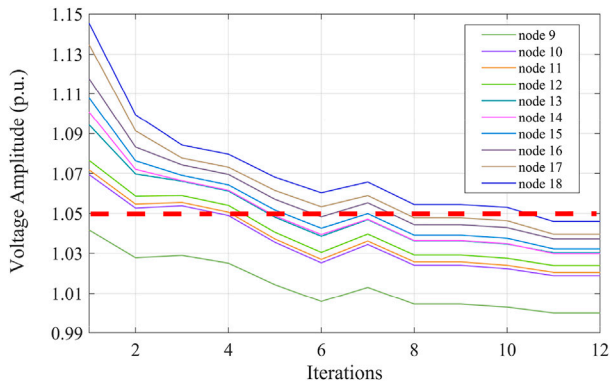


FIGURE 16
Voltage control effects of the PSO algorithm.

the output power of the DPVs and the ES in the control region, the voltage of the bus node can be quickly and effectively maintained within the safety threshold.

To demonstrate the advantages of the rapid calculation speed of the DQN-DDPG algorithm, this paper compared it with the PSO algorithm. Figure 16 shows the effect of testing the voltage control model using the PSO algorithm with a consumption time of 19.52 s, and it is obvious that the DDPG algorithm was able to control voltage more quickly. This is because the PSO must obtain the optimal control strategy through continuous iteration, and the solution process was an iterative process of the objective function, while DDPG achieved the optimal strategy through the exploration and learning of the environment by the agents; the optimal strategy was the trained optimization sequence.

7 Conclusion

In order to give full play to the voltage regulating potential of the high permeability DGs in the DN, a deep reinforcement learning algorithm was used to test the voltage control model of the corresponding DN. The voltage control model was converted to a Markov decision process, and the whole series of steps of the algorithm to improve its design depth according to the objective function and constraint conditions were put forward. By combining the DQN and the DDPG algorithms with deep reinforcement learning, the discrete and continuous variables could be processed simultaneously, and the algorithms could control the DN in real time according to the current state of the power grid. The algorithms were independent of changes in the DN environment, and the optimal strategy was obtained through the exploration and learning of agents in the environment. This method effectively solved the problems of large model dimensions and high data volumes, to complete complex tasks, and achieve cooperative control of different voltage regulation devices.

However, this paper still has some imperfections. When the controller issues voltage regulation instructions to the inverter, there is an unavoidable communication delay, which can affect the real-time performance and effectiveness of the voltage regulation equipment. Also, this paper only considered a DN with OLTC, DPV, ES and SVC access, and did not conduct in-depth research on newer DNs with large-scale access to wind power and hydrogen energy or flexible loads such as electric vehicles.

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

Conceptualization, MM and LW; methodology, CD and WD; software, MM; validation, SL; formal analysis, WD; investigation, CD and MM; resources, LW; data curation, WD; writing and preparation of the original draft, SL and LW; writing—reviewing and editing, LW and CD; visualization, WD; supervision, MM; funding acquisition, SL and WD. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

MM, WD, and LW were employed by the Electric Power Research Institute of the Guangdong Power Grid Co., Ltd.

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.

References

Amir, M., Prajapati, A. K., and Refaat, S. S. (2022). Dynamic performance evaluation of grid-connected hybrid renewable energy-based power generation for stability and power quality enhancement in smart grid. *Front. Energy Res.* 10, 861282. doi:10.3389/fenrg.2022.861282

Atia, R., and Yamada, N. (2016). Sizing and analysis of renewable energy and battery systems in residential microgrids. *IEEE Trans. Smart Grid.* 7 (3), 1204–1213. doi:10.1109/TSG.2016.2519541

Chen, H., Prasai, A., and Divan, D. (2018). A modular isolated topology for instantaneous reactive power compensation. *IEEE Trans. Power Electron.* 33, 975–986. doi:10.1109/TPEL.2017.2688393

Dai, N., Ding, Y., Wang, J., and Zhang, D. (2022). Editorial: Advanced technologies for modeling, optimization and control of the future distribution grid. *Front. Energy Res.* 10, 885659. doi:10.3389/fenrg.2022.885659

Feng, J., Wang, H., Xu, J., Su, M., Gui, W., and Li, X. (2018). A three-phase grid-connected microinverter for AC photovoltaic module applications. *IEEE Trans. Power Electron.* 33, 7721–7732. doi:10.1109/TPEL.2017.2773648

Gerdroodbari, Y. Z., Razzaghi, R., and Shahnia, F. (2021). Decentralized control strategy to improve fairness in active power curtailment of PV inverters in low-voltage distribution networks. *IEEE Trans. Sustain. Energy* 12 (4), 2282–2292. doi:10.1109/TSTE.2021.3088873

Gush, T., Kim, C.-H., Admasie, S., Kim, J.-S., and Song, J.-S. (2021). Optimal smart inverter control for PV and BESS to improve PV hosting capacity of distribution networks using slime mould algorithm. *IEEE Access* 9, 52164–52176. doi:10.1109/ACCESS.2021.3070155

Hu, J., Yin, W., Ye, C., Bao, W., Wu, J., and Ding, Y. (2021). Assessment for voltage violations considering reactive power compensation provided by smart inverters in distribution network. *Front. Energy Res.* 9, 713510. doi:10.3389/fenrg.2021.713510

Impram, S., Varbak Nese, S., and Oral, B. (2020). Challenges of renewable energy penetration on power system flexibility: A survey. *Energy Strategy Rev.* 31, 100539. doi:10.1016/j.esr.2020.100539

Kekatos, V., Wang, G., Conejo, A. J., and Giannakis, G. B. (2015). Stochastic reactive power management in microgrids with renewables. *IEEE Trans. Power Syst.* 30, 3386–3395. doi:10.1109/TPWRS.2014.2369452

Labash, A., Aru, J., Matiisen, T., Tampuu, A., and Vicente, R. (2020). Perspective taking in deep reinforcement learning agents. *Front. Comput. Neurosci.* 14, 69. doi:10.3389/fncom.2020.00069

Le, J., Zhou, Q., Wang, C., and Li, X. (2020). Research on voltage and power optimal control strategy of distribution network based on distributed collaborative principle. *Proc. CSEE.* 40 (04), 1249. doi:10.13334/j.0258-8013.pcsee.182229

Li, Z., Yan, J., Yu, W., and Qiu, J. (2020). Event-triggered control for a class of nonlinear multiagent systems with directed graph. *IEEE Trans. Syst. Man, Cybern. Syst.* 51 (11), 6986–6993. doi:10.1109/TSMC.2019.2962827

Liu, S., Ding, C., Wang, Y., Zhang, Z., Chu, M., and Wang, M. (2021). "Deep reinforcement learning-based voltage control method for distribution network with high penetration of renewable energy," in Proceedings of the in 2021 IEEE Sustainable Power and Energy Conference, Nanjing, China, December 2021, 287–291.

Liu, S., Zhang, L., Wu, Z., Zhao, J., and Li, L. (2022). Improved model predictive dynamic voltage cooperative control technology based on PMU. *Front. Energy Res.* 10, 904554. doi:10.3389/fenrg.2022.904554

Qin, P., Ye, J., Hu, Q., Song, P., and Kang, P. (2022). Deep reinforcement learning based power system optimal carbon emission flow. *Front. Energy Res.* 10, 1017128. doi:10.3389/fenrg.2022.1017128

Shuang, N., Chenggang, C., Ning, Y., Hui, C., Peifeng, X., and Zhengkun, L. (2021). Multi-time-scale online optimization for reactive power of distribution network based on deep reinforcement learning. *Automation Electr. Power Syst.* 45 (10), 77–85. doi:10.7500/AEPS20200830003

Vinnikov, D., Chub, A., Liivik, E., Kosenko, R., and Korkh, O. (2018). Solar optiverter—a novel hybrid approach to the photovoltaic module level power electronics. *IEEE Trans. Industrial Electron.* 66 (5), 38693869–38803880. doi:10.1109/TIE.2018.2850036

Wang, Y., He, H., Fu, Q., Xiao, X., and Chen, Y. (2021). Optimized placement of voltage sag monitors considering distributed generation dominated grids and customer demands. *Front. Energy Res.* 9, 717089. doi:10.3389/fenrg.2021.717089

Wu, W., Tian, Z., and Zhang, B. (2017). An exact linearization method for OLTC of transformer in branch flow model. *IEEE Trans. Power Syst.* 32, 2475–2476. doi:10.1109/TPWRS.2016.2603438

Yang, P., and Nehorai, A. (2014). Joint optimization of hybrid energy storage and generation capacity with renewable energy. *IEEE Trans. Smart Grid.* 5 (4), 1566–1574. doi:10.1109/TSG.2014.2313724

Zeraati, M., Golshan, M. E. H., and Guerrero, J. M. (2019). Voltage quality improvement in low voltage distribution networks using reactive power capability of single-phase PV inverters. *IEEE Trans. Smart Grid.* 10, 5057–5065. doi:10.1109/TSG.2018.2874381

Zhang, J., Li, Y., Wu, Z., Rong, C., Wang, T., Zhang, Z., et al. (2021). Deep-reinforcement-learning-based two-timescale voltage control for distribution systems. *Energies* 14 (12), 3540. doi:10.3390/en14123540

Zhang, X., Zhang, X., and Liu, X. (2014). Partition operation on distribution network based on theory of generalized node. *Power Syst. Prot. Control* 42 (7), 122–127.

Zhang, Z., Dou, C., Yue, D., Zhang, B., and Zhang, H. (2020). Event-triggered voltage distributed cooperative control with communication delay. *Proc. CSEE* 40 (17), 5426–5435. doi:10.13334/j.0258-8013.pcsee.200456

Zhou, W., Zhang, N., Cao, Z., Chen, Y., Wang, M., and Liu, Y. (2021). "Voltage regulation based on deep reinforcement learning algorithm in distribution network with energy storage system," in Proceedings of the in 2021 4th International Conference on Energy Electrical and Power Engineering, Chongqing, China, April 2021, 892–896.

Zimmerman, R. D., Murillo-Sánchez, C. E., and Thomas, R. J. (2011). Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Trans. Power Syst.* 26, 12–19. doi:10.1109/TPWRS.2010.2051168

Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction.* MIT Press.Cambridge, Massachusetts.

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 30 January 2023
DOI 10.3389/fenrg.2023.1124700

OPEN ACCESS

EDITED BY
Zhi-Wei Liu,
Huazhong University of Science and Technology, China

REVIEWED BY
Yun Feng,
Hunan University, China
Feng Liu,
Stevens Institute of Technology, United States
Tao Huang,
James Cook University, Australia

*CORRESPONDENCE
Xiaobing Xiao,
253570868@qq.com

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 15 December 2022
ACCEPTED 09 January 2023
PUBLISHED 30 January 2023

CITATION
Xiao X, Li Y, He X, Cai Y, Xiao Y, Huang B and Jin X (2023), Optimal topology control of monitoring sensor network based on physical layer security for smart photovoltaic power system.
Front. Energy Res. 11:1124700.
doi: 10.3389/fenrg.2023.1124700

Optimal topology control of monitoring sensor network based on physical layer security for smart photovoltaic power system

Xiaobing Xiao[1]*, Yue Li[1], Xiaomeng He[2], Yongxiang Cai[1], Yong Xiao[3], Boyang Huang[3] and Xin Jin[3]

[1]Electric Power Research Institute of Guizhou Power Grid Co, Ltd, Guiyang, China, [2]Guiyang Power Supply Bureau of Guizhou Power Grid Co, Ltd, Guiyang, China, [3]Electric Power Research Institute of China Southern Power Grid, Guangzhou, China

In order to realize the intelligent management of photovoltaic power generation system, wireless sensor network (WSN) is considered as a promised solution. However, the energy of sensor nodes in WSN is limited and the security is not guaranteed. In order to prolong the service life of WSN and to improve the security of monitoring, this paper comprehensively considers the factors those affect the security and efficiency of WSN, such as transmission power, sensor association and the choice of Cluster Head (CH). Therefore, this paper proposes a joint sensor secure rate and energy efficiency optimization algorithm to obtain the optimal topology and transmission power design. Then, an effective iterative algorithm based on Block Coordinate Descent (BCD) technology is proposed to obtain the optimal solution to minimize the energy consumption of WSN and maximize the secure rate of sensor networks, thus ensuring the security and reliability of monitoring. Sensor association is modeled as a 0–1 multi-knapsack optimization problem. Different complexity methods are implemented to solve this NP-hard problem, and their performance are compared with simulations.

KEYWORDS

photovoltaic, wireless sensor network, cluster head, sensor association, transmission power

1 Introduction

Under the global energy crisis and environmental problems, clean energy has attracted more attention and has become a research hot point in recent years. In this context, the use of light to generate electricity has become more popular. Photovoltaic power generation is a technology that converts light energy into electric energy through photovoltaic effect. Taking photovoltaic power generation as an example, the significance of photovoltaic energy storage for joint participation in electricity market was studied in (Xing et al., 2020), which shows that photovoltaic power generation has a good prospect in the power market.

After 2020, the grid-connected solar photovoltaic system has increased its global cumulative capacity by almost 250 times, exceeding 300 GW by the end of 2016, which is double in 2020 (Jäger-Waldau et al., 2017). In the European Union, renewable energy is also widely used at present, such as photovoltaic panels. The influence of series and parallel resistors on photovoltaic panels is described (Pavlík et al., 2022). Photovoltaic system is one of the fastest growing systems in the world. By smoothly switching the maximum power point tracking mode and Constant Power Generation (CPG) mode, high performance and reliable operation are ensured. The proposed control can be directly applied to the existing photovoltaic system

without additional hardware equipment, such as irradiance sensor, which gives up the design of the system by combining sensors in consideration of the cost in the application (Helmers and Bett, 2016; Suryawanshi and Patil, 2019). However, in practical application, the situation that the photovoltaic system does not use sensors at all is relatively limited.

In recent years, more and more sensors are applied to photovoltaic systems in the application of photovoltaic equipment. There are more and more attempts on the application of current sensor in photovoltaic system. For example, in practical application, when more distributed photovoltaic power generation systems are connected to the distribution network, the corresponding current sensor should be selected to detect the current Direct Current (DC) component. Transmitting optical power is a better scheme to supply power for wireless sensor nodes. At the sensor, the light power in the form of monochromatic laser is converted from photovoltaic laser power to electric energy by photovoltaic laser power converter (Zhu et al., 2021).

With the continuous opening of wireless network, the number of wireless terminals is also increasing. How to ensure the security of wireless communication is a problem worth thinking about. Compared with traditional encryption methods, the physical layer of wireless communication generally guarantees security from the perspective of information theory (Hu and Li, 2014; Kong et al., 2019). This paper introduces the security objectives and implementation constraints of wireless sensor networks, summarizes the possible attacks and main defense means of wireless sensor networks, and analyzes various security technologies that are currently hot research topics (Zhu et al., 2008; Kong et al., 2016; Kong et al., 2021). A centralized, secure and low-energy adaptive clustering system is proposed, which uses gateway resources to extend the network life and provide trust management. It provides security, energy efficiency and memory efficiency, and can prevent all kinds of attacks, from spoofed messages to damaged node attacks, and it provides confidentiality (Yang, 2011; Zhong et al., 2019; Zhong et al., 2022).

Based on the above description, a secure communication system for photovoltaic power generation is designed, which collects the voltage and current information of the system through sensors. The Frequency Division Multiple Access (FDMA) method is used to connect the sensor with the base station. The nodes of the model include acquisition node, eavesdropping node and base station. Considering that such a system with more nodes will consume more energy, all the acquisition nodes are clustered. The work of this paper is as follows.

1) This paper analyse the factors that affect the security and efficiency of Wireless sensor network (WSN), such as transmission power, sensor association and the choice of Cluster Heads (CHs). In order to prolong the service life of WSN and to improve the security of monitoring, a joint sensor secure rate and energy efficiency optimization algorithm (TPTCO) is proposed. The optimal topology and transmission power design are obtained.
2) A joint optimization problem of multi-constrained mixed integer programming is proposed, and an effective iterative algorithm based on Block Coodinate Descent (BCD) technology is proposed to obtain the second-best solution to minimize the power consumption of WSN and maximize the secure rate of sensor networks, thus ensuring the security and reliability of monitoring.
3) Topology control is build as a 0–1 multi-knapsack optimization problem. Different complexity methods are put forward to solve this problem, and their performance differences are compared by simulation, so that they can be selected according to the actual needs.

The rest of this paper is organized as follows: **Section 2** summarizes the literature review of related work. In **Section 3**, we introduce the system model of communication model. In **Section 4**, the formulation and optimization of related problems are put forward. In **Section 5**, the simulation analysis of the whole system is given.

2 Literature review

Based on the digital signal processing (DSP) chip of the controller, this paper presents the hardware circuit of the control system. Moreover, the circuit can be used normally in areas with strong light. The photovoltaic power generation system includes a driving circuit, a current and voltage sampling circuit, a main circuit DC-DC converter and a protection circuit, and has a micro-pulse transscleral ciliary body photocoagulation function (Kuang et al., 2020). When the photovoltaic power generation device works outdoors for a long time, it is easily disturbed by the external environment, which affects the absorption of solar energy, greatly reduces the power generation efficiency of the device, and also leads to further damage of the device. In order to prevent this problem, a photovoltaic monitoring system is proposed in this paper (Chen et al., 2021). This paper presents a coupling modeling method, which compares the energy consumption of three different vehicles in different places in the integrated photovoltaic system. The method provides the required energy through the battery. A transparent 3D printing package is proposed to prevent the battery from being impacted by the outside world and reduce the battery loss (Ntagios et al., 2020). Complementary metal oxide semiconductor (CMOS) imaging and wearable devices can also be applied in battery biomedicine and wearable devices (Gu et al., 2021). The above results show that photovoltaic systems are widely used in different fields, and sensors are gradually applied in photovoltaic systems (Gu et al., 2021). An optimization model of distributed photovoltaic power generation is proposed. The goal of this model is to punish the cost of distributed photovoltaic power generation (Saffih, 2017).

A photovoltaic power supply model is proposed, which regards the energy consumption of WSN as the energy consumption of photovoltaic energy collector. The low power consumption $C\,O_2$ gas sensor is used for simulation analysis (Yue et al., 2017). Generally, the photovoltaic power generation management system based on the Internet of Things will use a third-party platform to manage IoT devices, but the security of the system is poor. In order to improve the flexibility and security of the system, a photovoltaic power generation management system based on narrow-band Internet of Things and a new photovoltaic equipment access protocol are proposed, which uses multi-sensor reporting data analysis rules (Yang and Zhou, 2021). A PV module is proposed, which supplies power to the implanted temperature sensor and is suitable for remote self-powered systems in various applications (Hung et al., 2018).

In the complex power field, intelligent sensors are often used to monitor the status of power equipment, as the core of the monitoring system, but how to supply energy to the sensors is

an important problem, which attracts much attention. Solar energy can be easily obtained from the outside world. A photovoltaic cell model is proposed to ensure a stable energy supply for monitoring equipment. In the complex power field, intelligent sensors are often used to monitor the status of power equipment, as the core of the monitoring system, but how to supply energy to the sensors is an important problem, which attracts much attention. Solar energy can be easily obtained from the outside world. A photovoltaic cell model is proposed to ensure a stable energy supply for monitoring equipment (Xiang et al., 2021).

A photovoltaic optimizer based on single inductor and single sensor is proposed. The optimizer includes inductor, sensor and MPPT unit, which improves the maximum output power of submodules, reduces the number of components, realizes the best current point tracking strategy and reduces the cost of photovoltaic optimizer (Pichel et al., 2018). A photovoltaic charging scheme based on MPPT strategy with single sensor is proposed, and a new optimization algorithm based on single current sensor is proposed to track MPPT (Kumar et al., 2017).

3 System model

As shown in **Figure 1**, we propose a photovoltaic power generation system that includes power generation equipment, energy storage equipment, and energy consumption equipment. In order to make the system stable, the paper uses a wireless sensor network to monitor the operation of the system. The energy consumption devices in this system include not only the energy consumed by the user but also the energy consumed by the wireless sensor network. Therefore, we consider a secure communication system for photovoltaic power generation systems, which collects information about the voltage and current of the system through sensors. In order to realize the intelligent management of the whole platform and ensure the safe operation of the whole system, the model consists of acquisition nodes, eavesdropping nodes and base stations. The acquisition node is responsible for transmitting the collected confidential data to the base station, while the eavesdropping node attempts to eavesdrop on the confidential data. It is assumed that the locations of all nodes are fixed, and all acquisition nodes will be eavesdropped when transmitting data. Considering that all nodes directly transmit data to the base station will consume a lot of energy, in order to reduce the energy consumption of the whole sensor network, all the acquisition nodes are clustered in this paper. After the sensor nodes are divided into multiple clusters, a central node is selected from each cluster as CHs, and CHs is used as a collection node, and the data of surrounding nodes are collected and transmitted to the base station. The sensor set is $\mathcal{N}$, and the total number of nodes is N, which can be expressed as $|\mathcal{N}| = N$. Because the location of the equipment in the photovoltaic system is not fixed, the sensor nodes are randomly deployed in a certain range. In this system, all sensor nodes have the same initial value and the energy consumed when collecting data is the same. In the whole system, all CHs are defined as the set $\mathcal{K}$, so the number of CHs included in WSN is K.

In this paper, FDMA is used as the communication connection between the sensor and the base station. FDMA is usually used for uplink transmission between the device and the server. It orthogonalizes the transmission of different users and completely reduces the intra-cell interference (Wang et al., 2020).

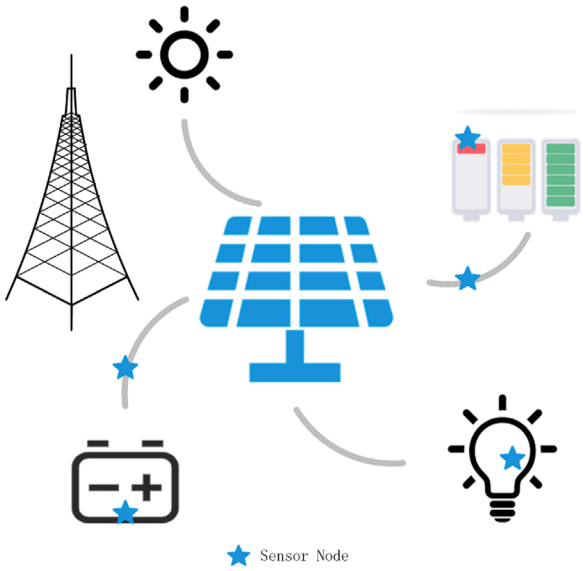


FIGURE 1
A photovoltaic power generation system based on WSN.

According to Shannon formula, the achievable transmission rate from node i to node j is:

$$r_{i,j} = B\log_2\left(1 + \frac{p_i h_{i,j}}{\sigma}\right), \tag{1}$$

Similarly, the reachable rate from eavesdropping node to node j can be expressed as:

$$r_{e,j} = B\log_2\left(1 + \frac{p_e h_{e,j}}{\sigma}\right), \tag{2}$$

p_i is the transmission power of the acquisition node. p_e is the transmitting power of the eavesdropping node. σ is Gaussian white noise. $h_{i,j} = \frac{|g_0|^2}{(d_{i,j})^2}$, $h_{e,j}$ is the channel power gain between the eavesdropping node and the node j. $h_{i,j}$ is the channel power gain between node i and node j. $d_{i,j}$ is the distance between node i and node j, g_0 is the Rayleigh fading channel coefficient, which accords with the complex Gaussian distribution $\zeta(0,1)$.

3.1 Communication model

Due to the position of each sensor node is fixed, the energy consumed by the sensor node when transmitting data to the base station is mainly determined by the power of the sensor. Therefore, this paper defines a cost function related to the power of sensor nodes, which increases in proportion to the power. The greater the power of sensor nodes, the greater the cost function and the more energy consumed. The energy consumption of sensor transmitting data once is shown as follows:

$$Cost(p), \tag{3}$$

$Cost(\cdot)$ is the monotonic increasing function of transmission power, which comprehensively reflects the influence of power, transmission

TABLE 1 Symbols parameters.

Variable	Parameter
$\beta_{i,j}$	Association decision
p_i	Transmission power of sensor node i
$N_j^{\max}$	Maximum access number of CH
$Cost(\cdot)$	Monotone increasing function of transmission power
$r_{i,j}$	Privacy rate from CHs to base station
$r_{e,j}$	Eavesdropping the achievable rate from node eavesdropping to node j
h_{ij}	Channel gain from node i to node j

distance and channel condition on energy consumption. The energy consumption of CMs is expressed as $Cost(p_i)$, and the energy consumption of CHs is expressed as $Cost(p_j)$.

3.2 Secure rate

In the system model proposed in this paper, a safe rate is proposed to represent the rate at which data can arrive safely when data is transmitted between different sensor nodes. Security rate is defined as the transmission rate between node i and node j minus the transmission rate between eavesdropping node and node j. Only when the transmission rate between node j and node j is greater than that between eavesdropping node and node j, data can be safely transmitted from node j to node j. The security rate of data transmission from node j to node j can be expressed as follows:

$$r_{i,j} = \begin{cases} r_{i,j} - r_{e,j}, & p_i h_{i,j} > p_e h_{e,j} \\ 0, & p_i h_{i,j} < p_e h_{e,j} \end{cases} \tag{4}$$

The transmission rate from node i to node j is $r_{i,j}$ on the left side of the formula, and the transmission rate from eavesdropping node to node j is $r_{e,j}$. The final $r_{i,j}$ is the security rate between node i and node j. CHs the secure rate to the base station is expressed as $r_{j,c}$. All sensor nodes are divided into multiple clusters, and each cluster has a CH to collect all CMs data and transmit it to the base station. Other sensor nodes in the cluster automatically join the CMs. Every CM can select one CH for association, Assume that is the association silver between CMs and CHs, and represents the CMs association CHj, otherwise. In addition, is expressed as the maximum associated number of CHs. Then the secure rate that all CMs can achieve is:

$$\sum_i \beta_{i,j} r_{i,j}, \tag{5}$$

The above symbols are summarized in **Table 1**.

4 Problem formulation and problem solution

The core of this paper is to reduce the power loss of the sensor and improve the use efficiency. Therefore, considering the topology of the sensor and the transmission power when transmitting data, the topology of the sensor can be optimized by CHs selection and sensor correlation. Therefore, this paper considers three optimization problems to reduce the energy consumption of sensor nodes, including transmission power design, CHs selection and sensor association. This paper selects CHs according to the designed CHs selection algorithm. Then, a multi-constrained mixed integer optimization problem based on joint transmission power and topology control is proposed. The algorithm proposed in this paper is used to solve the optimization problem, and at the same time, the correlation between the solved transmission power and the sensor under the current CHs is obtained. However, it is uncertain that the transmission power and sensor correlation obtained at this time are the optimal solution of this optimization problem, so it is necessary to iterate continuously until the CHs, transmission power and sensor correlation obtained is fixed. Through the above scheme, it can be considered that a suboptimal solution is obtained. Our proposed solution almost traverses all sensor nodes, so K CHs are randomly selected at the beginning of the simulation experiment.

4.1 Problem formulate

In this part, we formulate a joint optimization problem based on joint transmission power design and sensor association, aiming at maximizing the security of sensor transmission and reducing the energy consumption of sensor. Note that the transmission power of all sensors is $\boldsymbol{p}$. The relationship between CMs and CHs is $\boldsymbol{\beta} = \{\beta_{i,1},\ldots,\beta_{i,j},\ldots,\beta_{i,K}, \forall i \in \mathbb{N}, \forall j \in \mathcal{K}\}$. In order to ensure the security of data uploaded by sensors and reduce the energy consumption in the transmission process, we formulate a multi-constrained mixed integer optimization problem about transmission power $\boldsymbol{p}$ and node association $\boldsymbol{\beta}$. The optimization problem is expressed as follows:

$$\mathbf{P}: \max_{\boldsymbol{p},\boldsymbol{\beta}} \ \omega_1 \left(\sum_i \sum_j \beta_{i,j} r_{i,j} + \sum_j r_{j,c} \right) - \omega_2 \left(\sum_i Cost(p_i) + \sum_j Cost(p_j) \right) \tag{6}$$

$$s.t. p_{\min} < p_i < p_{\max}, \tag{6a}$$

$$\beta_{i,j} \in \{0,1\}, \tag{6b}$$

$$\sum_i \beta_{i,j} = 1, \tag{6c}$$

$$\sum_j \beta_{i,j} = N_j^{\max}, \tag{6d}$$

The weight factors $\omega_1 > 0, \omega_2 > 0$ are the tradeoff between secure rate and energy consumption. Constraints 6a indicates the limit of transmission power of sensor nodes of CMs; Constraints 6b and 6c mean that each CM can only access one CH; A constraint of 6 d means that the maximum number of CMs accessed by each CH is $N_j^{\max}$.

Because the correlation factor is a binary variable, problem $\boldsymbol{P}$ is a non-convex mixed integer optimization problem. It is difficult to obtain the optimal solution of this problem by standard convex optimization techniques. Therefore, this paper adopts the block coordinate descent algorithm to decompose the optimization problem into two sub-problems, and solves problem $\boldsymbol{P}$ by iteratively solving the optimal solution of the sub-problems.

4.2 Problem solving

In this part, we use the BCD method of convex optimization to decompose problem $\boldsymbol{P}$ into two optimization problems: transmission power $\boldsymbol{p}$ and sensor node association $\boldsymbol{\beta}$, and get a suboptimal solution through continuous iterative optimization. In each iteration, the transmission power design problem in the fixed sensor correlation result $\boldsymbol{\beta}$ is first solved. Then, we optimize the sensor association under the given power control design $\boldsymbol{p}$. When the proposed algorithm converges to the predetermined accuracy, the suboptimal solution can be obtained.

4.2.1 Power control optimization

When the node association $\boldsymbol{\beta}$ is given, the power control optimization problem can be simplified as follows:

$$\begin{aligned}\mathbf{P1}:\ \max_{\boldsymbol{p}}\quad & \omega_1\left(\sum_i\sum_j\beta_{i,j}r_{i,j}+\sum_j r_{j,c}\right)-\omega_2\left(\sum_i Cost(p_i)+\sum_j Cost(p_j)\right)\\ &=\omega_1\left(\sum_i\sum_j\beta_{i,j}B\left(\log_2\left(1+\frac{p_i h_{i,j}}{\sigma}\right)-\log_2\left(1+\frac{p_e h_{e,j}}{\sigma}\right)\right)+\sum_j r_{j,c}\right)\\ &-\omega_2\left(\sum_i Cost(p_i)+\sum_j Cost(p_j)\right)\\ s.t.\quad & p_{\min}<p_i<p_{\max},\end{aligned}\tag{7}$$

When $Cost(\cdot)$ is a convex function, the problem is a standard convex optimization problem, so the strong dual condition holds. In the next part, we use Lagrange duality technique to formulate the optimal solution p^* of this optimization problem. Then, the Lagrange function of the dual problem can be formulated as:

$$\begin{aligned}&\mathcal{L}(\boldsymbol{p},\lambda,\nu)\\ &=\omega_1\left(\sum_i\sum_j\beta_{i,j}r_{i,j}+\sum_j r_{j,c}\right)-\omega_2\left(\sum_i Cost(p_i)+\sum_j Cost(p_j)\right)\\ &+\sum_i\lambda_i(p_i-p_{\min})+\sum_i\nu_i(p_{\max}-p_i)\end{aligned}\tag{8}$$

Among $\nu=\{\nu_i,\forall i\}$, $\nu=\{\nu_i,\forall i\}$ is a non-negative Lagrange multiplier related to their respective constraints. The formula of dual function is as follows:

$$D(\lambda,\nu)=\max_p \mathbb{L}(\boldsymbol{p},\lambda,\nu)\tag{9}$$

Then, the dual problem is given through.

$$\begin{aligned}&\min_{\lambda,\nu}\ D(\lambda,\nu)\\ &s.t.\quad \lambda\geq 0,\\ &\qquad \nu\geq 0,\end{aligned}\tag{10}$$

The original problem is convex about p_i, so the solution satisfying KKT condition is also the optimal solution of the original problem. By finding the first derivative of the objective function to p_i, the optimal transmission power allocation, denoted as p_i^*, can be obtained by:

$$p_i^*=\frac{\omega B\sum_j\beta_{i,j}}{\ln 2\left(\omega_2 a+\lambda_i-\nu_i\right)}-\frac{\sigma}{\sum_j h_{i,j}}\tag{11}$$

After we find p_i^*, we get the dual function $D(\lambda,\nu)$ by the above formula. The optimal dual variable cannot be obtained by one-time solution, but it needs to be obtained by projected sub-gradient. The formula of projected sub-gradient method is as follows:

$$\lambda_i^{(k+1)}=\left[\lambda_i^{(k)}+s_k\left(p_i-p_{\min}\right)\right]^+\tag{12}$$

$$\nu_i^{(k+1)}=\left[\nu_i^{(k)}+s_k\left(p_{\max}-p_i\right)\right]^+\tag{13}$$

where $s_k = dfrac\ gamma\|gk\|$ is defined as the step size of the kth cycle, and $s_k>0$. g^k is defined as the gradient corresponding to the kth cycle. On the basis of Lagrange duality and projection sub-gradient method, we propose a cyclic optimization algorithm, as shown in **Algorithm 2**.

4.2.2 Node association optimization

Given the transmission power $\boldsymbol{p}$, the optimization problem of node association can be simplified as follows:

$$\begin{aligned}\mathbf{P2}:\ \max_{\beta_{i,j}}\quad & \omega_1\sum_i\sum_j\beta_{i,j}r_{i,j}\\ &=\omega_1\sum_i\sum_j\beta_{i,j}B\left(\log_2\left(1+\frac{p_i h_{i,j}}{\sigma}\right)-\log_2\left(1+\frac{p_e h_{e,j}}{\sigma}\right)\right)\\ s.t.\quad & \beta_{i,j}\in\{0,1\},\\ & \sum_i\beta_{i,j}=1,\\ & \sum_j\beta_{i,j}=N_{\max}\end{aligned}\tag{14}$$

The problem is a 0–1 multi-knapsack problem. With the increase of the dimension of the problem, the time complexity of the algorithm for solving the problem will also increase dramatically. How to solve the integer optimization problem is a problem worth thinking about. With the increase of the dimension of the problem, the time complexity of the algorithm for solving the problem will also increase dramatically. How to solve the integer optimization problem is a problem worth thinking about. In this paper, an algorithm is proposed to get the sub-optimal solution of the node association optimization problem. f_{ij} will be defined as the income of sensor node i accessing CHj, The expression f_{ij} is:

$$f_{ij}=B\left(\log_2\left(1+\frac{p_i h_{i,j}}{\sigma}\right)-\log_2\left(1+\frac{p_e h_{e,j}}{\sigma}\right)\right)\tag{15}$$

In the iteration of the algorithm, all unallocated sensor nodes are considered first. If the maximum distance between the current node and the second node is the largest, the sensor i is connected to the CHj until all sensor nodes are connected to the CH, and the iteration of the algorithm is stopped. The proposed algorithm of sensor association scheme based on MTHG is shown in **Algorithm 1**.

4.3 CH selection

In order to improve the security of the whole photovoltaic system and reduce the energy consumption of the system, the CHs selection algorithm is proposed in this paper. CHs selection algorithm divides sensor nodes into multiple clusters, and selects CHs according to the minimum transmission distance between CHs and CMs, thus making the transmission energy consumption between each CM smaller. At the same time, the transmission distance between CH and the base station should be as small as possible to reduce the transmission energy consumption of CHs to the base station. We use the sum of the minimum Euclidean distances between CHs and CMs to select CHs. CHs selection formula is as follows:

$$CH_j = arg\min_m \sum\{dist_m,\forall m\in\mathfrak{O}\},\forall j\in\mathcal{K}\tag{16}$$

Among, $\mathfrak{O}$ represents the deletion of the set of nodes m. $\sum dist_m,\forall m\in\mathfrak{O}$ is the sum of the distances from sensor m to other sensor in the cluster.

Evaluation random CHs and topology control β^0.
Evaluation $(\lambda,v)^1$ and $r_1,r_2,r_3=0$.
repeat
 repeat
  repeat
   Calculate p^{r_1} by Eq. **11**.
   Calculate $(\lambda,v)^{r_1+1}$ by Eqs **12**, **13**.
   Calculate $r_1=r_1+1$.
  until Sub-gradient algorithm reaches a certain value.
  Update f_{ij} by Eq. **15**.
  while $\mathcal{G}\neq\varnothing$ **do**
   $d^*=-\infty$.
   for all $i\in\mathcal{G}$ **do**
    $j'=\underset{j}{argmax}\{f_{ij},\forall i\in\mathcal{G}\}$.
    Update $d=f_{ij'}-max\{f_{ij},\forall j\in\mathcal{K}\}$.
    if $d>d^*$ **then**
     $d^*=d$.
     $\beta_{ij'}=1$.
    end if
   end for
   delete sensor i in $\mathcal{G}$.
  end while
  Calculate $r_2=r_2+1$.
 until The target reaches a certain value.
 Acquire the CHs by CHs selection algorithm.
 Update $r_3=r_3+1$.
until CHs and topology control $\boldsymbol{\beta}$ is fixed.

Algorithm 1. MTHG-Based Topology Control (MTHG)

4.4 Overall algorithm design

We propose a whole iterative loop algorithm, as shown in **Algorithm 2**, aiming at effectively solving the optimization problem with reduced algorithm complexity. Firstly, the topological association of sensor networks is randomly initialized, and the transmission power is solved by sub-gradient algorithm. Then, the price of knapsack problem is calculated by the obtained transmission power, and the corresponding sensor association result is obtained by traversing all CMs. Iterate for many times, reach the accuracy given by BCD algorithm, and get a solution. The solution obtained at this time is not suboptimal, because CHs is random. Finally, the CHs is selected according to the given CHs selection algorithm, and the transmission power and sensor correlation results are recalculated. After repeated iterations, until the CH sum does not change any more and the target value converges to the predetermined accuracy, we think that a suboptimal solution has been found for this problem. Repeated iterations are because CHs is randomly initialized in the initialization process, and after repeated iterations, high-quality topology control strategies can be obtained. The details of this algorithm are summarized as algorithm TPTCO. The flowchart of the TPTCO algorithm is shown in **Figure 2**.

Evaluation random CHs and topology control β^0.
Evaluation $(\lambda,v)^1$ and $r_1,r_2,r_3=0$.
repeat
 repeat
  repeat
   Calculate p^{r_1} by Eq. **11**.
   Calculate $(\lambda,v)^{r_1+1}$ by Eq.**12** and **13**.
   Update $r_1=r_1+1$.
  until Sub-gradient algorithm reach to a fixed value.
  Acquire topology control through **Algorithm 1** (MTHG).
 until BCD method reduce to a certain value.
 Calculates the sum of Euclidean distances between each node and other nodes.
 CH is the node of the sum of the minimum Euclidean distances.
until Topology control and CHs is fixed.

Algorithm 2. Transmission Power and Topology Control Optimization (TPTCO)

5 Simulation analysis

We prove the effectiveness of the algorithm proposed in this paper by simulation results. In the simulation, it is assumed that a WSN topology is randomly deployed, and the sensor nodes are randomly deployed in a certain range. Where d > 0 represents the Euclidean distance from CH to BS and from CMs to CH (Heinzelman et al., 2000). The MTHG algorithm proposed in this paper is compared with GREEDY, Relax-CVX and other algorithms by simulation.

5.1 Simulation parameters

This paper evaluates the performance of the proposed algorithm by analyzing the clustering results, the energy consumption of the algorithm, the convergence of the algorithm, the transmission security and the load of the cluster head. The simulation results are obtained by running Matlab 2016b simulation software on a computer equipped with Intel Core i7-7700,296 3.6 GHz CPU and 16 GB memory. The parameters used in the simulation experiment are shown in **Table 2**.

5.2 Analysis of simulation results

Figure 3 shows the topology of WSN in the first round of operation, and the results of sensor association after running MTHG algorithm in this topology. In the simulation process, all sensor nodes are randomly deployed according to PPP density, and MTHG algorithm divides sensor nodes into CHs and CMs. CM is associated with CH for data transmission. Using different sensor node deployment algorithms will lead to uneven density of sensor nodes in different areas, and higher CH load in denser areas, which will lead to increased energy consumption of CH and thus the energy consumption of the whole system. Therefore, it is necessary to limit the

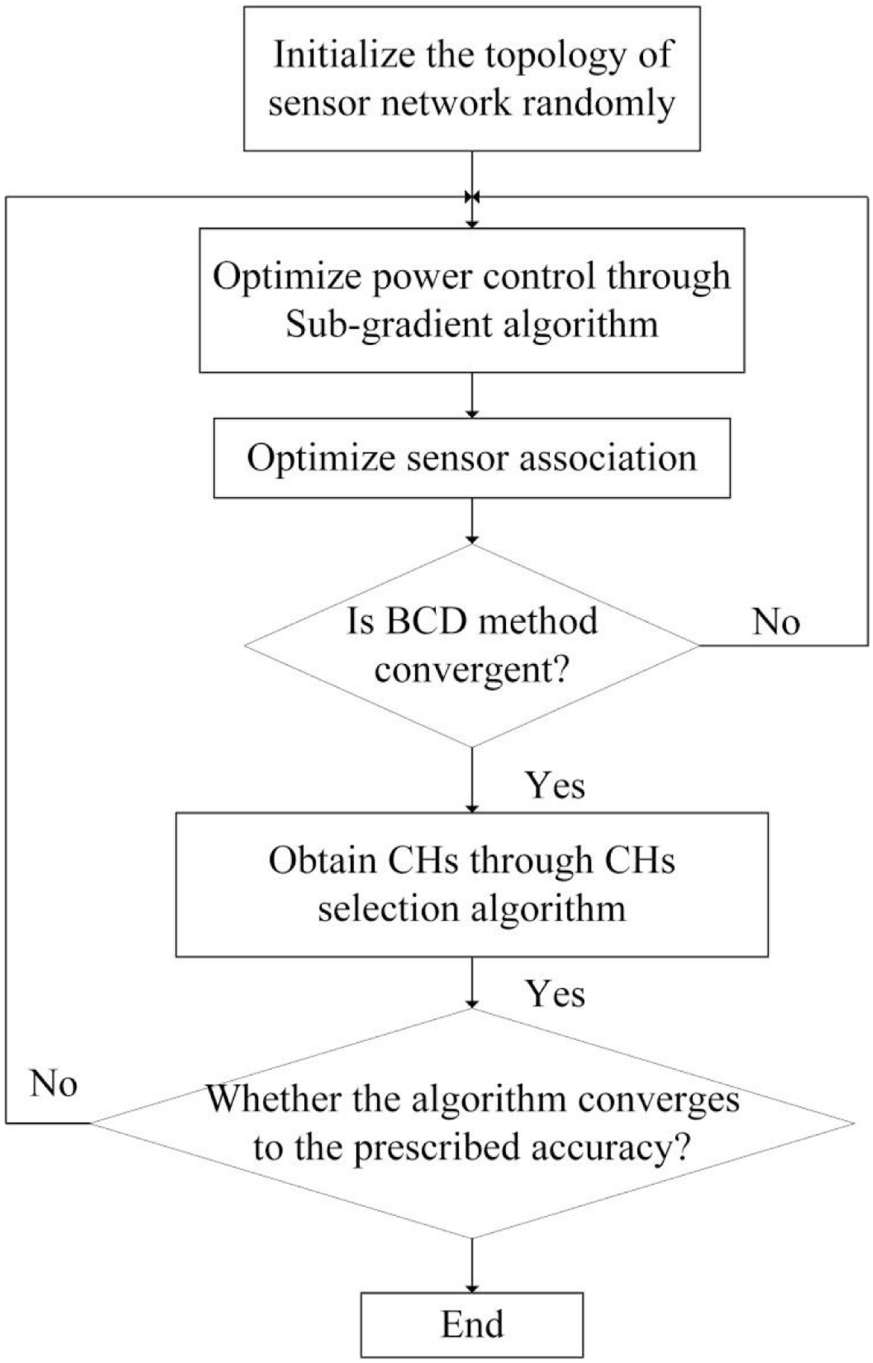


FIGURE 2
The flowchart of the TPTCO algorithm.

maximum number of CH associations. In this paper, sensor nodes are divided into five clusters. It can be seen from the figure that according to the CHs selection algorithm, each CH is distributed in the center of each cluster as much as possible, which will shorten the Euclidean distance between CH and CM, thus reducing the energy consumption during data transmission.

Figure 4 shows the convergence of the algorithm proposed in this paper compared with other algorithms. It can be seen from the figure that, at the beginning, the convergence curves of all algorithms keep rising with the number of iterations. After iteration for a period of time, the correlation between CHs and sensors will not change, and then all algorithms will gradually converge. It can be seen from the convergence curves of different algorithms that CVX has the best convergence performance. The convergence curve gap between MTHG algorithm and Relax algorithm proposed in this paper is very small, while the performance of GREEDY algorithm is worse. The reason why there is little difference between MTHG algorithm and GREEDY algorithm may be that both of

TABLE 2 Simulation parameters.

Variable	Parameter	Value
S	Distribution area	500×500
ρ	Deployment density of WSN nodes	250
$N_j^{\max}$	Maximum access number of CH	30
w_1	Weight of security	100*s*
w_2	Weight of energy consumption	100*s*
$P_{\min}$	Minimum transmission power	100*s*
$P_{\max}$	Maximum transmission power	500*s*

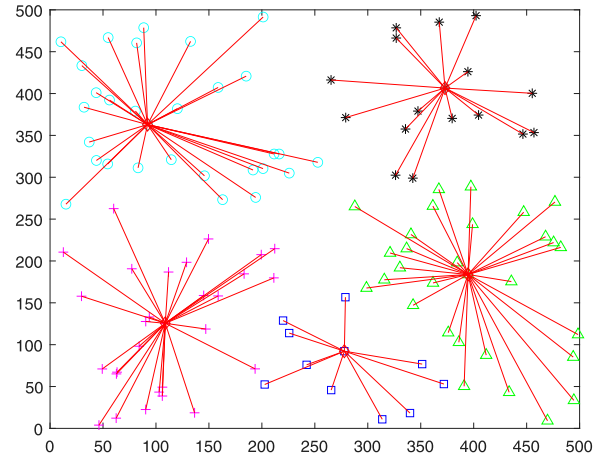


FIGURE 3
Topology of WSN and result of sensor association by MTHG.

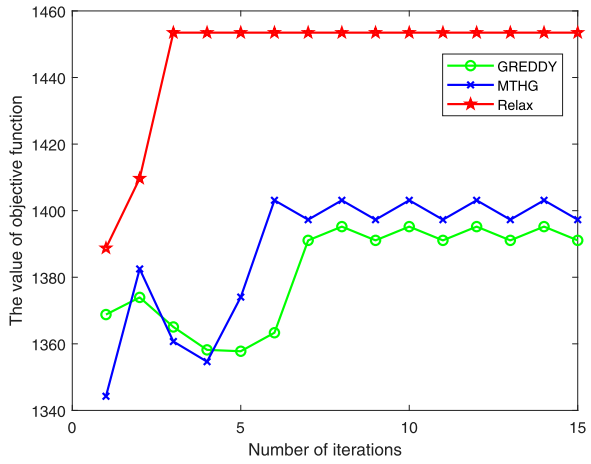


FIGURE 4
Convergence behaviour of the proposed algorithm.

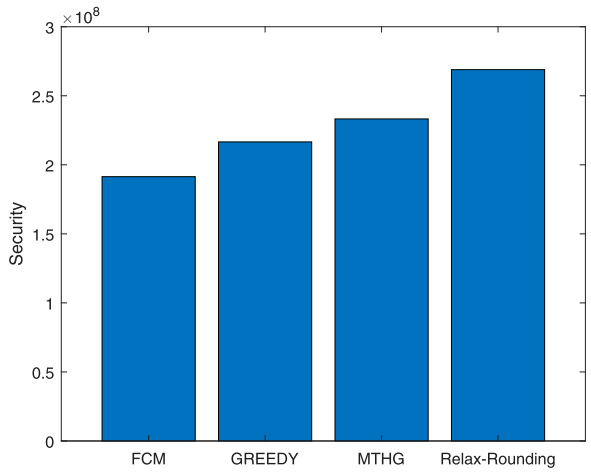


FIGURE 5
Security of the proposed algorithm.

them use the minimum Euclidean distance for clustering. Therefore, we can know that the MTHG algorithm proposed in this paper can also get high-quality approximate suboptimal solutions. Even if there is a certain gap between the solution and the optimal solution. GREEDY got a feasible solution with the lowest algorithm complexity. Compared with GREEDY algorithm, MTHG algorithm has more obvious improvement when the complexity increases little.

Figure 5 shows the security of different algorithms for unit energy consumption, with Realx-Rounding being the best. MTHG proposed in this paper is the closest to Realx-Rounding, GREEDY algorithm is the third, and Fuzzy C-Means (FCM) algorithm is the worst. Realx-Rounding takes the most time. For FCM algorithm, security in the transmission process is not considered, so the security of FCM algorithm is poor. For GREEDY algorithm, in the process of optimizing sensor association, only the minimum energy consumption is considered, so CMs will associate with its nearest CHs as much as possible, without considering the security of data transmission, which will also lead to the reduction of security during transmission. However, the computational complexity of MTHG proposed in this paper is significantly reduced. It can be seen from the above that MTHG proposed in this paper improves the security requirements in the process of data transmission.

Figure 6 shows the CHs load under different algorithms. For CMs, they will always reduce the energy consumed in the transmission process as much as possible under the CHs constraint, so CMs will choose to associate with its nearest CHs as much as possible. It can be seen from the figure that under the same constraints, the load calculated by MTHG algorithm is the most balanced, followed by CVX, and the CHs load calculated by Greedy algorithm is the most unbalanced. For GREEDY algorithm, in the process of optimizing sensor association, only the minimum energy consumption is considered, so CMs will associate with its nearest CHs as much as possible, which will lead to the unbalanced load of CHs. The MTHG algorithm proposed in this paper iteratively solves the transmission power, sensor association and CHs selection for many times until the obtained results are unchanged. The CVX algorithm optimizes the problem as a whole, and does not consider the load balance too much, so the CHs load is not so balanced. Therefore, MTHG can balance the load of CHs as much as possible, which can greatly improve the service life of sensor nodes.

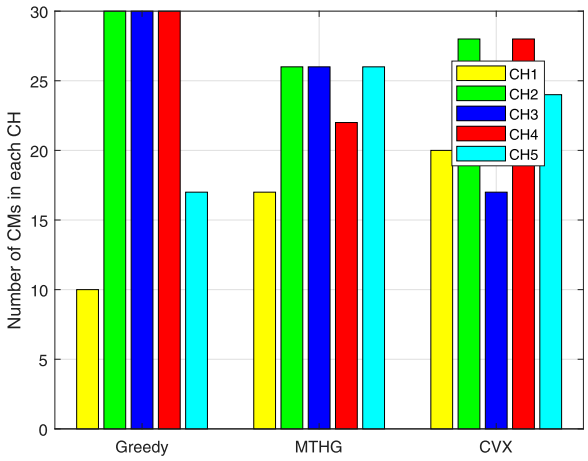


FIGURE 6
CHs load of the proposed algorithm.

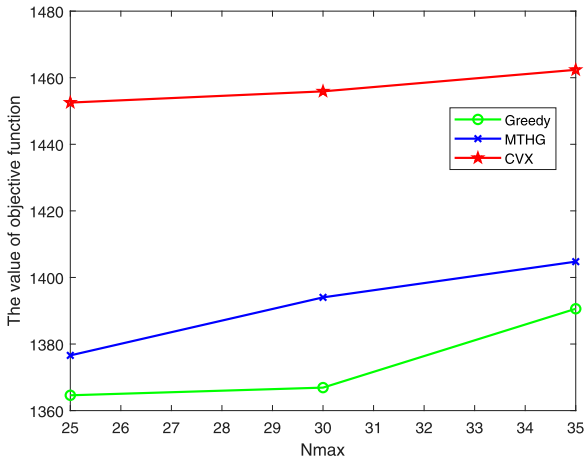


FIGURE 7
The relationship between n value and target value.

Figure 7 shows the convergence performance of the algorithm under the changed conditions. CVX has the best convergence, because CVX considers the optimization problem as a whole in the optimization process, and obtains the optimal correlation result. Therefore, when it changes, the target value calculated by CVX algorithm does not change much, but CVX algorithm consumes the most time. The MTHG algorithm proposed in this paper has better convergence than GREEDY algorithm. GREEDY algorithm only considers the minimum energy consumption in the process of optimizing sensor association. When it increases, CMs will be associated with its nearest CHs as much as possible, and the energy consumption will be correspondingly reduced. The objective function value of the optimization problem will be improved, but at the same time, the load imbalance of CHs will be more serious. The MTHG proposed in this paper reduces the complexity of CVX algorithm and is very important for the implementation of solutions.

6 Conclusion

The intelligent management of photovoltaic power generation system needs to collect information from all parts of the system through sensors. In order to improve the security of WSN and reduce the energy consumption of sensors, this paper uses BCD method to solve the joint optimization problem of transmission power and sensor association, and proposes MTHG algorithm to solve the problem of sensor association optimization, and uses sub-gradient method to solve the transmission power. Select CHs again according to the association of sensors. The simulation results show that the algorithm can effectively improve the security of the transmission process, reduce the energy consumption of the sensor, and thus prolong the service life of the wireless sensor network. In the future work, we will try to improve the algorithm by reducing the gap between the suboptimal solution and the optimal solution and reducing its complexity. In addition, a more realistic environment will be considered in the optimization problem.

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

Methodology, XX; software, YL; formal analysis, XX and YL; investigation, YC; resources, XX; data curation, YL and XH; writing—original draft preparation, YX; visualization, BH; supervision, YX; project administration, XJ; funding acquisition, BH All authors have read and agreed to the published version of the manuscript.

Funding

China Southern Power Grid Corporation key technology project: GZKJXM20220052.

Conflict of interest

XX, YL, and YC were employed by Electric Power Research Institute of Guizhou Power Grid Co, Ltd; XH was employed by Guiyang Power Supply Bureau of Guizhou Power Grid Co, Ltd.; YX, BH, and XJ are employed by Electric Power Research Institute of China Southern Power Grid.

The authors declare that this study received funding from China Southern Power Grid Corporation. The funder had the following involvement in the study: data collection and analysis, preparation of the manuscript.

Publisher's note

References

Chen, Y., Ma, J., Ma, Y., Zhang, Y., Zhou, X., and Zhou, R. (2021). "The design and implementation of dust monitoring system for photovoltaic power generation," in 2021 IEEE 23rd Int Conf on High Performance Computing and Communications; 7th Int Conf on Data Science and Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud and Big Data Systems and Application (HPCC/DSS/SmartCity/DependSys) (Haikou, Hainan, China: IEEE), 1458–1465. doi:10.1109/HPCC-DSS-SmartCity-DependSys53884.2021.00218

Gu, J., Zhao, L., and Zhu, C. (2021). "Optimal planning of distributed photovoltaic in integrated electricity and heat systems," in 2021 4th International Conference on Energy, Electrical and Power Engineering (CEEPE), 982–986. doi:10.1109/CEEPE51765.2021.9475729

Heinzelman, W., Chandrakasan, A., and Balakrishnan, H. (2000). "Energy-efficient communication protocol for wireless microsensor networks," in Proceedings of the 33rd Annual Hawaii International Conference on System Sciences (Maui, HI, USA: IEEE), 10. doi:10.1109/HICSS.2000.926982

Helmers, H., and Bett, A. W. (2016). "Photovoltaic laser power converters for wireless optical power supply of sensor systems," in 2016 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE) (Aachen, Germany: IEEE), 152–154. doi:10.1109/WiSEE.2016.7877321

Hu, A., and Li, G. (2014). Physical layer security in wireless communication: Survey. *J. Data Acquis. Process.* 5, 237. doi:10.23919/JCIN.2020.9200889

Hung, Y. J., Cai, M. S., Chen, J. F., Su, H. W., Jen, P. C., Chen, P., et al. (2018). High-voltage backside-illuminated cmos photovoltaic module for powering implantable temperature sensors. *IEEE J. Photovoltaics* 8, 342–347. doi:10.1109/JPHOTOV.2017.2775440

Jäger-Waldau, A., Huld, T., and Szabo, S. (2017). "Residential photovoltaic electricity generation in the European Union 2017-opportunities and challenges," in 2017 IEEE 44th Photovoltaic Specialist Conference (PVSC) (Washington, DC, USA: IEEE), 2167–2169. doi:10.1109/PVSC.2017.8366336

Kong, Z., Song, J., Wang, C., Chen, H., and Hanzo, L. (2021). Hybrid analog-digital precoder design for securing cognitive millimeter wave networks. *IEEE Trans. Inf. Forensics Secur.* 16, 4019–4034. doi:10.1109/TIFS.2020.3039697

Kong, Z., Yang, S., Wang, D., and Hanzo, L. (2019). Robust beamforming and jamming for enhancing the physical layer security of full duplex radios. *IEEE Trans. Inf. Forensics Secur.* 14, 3151–3159. doi:10.1109/TIFS.2019.2908481

Kong, Z., Yang, S., Wu, F., Peng, S., Zhong, L., and Hanzo, L. (2016). Iterative distributed minimum total mse approach for secure communications in mimo interference channels. *IEEE Trans. Inf. Forensics Secur.* 11, 594–608. doi:10.1109/TIFS.2015.2493888

Kuang, S., Chang, K. C., Chu, K. C., Lin, Y. C., Wang, H. C., Hsu, T. L., et al. (2020). "Study of dsp (tms320f2812) chip for advanced photovoltaic power generation maximum power point tracking control system," in 2020 15th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT) (Taipei, Taiwan: IEEE), 294–297. doi:10.1109/IMPACT50485.2020.9268608

Kumar, N., Hussain, I., Singh, B., and Panigrahi, B. K. (2017). Single sensor-based mppt of partially shaded pv system for battery charging by using cauchy and Gaussian sine cosine optimization. *IEEE Trans. Energy Convers.* 32, 983–992. doi:10.1109/TEC.2017.2669518

Ntagios, M., Escobedo, P., and Dahiya, R. (2020). "3d printed packaging of photovoltaic cells for energy autonomous embedded sensors," in 2020 IEEE SENSORS (Rotterdam, Netherlands: IEEE), 1–4. doi:10.1109/SENSORS47125.2020.9278635

Pavlík, M., Kolcun, M., Bucko, S., and Kohan, V. (2022). "Analysis and evaluation of photovoltaic cell defects and their impact on electricity generation," in 2022 22nd International Scientific Conference on Electric Power Engineering (EPE) (Kouty nad Desnou, Czech Republic: IEEE), 1–4. doi:10.1109/EPE54603.2022.9814105

Pichel, N., Vivar, M., and Fuentes, M. (2018). "Optimization study of a photovoltaic-photochemical hybrid system (solwat) for meeting the needs of electricity and clean water," in 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC and 34th EU PVSEC) (Waikoloa, HI, USA: IEEE), 1222–1224. doi:10.1109/PVSC.2018.8547320

Saffih, F. (2017). "Artificially-intelligent imaging (AI^2) sensors: How intelligent CMOS imaging devices can benefit photovoltaics?," in 2017 7th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO) (Sharjah, UAE: IEEE). doi:10.1109/ICMSAO.2017.7934899

Suryawanshi, S. Y., and Patil, M. D. (2019). "Implementation of a DC side sensor-less cascaded H-bridge multilevel converter for photovoltaic system," in 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (Tirunelveli, India: IEEE), 1104–1108. doi:10.1109/ICOEI.2019.8862535

Wang, Q., Tan, L. T., Hu, R. Q., and Qian, Y. (2020). Hierarchical energy-efficient mobile-edge computing in iot networks. *IEEE Internet Things J.* 7, 11626–11639. doi:10.1109/JIOT.2020.3000193

Xiang, S., Du, L., Li, C., Li, Y., Yu, H., and Huang, P. (2021). "An optimization method based on lm-ga for parameter identification of photovoltaic cell," in 2021 6th Asia Conference on Power and Electrical Engineering (ACPEE) (Chongqing, China: IEEE), 405–409. doi:10.1109/ACPEE51499.2021.9437110

Xing, C., Xi, X., He, X., and Liu, M. (2020). "Research on the mppt control simulation of wind and photovoltaic complementary power generation system," in 2020 IEEE Sustainable Power and Energy Conference (iSPEC) (Chengdu, China: IEEE), 1058–1063. doi:10.1109/iSPEC50848.2020.9350965

Yang, L. (2011). *Centralized security protocol for wireless sensor networks presented to.*

Yang, M., and Zhou, G. (2021). "Design and implementation of photovoltaic power generation management system based on NB-IoT," in 2021 IEEE International Conference on Information Communication and Software Engineering (ICICSE) (Chengdu, China: IEEE), 255–259. doi:10.1109/ICICSE52190.2021.9404117

Yue, X., Kauer, M., Bellanger, M., Beard, O., Brownlow, M., Gibson, D., et al. (2017). Development of an indoor photovoltaic energy harvesting module for autonomous sensors in building air quality applications. *IEEE Internet Things J.* 4, 2092–2103. doi:10.1109/JIOT.2017.2754981

Zhong, L., Ge, M.-F., Zhang, S., and Liu, Y. (2022). Rate-aware fuzzy clustering and stable sensor association for load balancing in wsns. *IEEE Internet Things J.* 9, 3559–3573. doi:10.1109/JIOT.2021.3098352

Zhong, L., Li, M., Cao, Y., and Jiang, T. (2019). Stable user association and resource allocation based on stackelberg game in backhaul-constrained hetnets. *IEEE Trans. Veh. Technol.* 68, 10239–10251. doi:10.1109/TVT.2019.2937941

Zhu, Y., Wen, H., Chu, G., Hu, Y., Li, X., and Ma, J. (2021). High-performance photovoltaic constant power generation control with rapid maximum power point estimation. *IEEE Trans. Industry Appl.* 57, 714–729. doi:10.1109/TIA.2020.3029128

Zhu, Z. J., Tan, Q. P., and Zhu, P. D. (2008). A survey of wsn security research. *Comput. Eng. Sci.*

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 14 February 2023
DOI 10.3389/fenrg.2023.1141954

OPEN ACCESS

EDITED BY
Chao Deng,
Nanjing University of Posts and Telecommunications, China

REVIEWED BY
Weihua Wu,
Shaanxi Normal University, China
Haitao Zhao,
National University of Defense Technology, China
Liu Yang,
Jiangnan University, China

*CORRESPONDENCE
Yiyang Ni,
niyy@jssnu.edu.cn

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 11 January 2023
ACCEPTED 30 January 2023
PUBLISHED 14 February 2023

CITATION
Zhao J, Ni Y and Zhu H (2023), Multi-armed bandit based device scheduling for crowdsensing in power grids.
Front. Energy Res. 11:1141954.
doi: 10.3389/fenrg.2023.1141954

Multi-armed bandit based device scheduling for crowdsensing in power grids

Jie Zhao[1,2], Yiyang Ni[1,2]* and Huisheng Zhu[1,2]

[1]College of Physics and Information Engineering, Jiangsu Second Normal University, Nanjing, China, [2]Jiangsu Province Engineering Research Center of Basic Education Big Data Application, Jiangsu Second Normal University, Nanjing, China

With the increase of devices in power grids, a critical challenge emerges on how to collect information from massive devices, as well as how to manage these devices. Mobile crowdsensing is a large-scale sensing paradigm empowered by ubiquitous devices and can achieve more comprehensive observation of the area of interest. However, collecting sensing data from massive devices is not easy due to the scarcity of wireless channel resources and a large amount of sensing data, as well as the different capabilities among devices. To address these challenges, device scheduling is introduced which chooses a part of mobile devices in each time slot, to collect more valuable sensing data. However, the lack of prior knowledge makes the device scheduling task hard, especially when the number of devices is huge. Thus the device scheduling problem is reformulated as a multi-armed bandit (MAB) program, one should guarantee the participation fairness of sensing devices with different coverage regions. To deal with the multi-armed bandit program, a device scheduling algorithm is proposed on the basis of the upper confidence bound policy as well as virtual queue theory. Besides, we conduct the regret analysis and prove the performance regret of the proposed algorithm with a sub-linear growth under certain conditions. Finally, simulation results verify the effectiveness of our proposed algorithm, in terms of performance regret and convergence rate.

KEYWORDS

crowdsensing, device scheduling, multi-armed bandit (MAB), edge intelligence, power grid

Introduction

Nowadays, the development of smart power grids brings much convenience to human life and production. Meanwhile, more and more devices, such as sensors and actuators, are deployed in power grids, e.g., substations, transformers, and generators. Consequently, A critical challenge arises on how to collect information from massive devices and how to manage these devices. Mobile crowdsensing is a large-scale sensing paradigm empowered by ubiquitous devices. These devices interact with each other by sharing local knowledge according to the data they have perceived, and then the information can be further aggregated and fused in a central node for crowd intelligence extraction, decision-making, and service delivery (Guo et al., 2014).

However, collecting sensing data from massive devices is not easy due to the following reasons. Firstly, the scarce channel resource limits the number of devices that simultaneously access to an edge server. That is to say, the available wireless channels are fewer than the

sensing devices. Secondly, the overlap of perception areas of different devices introduces sensing data redundancy. Besides, the system heterogeneity of sensing devices, such as processing capability, network connectivity, and battery capacity, leads to different processing capabilities (Xia et al., 2021). The system heterogeneity causes a drift of global statistical characteristics since the fast devices can collect more data according to their local observations. To achieve a more comprehensive observation of the area of interest, one should guarantee the participation fairness of sensing devices with different coverage regions. Therefore, the edge server has to perform device scheduling, i.e., choosing a part of sensing devices in each time slot, to collect more valuable sensing data. However, the lack of prior knowledge makes the device scheduling task hard, especially when the number of sensing devices is huge.

Actually, there have been some works on device scheduling in crowdsensing tasks. For example, The authors in (Chu et al., 2013) proposed a selection scheme of individual sensors to collect data in different regions in order to optimize some specified objective while satisfying constraints in the number and costs of sensors. The authors in (Han et al., 2016) chose from a set of available participants to maximize sensing revenue under a limited budget. The authors in (Sun and Tang, 2019) proposed a greedy scheduling algorithm to find data-giver vehicles for every subtask with minimized cost in vehicular crowdsensing. The work in (Han et al., 2015) considered an online scheduling problem that determined sensing decisions for smartphones that were distributed over different regions of interest. (Nguyen and Zeadally, 2021). studied a participant selection problem that aimed to maximize the number of event records reported by fewer users. Different from (Han et al., 2015; Han et al., 2016; Sun and Tang, 2019; Nguyen and Zeadally, 2021), the work in (Gendy et al., 2020) aimed to maximize the percentage of the accomplished sensing tasks in a given period, by modeling the interaction between the participating devices and sensing task publishers as auctions. However, these works did not take into account the effects of dynamic wireless channels on sensing performance. Besides, most of them performed sensing device scheduling under the assumption that some statistical information is available in advance, which is usually resource-consuming and even impractical especially when the number of sensing devices is huge. Motivated by this fact, we aim to propose an online scheduling algorithm to find device scheduling decisions in crowdsensing tasks.

Recently, the rapid development of reinforcement learning (RL) techniques sheds light on the considered problem. Among these RL techniques, the multi-armed bandit (MAB) program is thought of as an important tool and has been widely adopted for scheduling and resource allocation problems. For example, MAB has been applied to advertisement placement, multi-antenna beam selection (Cheng et al., 2019), packet routing, offloading (Sun et al., 2018; Chen and Xu, 2019), caching (Blasco and Gündüz, 2014; Sengupta et al., 2014), and so on. In this work, we reformulate the sensing device scheduling problem as an MAB program, based on which a device scheduling algorithm is also proposed. The contributions of this work are summarized as follows.

- Considering the scarcity of wireless channel resources, we formulate a device scheduling problem in crowdsensing scenarios. We take into account not only the availability of devices caused by dynamic wireless channels but also fairness among the devices for better comprehensive observation of the area of interest. Besides, no prior information about devices is available.
- Then, the device scheduling problem is reformulated as an MAB problem, based on which an online scheduling algorithm is also proposed. The proposed algorithm propose incorporates the upper confidence bound (UCB) policy and virtual queue theory, whose regret performance is also analyzed in this work.
- Finally, simulation results are conducted to verify the effectiveness of the proposed algorithm. The balance between the time used to reach a point that meets the fairness constraints of devices and the performance regret is revealed.

System model

Consider a system consisting of an edge server and a set $\mathcal{K} = \{1, 2, \ldots, K\}$ of crowdsensing devices (e.g., sensors, cameras, and so on), as shown in **Figure 1**. These devices are responsible for collecting raw data from the observed events or objects and then pre-processing the raw data into samples, finally transmitting these samples to the edge server for processing tasks, such as statistical analysis and training a neural network for classification. For simplicity, we assume that the samples generated by different devices have the same size δ. Since the observed events can be periodic or aperiodic, or the observed objects have different activity characteristics, the amount of raw data collected by different devices is different. Other factors such as device location and perception ability also have influences on the amount of raw data collected by different devices. In addition, the processing capabilities of different devices are heterogeneous. Taking into account these facts mentioned above and for simplicity, we assume that time is slotted and the number of the newly generated samples of device $k \in \mathcal{K}$ in time slot τ, $N_k(\tau)$, is independently and identically distributed (i.i.d.) according to some unknown distribution whose expectation v_k is also unknown *a priori*. Thus, the total number of the samples waiting for uplink transmission of device k at the beginning of time slot τ is

$$M_k(\tau) = \min\left\{M_{\max}, \left[M_k(\tau-1) + N_k(\tau-1) - L_k(\tau-1)\right]^+\right\}, \quad (1)$$

where $[x]^+ = \max\{0, x\}$, $M_{\max}$ is the largest number of the samples that each device can store due to the limited storage space, and $L_k(\tau)$ is the number of the samples of device k has been transmitted the edge server in time slot τ, which will be specified in the following.

Transmission model

The orthogonal frequency-division multiple access technique is adopted and there are $F_{\max}$ orthogonal channels, each with the same bandwidth w, that can be used for uplink transmission simultaneously. The channel h_k between the edge server and device k is i.i.d., which is assumed to be constant within a time slot but varies independently across different time slots. The achievable uplink rate of device k in time slot τ is computed as

$$R_k(\tau) = w\log_2\left(1 + \frac{p_k|h_k(\tau)|^2}{\sigma^2}\right), \quad (2)$$

where σ^2 denotes the noise power and p_k denotes the transmit power of device k. Then, the number of samples that can be transmitted to the edge server is

$$l_k(\tau) = \min\left\{\frac{\Delta t R(\tau)}{\delta}, M_k(\tau) + N_k(\tau)\right\}, \tag{3}$$

where Δt is the duration length of a time slot.

Available channel constraint

When the edge server collects the generated samples from the devices, some devices can be unavailable for uplink transmission. For example, the devices experience poor channel conditions due to external interference, or the devices cannot work in the transmission mode when collecting raw data due to power constraints. We introduce the binary variable $a_k(\tau)$ to indicate the availability state of device k in time slot τ. Specifically, $a_k(\tau) = 1$ represents that device k can work in the transmission mode in time slot τ, otherwise not. Let $\mathcal{Z}(\tau) = \{k \in \mathcal{K} | a_k(\tau) = 1\} \in \mathcal{B}(\mathcal{K})$ denote the set of available devices that can work in the transmission mode in time slot τ where $\mathcal{B}(\mathcal{K})$ is the power set of $\mathcal{K}$. We assume the distribution of available devices, $\hat{P}_{\mathcal{Z}}(e) = \hat{P}(\mathcal{Z}(\tau) = e), e \in \mathcal{B}(\mathcal{K})$, is i.i.d. over time and unknown *a priori*, but $\mathcal{Z}(\tau)$ is unmasked to the edge server at the beginning of each time slot τ. Then, L_k is specified by

$$L_k(\tau) = \begin{cases} l_k(\tau) & \text{if } k \in \mathcal{Z}(\tau), \\ 0, & \text{else.} \end{cases} \tag{4}$$

In the considered system, there can be a huge number of devices, but the number of available channels at the same time is constrained. Due to the limited number of available channels, the edge server has to select a subset $\mathcal{W}(\tau)$ from the available devices, which should meet the available channel constraint, i.e.,

$$\mathcal{W}(\tau) \triangleq \{\mathcal{W}(\tau) \subseteq \mathcal{Z}(\tau) : |\mathcal{W}(\tau)| \le F_{\max}\} \in \mathcal{B}(\mathcal{Z}(\tau)), \tag{5}$$

where $|\mathcal{W}(\tau)|$ denotes the cardinality of $\mathcal{W}(\tau)$.

Fairness constraint

In order to achieve more comprehensive observation of the area of interest or better performance of computational tasks such as training a neural network, besides collecting as many samples as possible, the edge server is required to collect samples from different devices to ensure the diversity of samples. Thus, fairness among the devices is also an important issue that should be addressed in many practical applications. Here, a binary variable $b_k(\tau)$ is introduced with $b_k(\tau) = 1$ if device k is chosen to transmit its samples to the edge server in time slot τ, otherwise, $b_k(\tau) = 0$. With the definition of $b_k(\tau)$, we formulate the fairness constraint as follows

$$\lim_{T\to\infty} \inf \sum_{\tau=1}^{T} \mathbb{E}[b_k(\tau)] \ge c_k, \quad \forall k \in \mathcal{K}, \tag{6}$$

where T represents the total number of time slots, $c_k \in (0,1)$ represents the minimum of the portion of time slots required to transmit the samples of device k, and $\mathbb{E}[\bullet]$ is the expectation operator. We incorporate $c_k, k \in \mathcal{K}$ into a vector $\mathbf{c} = [c_1, c_2, \ldots, c_K]^T$ and c is thought of as a feasible fairness constraint if there is a policy which generates a decision sequence $\{\mathcal{W}(\tau), \tau \ge 1\}$ such that the fairness constraint 6) is satisfied.

Problem formulation

In this work, we aim to optimize a time sequence $\{\mathcal{W}(\tau), \tau \ge 1\}$ which maximizes the number of samples received at the edge server with a given time horizon of T time slots. The underlying problem with the fairness constraint and the available channel constraint can be formulated as

$$\max_{\{\mathcal{W}(\tau), \tau \ge 0\}} \sum_{\tau=1}^{T} \sum_{k \in \mathcal{W}(\tau)} L_k(\tau) \quad \text{s.t. (5) and (6),} \tag{7}$$

which is hard to solve because we have no idea about the distribution of the number of newly generated samples, as well as the distribution of wireless channels. Besides, the fairness constraint and the available channel constraint make problem Eq. 7 more challenging. Thanks to the development of the MAB framework, which sheds light on solutions to problem Eq. 7.

Proposed algorithm

In this section, we first introduce a stationary policy optimization program to deal with the uncertainty of device availability. Then, the device scheduling problem is reformulated as an MAB program, based on which an arm-pull algorithm is proposed to determine the decision sequence.

Problem reformulation

In this work, to simplify the scheduling complexity, a stationary policy named $\mathcal{Z}$-only policies is introduced, in which a super arm $\mathcal{W}(\tau) \in \mathcal{Y}(\mathcal{Z}(\tau))$ is selected according to the observed $\mathcal{Z}(\tau)$ only in each time slot τ (Neely, 2010), where $\mathcal{Y}(\mathcal{Z}(\tau))$ denotes the set of all possible subsets when $\mathcal{Z}(\tau)$ is observed. According to Theorem 4.5 in (Neely, 2010), if c belongs to the maximum feasibility region $\mathcal{C}$ strictly, a $\mathcal{Z}$-policy which can meet the fairness constraint in Eq. 6 always exists.

We further use a vector of probability distributions $\mathbf{q} = [q_{\mathcal{W}}(e), \forall \mathcal{W} \in \mathcal{Y}(e), \forall e \in \mathcal{B}(\mathcal{K})]$ to describe an $\mathcal{Z}$-only policy π with $\sum_{\mathcal{W} \in \mathcal{Y}(e)} q_{\mathcal{W}}(e) = 1, \forall e \in \mathcal{B}(\mathcal{K})$. Then, we compute the mean of $b_k(\tau)$ as

$$\mathbb{E}\left[b_k^{\pi}(\tau)\right] = \sum_{e \in \mathcal{B}(\mathcal{K})} \hat{P}_{\mathcal{Z}}(e) \sum_{\mathcal{W} \in \mathcal{Y}(e): k \in \mathcal{W}} q_{\mathcal{W}}(e), \tag{8}$$

and have an equivalent expression of constraint 6), i.e., $\mathbb{E}[b_k^{\pi}(\tau)] \ge c_k$. Besides, we assume $M_{\max}$ is large enough and define $\bar{l}_k = \frac{1}{M_{\max}} \mathbb{E}[l_k] \in [0,1]$ as the normalized expectation of L_k. Then, problem Eq. 7 can

be reformulated as

$$\begin{aligned}\max_{\mathrm{q}} \quad & \sum_{e\in\mathcal{B}(\mathcal{K})} \hat{P}_{\mathcal{Z}}(e) \sum_{\mathcal{W}\in\mathcal{Y}(e)} q_{\mathcal{W}}(e) \sum_{k\in\mathcal{W}} \bar{l}_k \\ \text{s.t.} \quad & \sum_{e\in\mathcal{B}(\mathcal{K})} \hat{P}_{\mathcal{Z}}(e) \sum_{\mathcal{W}\in\mathcal{Y}(e):k\in\mathcal{W}} q_{\mathcal{W}}(e) \geq c_k, \quad \forall k\in\mathcal{K}, \\ & \sum_{\mathcal{W}\in\mathcal{Y}(e)} q_{\mathcal{W}}(e) = 1, \quad \forall e\in\mathcal{B}(\mathcal{K}), \\ & q_{\mathcal{W}}(e)\in[0,1], \forall\mathcal{W}\in\mathcal{Y}(e), \quad \forall e\in\mathcal{B}(\mathcal{K}).\end{aligned} \tag{9}$$

which is a linear problem if the expectation $\bar{L}_k$ of L_k is known *a priori*. However, this assumption does not usually hold in practice and the edge server needs to estimate the average number of samples received from device k per time slot to make scheduling decisions. To address this issue, we introduce the MAB program.

Multi-armed bandit program

An MAB program is a machine learning framework where a player chooses a sequential of actions (arms) in order to maximize its cumulative reward in the long term (Lattimore and Szepesvári, 2020). Thankfully, we can model problem Eq. 7 as an MAB problem, in which the edge server and the devices play the roles of the player and the arms, respectively. Each subset $\mathcal{W}(\tau)$ of available arms is also treated as a super arm. Correspondingly, we can interpret the objective of problem Eq. 7 as determining a time sequence of the super arm to maximize the cumulative reward (i.e., the number of samples received at the edge server).

In the MAB program, there is an expected reward for each arm, but such statistical information is unknown by the player, which brings challenges to the arm selection of the player. The main basis that can be used to determine actions is some observation about the state in the current round and the experience gathered in previous rounds. More specifically, the arms which performed well in the past should be associated with higher priority. In the meantime, the player continues to explore the expected payoffs of the other arms. In other words, the player has to balance between the need to acquire more knowledge about the reward distributions of each arm (exploration) and the need to optimize rewards based on its current knowledge (exploitation) (Bubeck and Cesa-Bianchi, 2012). The exploration-exploitation dilemma inevitably causes performance loss and regret is the most popular metric for evaluating the learning performance in the MAB works, which is defined as the difference between the reward r^* and the average reward in a given period of time (Lai and Robbins, 1985). Here, r^* is the achievable maximum reward of problem Eq. 9 with the known $\bar{L}_k, \forall k\in\mathcal{K}$. Therefore, the original problem Eq. 7 can be reformulated as a cumulative regret minimization under policy π by determining a super arm $\mathcal{W}(\tau)$ in each time slot τ, i.e.,

$$\begin{aligned}\min_{\{\mathcal{W}(\tau),\tau\geq 1\}} \quad & R^{\pi} = Tr^* - \mathbb{E}\left\{\sum_{\tau=1}^{T}\sum_{k\in\mathcal{W}(\tau)} \hat{l}_k(\tau)\right\} \\ \text{s.t.} \quad & (5) \text{ and } (6),\end{aligned} \tag{10}$$

where $\hat{l}_k(\tau) = l_k(\tau)/M_{\max}$.

Algorithm design

When designing an algorithm for problem Eq. 10, three challenges need to be addressed: 1) how to maximize the cumulative reward when the reward expectation of each arm is unknown, 2) how to choose a super arm under the available channel constraint, and 3) how to meet the fairness constraint. The first two challenges can be dealt with with the extension of the classic UCB algorithm (Auer et al., 2002), but how to meet the fairness constraint requires the introduction of novel methods. Encouraged by (Neely, 2010; Li et al., 2019), the virtual queue technique has the potential to handle the fairness constraint. Specifically, a virtual queue is built for each arm k, i.e.,

$$Q_k(\tau) = \left[Q_k(\tau-1) + c_k - b_k(\tau-1)\right]^+, \tag{11}$$

where $[x]^+ = \max\{0,x\}$ and $D_k(\tau)$ represents the length of virtual queue of arm k at the beginning of time slot τ.

Define $\varrho_k(\tau) = \sum_{\tau'=1}^{\tau} b_k(\tau')$ as the number of times arm k has been chosen and $\upsilon_k(\tau)$ as the empirical mean of the reward of arm k by the end of time slot τ. The update rules of $\upsilon_k(\tau)$ and $\varrho_k(\tau)$ are given as

$$\upsilon_k(\tau) = \begin{cases} \dfrac{\upsilon_k(\tau-1)\varrho_k(\tau-1) + \hat{l}_k(\tau-1)}{\varrho_k(\tau-1)+1}, & \text{if } k\in\mathcal{W}(\tau), \\ \upsilon_k(\tau-1), & \text{else,} \end{cases} \tag{12}$$

and

$$\varrho_k(\tau) = \begin{cases} \varrho_k(\tau-1)+1, & \text{if } k\in\mathcal{W}(\tau), \\ \varrho_k(\tau-1), & \text{else,} \end{cases} \tag{13}$$

respectively. If $\varrho_k(\tau) = 0$, we set $\upsilon_k(\tau) = 0$. Note that both $\varrho_k(0)$ and $\upsilon_k(0)$ are initialized to be 0.

We estimate the mean reward of each arm k according to a truncated UCB method (Li et al., 2019), i.e.,

$$\hat{\upsilon}_k(\tau) = \min\left\{\upsilon_k(\tau-1) + \sqrt{\frac{2\ln\tau}{\varrho_k(\tau-1)}}, 1\right\}, \tag{14}$$

where $\hat{\upsilon}_k(\tau)$ is set to be 1, if $\varrho_k(\tau-1) = 0$. Then, a super arm is selected in each time slot τ according to

$$\mathcal{W}^*(\tau) \in \underset{\substack{\mathcal{W}\in\mathcal{Y}(\mathcal{Z}(\tau)),\\ |\mathcal{W}|=\min\{N,|\mathcal{Z}(\tau)|\}}}{\operatorname{argmax}} \sum_{k\in\mathcal{W}} (1-\alpha)\hat{\upsilon}_k(\tau) + \alpha Q_k(\tau) \tag{15}$$

where $\alpha\in(0,1]$ is a weighting value.

Finally, the whole algorithm is summarised in **Algorithm 1**.

Regret analysis

We first introduce a lemma that specifies the upper bound on the expected regret of the proposed algorithm.

Lemma 1. *The regret of the proposed algorithm is upper bounded by*

$$R^{\pi} \leq \alpha\frac{KT}{2} + (1-\alpha)\left[\left(\frac{\pi^2}{3}+1\right)K + 4\sqrt{2KFT\ln T}\right]. \tag{16}$$

Initialization: Set $\varrho_k(1) = \upsilon_k(1) = Q_k(1) = 0, \forall k \in \mathcal{K}$
for $\tau \in$ **do**
  for $k \in \mathcal{K}$ **do**
    if $v_k(\tau) > 0$ **then** update $\hat{v}_k(\tau)$ using **(14)**
    else set $\hat{v}_k(\tau) = 1$ **end if**
    Update Q_k using **(11)**
  end for
  Select $\mathcal{W}(\tau)$ using **(15)** and update $b_k(\tau), \forall k \in \mathcal{K}$
  Update $v_k(\tau)$ and $\varrho_k(\tau)$ using **(12)** and **13**, respectively
end for

Algorithm 1. Proposed algorithm for problem Eq. 10.

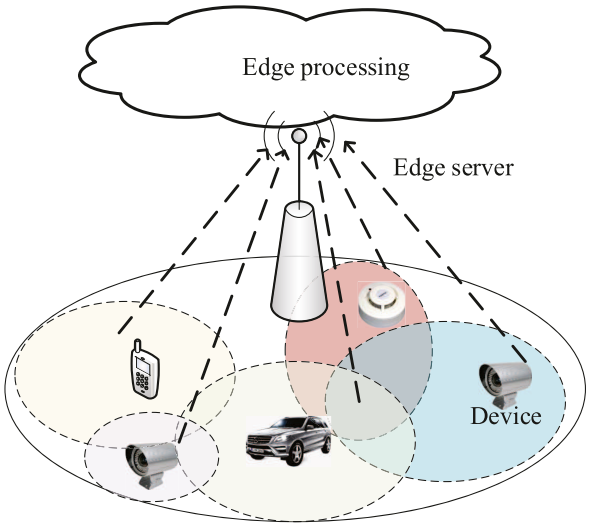


FIGURE 1
The considered system with an edge server and K devices. Different colored areas represent the perception areas of different devices.

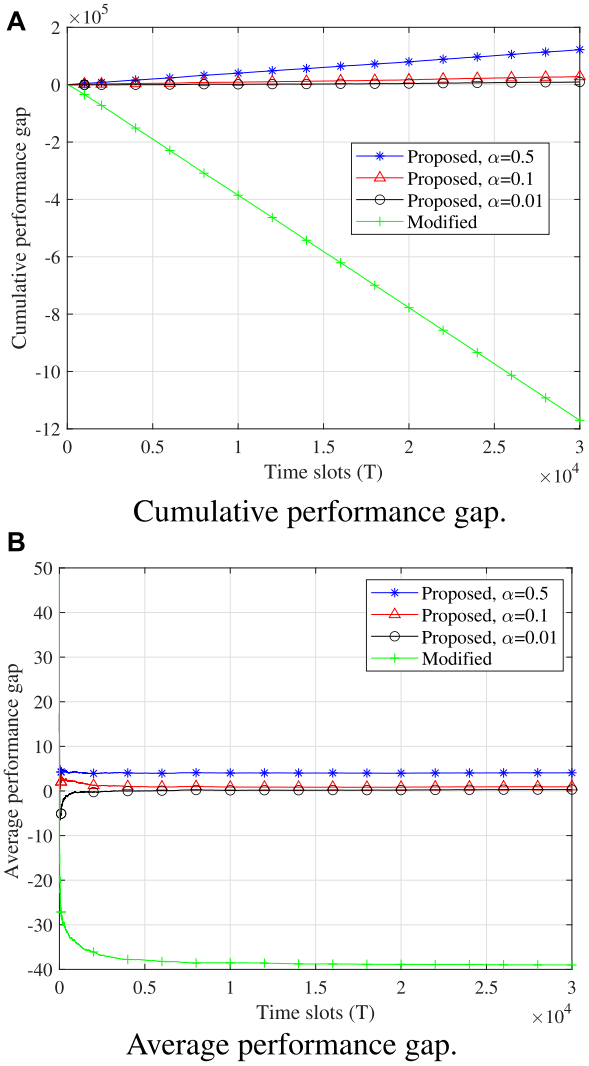


FIGURE 2
Performance comparison of different algorithms. **(A)** Cumulative performance gap over the time slots. **(B)** Average performance gap over the time slots.

Proof: Since similar proof has been presented in (Hsu et al., 2018; Li et al., 2019; Xia et al., 2021), here we only provide the sketch of the proof. Denote by $\mathcal{W}^*(\tau)$ the super arm selected according to the optimal $\mathcal{Z}$-policy $\pi^\star$ and by $b_k^*(\tau)$'s the corresponding the indicator variables. Then, we have

$$\begin{aligned} R^\pi &= \sum_{\tau=1}^{T} \mathbb{E}\left[\sum_{k\in\mathcal{W}^*(\tau)} \hat{l}_k(\tau) - \sum_{k\in\mathcal{W}(\tau)} \hat{l}_k(\tau)\right] \\ &= \sum_{\tau=1}^{T} \mathbb{E}\left[\sum_{k\in\mathcal{K}} \left(b_k^*(\tau) - b_k(\tau)\right)\bar{l}_k\right] \\ &\le \frac{\alpha K T}{2} + \sum_{\tau=1}^{T} \mathbb{E}\left[\Lambda_1(\tau)\right]. \end{aligned} \tag{17}$$

in which $\Lambda_1(\tau) = \sum_{k\in\mathcal{K}}[\alpha Q_k(\tau) + (1-\alpha)\bar{l}_k]\left(b_k^*(\tau) - b_k(\tau)\right)$ (Li et al., 2019, Appednix C). Then, according to (Xia et al., 2021, Lemma 2), we have

$$\Lambda_1(\tau) \le (1-\alpha)\left[\Lambda_2(\tau) + \Lambda_3(\tau)\right], \tag{18}$$

where $\Lambda_2(\tau) = \sum_{i\in\mathcal{W}(\tau)}\left(\hat{v}_k(\tau) - \bar{l}_k\right)$, $\Lambda_3(\tau) = \sum_{i\in\mathcal{W}^\ddagger(\tau)}\left(\bar{l}_k - \hat{v}_k(\tau)\right)$, and $\mathcal{W}^\ddagger(\tau)$ is chosen according to the following rule:

$$\mathcal{W}^\ddagger(\tau) \in \underset{\mathcal{W}\in\mathcal{Y}(\mathcal{Z}(\tau))}{\operatorname{argmax}} \sum_{k\in\mathcal{W}} \alpha Q_k(\tau) + (1-\alpha)\bar{l}_k. \tag{19}$$

Here, the upper bounds of $\Lambda_2(\tau)$ and $\Lambda_3(\tau)$ are directly given as follows:

$$\begin{aligned} &\sum_{\tau=1}^{T} \mathbb{E}\left[\Lambda_2(\tau)\right] \le \left(\frac{\pi^2}{6} + 1\right)K + 4\sqrt{2KFT\ln T}, \\ &\sum_{\tau=1}^{T} \mathbb{E}\left[\Lambda_3(\tau)\right] \le \frac{\pi^2}{6}K. \end{aligned} \tag{20}$$

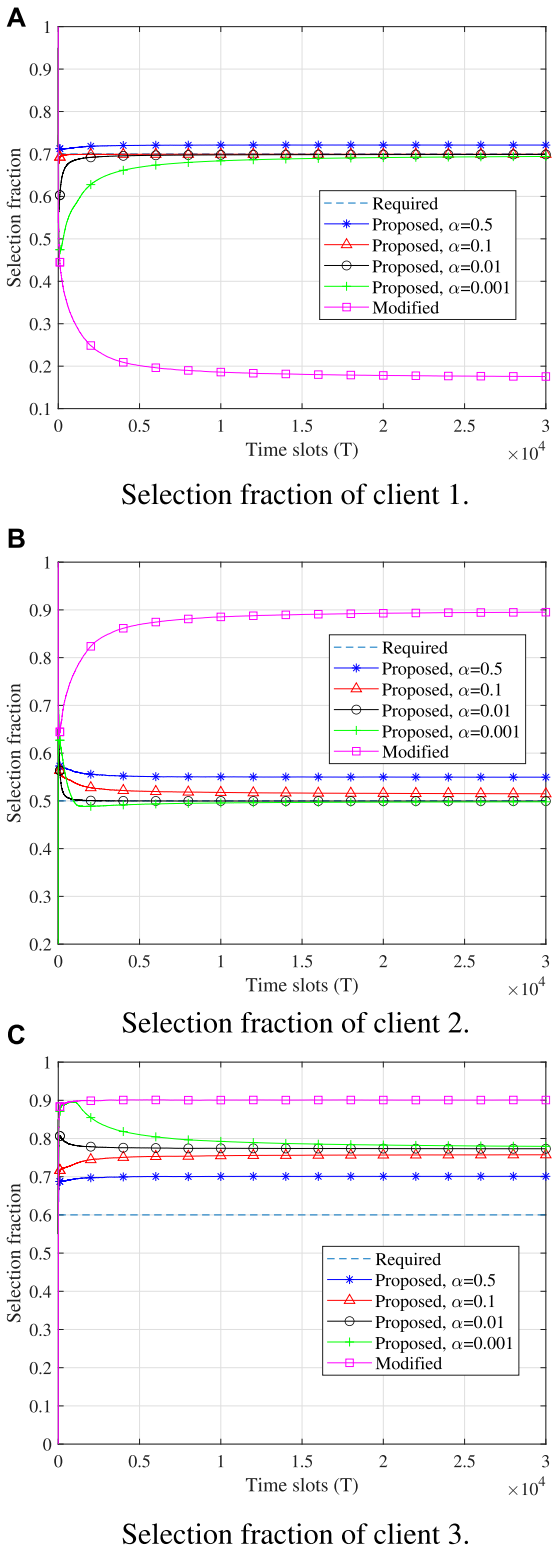


FIGURE 3
Client selection over the time slots. **(A)** Section fraction of client 1. **(B)** Section fraction of client 2. **(C)** Section fraction of client 3.

The corresponding analysis is similar to that in (Li et al., 2019; Appendices D and E). Finally, we finish the proof by substituting Eq. 20 into Eq. 18 and further into Eq. 17.

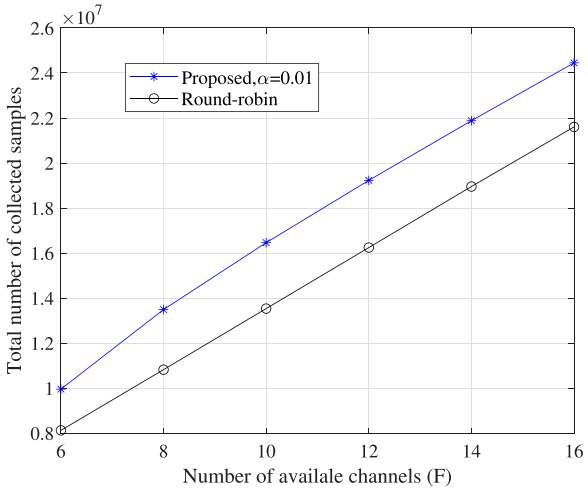


FIGURE 4
Performance comparison of the proposed algorithm with the round-robin algorithm.

Remark 1. Given $0 < \alpha \leq \frac{1}{\sqrt{T}}$ and a large value T, then we can simplify the upper bound in Eq. 16 as

$$R^{\pi} \leq \frac{K\sqrt{T}}{2} + \left(\frac{\pi^2}{3} + 1\right)K + 4\sqrt{2KFT\ln T}, \tag{21}$$

which suggests that the time-average performance regret increases at a sub-linear rate (i.e., $\mathcal{O}(\sqrt{T\ln T})$) over time.

Simulation results

In this section, we provide simulation results to verify the effectiveness of the proposed algorithm. We consider a disc area with a radius of 200 m and a single-antenna access point (AP) equipped with an edge server located in the center of the considered area. The transmit power of each sensing device k is set as 23 dBm and the noise power σ^2 is set as −107 dBm. The channel response h_k is computed as $h_k = \sqrt{\beta_k}\tilde{h}_k$ where $\tilde{h}_k$ and β_k stand for small-scale fading and large-scale fading, respectively. The small-scale fading is represented by i.i.d. zero-meaned complex Gaussian variables with unit variance. The large-scale fading is determined according to the path-loss model: PL [dB] = $128.1 + 37.6\log_{10}(d)$ where d stands for the distance in km (Dahrouj and Yu, 2010). The number of the newly generated samples $N_k(\tau)$ is assumed to be uniformly distributed in $[N_k^{\mathrm{LB}}, N_k^{\mathrm{UB}}]$, where N_k^{LB} and N_k^{UB} are set as $N_k^{\mathrm{LB}} = (0.5k + 0.5) \times 20$ and $N_k^{\mathrm{UB}} = (0.5k + 1.5) \times 80, \forall k \in \mathcal{K}$. We also assume the availability of each sensing device to be i.i.d. using a binary random variable with a mean of 0.9. Besides, we assume $M_{\max} = 500$ samples, $\delta = 100$ bits/sample, and the length of a time slot $\Delta t = 0.1$ s.

We consider a system with $K = 3$ sensing devices randomly distributed within the coverage of the AP. However, only $F = 2$ channel links are available and the bandwidth of each orthogonal channel is set as 15 KHz. The fairness constraint factors are $c_1 = 0.7$, $c_2 = 0.5$, and $c_3 = 0.6$. Here, we define Ω_1 and Ω_2 as

the cumulative performance gap and average performance gap, respectively, with $\Omega_1 = \Sigma^{\pi} M_{\max}$ and $\Omega_2 = \Omega_1 / T$, which are used to describe the difference between the optimal solution found by solving problem Eq. 9 and the solution found using the proposed algorithm (or baseline algorithms). Note that the optimal solution found by solving problem Eq. 9 satisfies the fairness constraint. For comparison, we introduce a modified version of the proposed algorithm, which does not take into account the fairness constraint. More specifically, the UCB algorithm for the modified version does not introduce the virtual queue technique. In Figure 2, we compare the proposed algorithms under different α values with the modified version. We find that the performance gap of the modified algorithm is the smallest, whose value is even negative because the modified algorithm does not need to meet the fairness constraint and may lead to biased observations of the area of interest. We also observe that the time-average performance gap of the proposed algorithm grows at a sub-linear trend. Besides, at first glance, the proposed algorithm with a smaller α value meets the fairness constraint but also enjoys better performance, which is more attractive. This is because a smaller α value makes the reward of each device dominant and the fairness constraint insignificant. However, what is missing in Figure 2 is the convergence time used to meet the fairness constraint, which is also an important metric that should be taken into account in practice.

Figure 3 shows the change in the selection fractions of different devices over the time slots. Here, the selection fraction is defined as the ratio of the chosen time of a certain device to the total number of time slots. We find that the curves of all the arms obtained by the proposed algorithm meet the fairness constraints eventually, no matter which α value is taken. The modified algorithm has no idea of the fairness constraints and thus does not need to meet the fairness constraints. In addition, it is observed that a smaller α value leads to more time consumption before the convergence is achieved and the convergence rate of the curve with $\alpha = 0.001$ is the slowest.

To further validate the effectiveness of the proposed algorithm, we consider a scenario with $K = 20$ devices and introduce the round-robin algorithm as a baseline, as shown in Figure 4. According to the results in Figure 4, we find that more samples are collected with the increase of the number of available channels. In addition, the proposed algorithm always achieves better performance than the round-robin algorithm.

Conclusion

In this work, we considered sensing device scheduling problem in mobile crowdsensing tasks, which suffers from the scarcity of wireless channel resource and the lack of prior knowledge, as well as different capabilities among devices. To address these challenges, we reformulated the device scheduling problem as an MAB program, one should guarantee the participation fairness of sensing devices with different coverage regions. Then, we proposed a device scheduling algorithm on the basis of the UCB policy and virtual queue theory, whose performance regret was also analyzed. Finally, numerical results were conducted to verify the effectiveness of the proposed algorithm.

Data availability statement

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

Author contributions

JZ contributed to the conception and prepared the first draft of the manuscript. YN performed the numerical simulations. HZ improved the writing of the manuscript. All authors approved the submitted version of the manuscript.

Funding

This work was supported in part by National Key R&D Program of China (2021ZD0140405), in part by Natural Science Foundation of Jiangsu Province (BK2021022532), in part by Jiangsu University Philosophy and Social Science Research Fund (2022SJYB0517).

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

Auer, P., Cesa-Bianchi, N., and Fischer, P. (2002). Finite-time analysis of the multiarmed bandit problem. *Mach. Learn.* 47, 235–256. doi:10.1023/a:1013689704352

Blasco, P., and Gündüz, D. (2014). "Learning-based optimization of cache content in a small cell base station," in 2014 IEEE International Conference on Communications (ICC), Sydney, Australia, 10-14 June 2014 (IEEE), 1897–1903.

Bubeck, S., and Cesa-Bianchi, N. (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Found. Trends® Mach. Learn.* 5, 1–122. doi:10.1561/2200000024

Chen, L., and Xu, J. (2019). "Task replication for vehicular cloud: Contextual combinatorial bandit with delayed feedback," in IEEE INFOCOM 2019 - IEEE

Conference on Computer Communications, Paris, France, 29 April 2019 - 02 May 2019, 748–756.

Cheng, M., Wang, J.-B., Wang, J.-Y., Lin, M., Wu, Y., and Zhu, H. (2019). "A fast beam searching scheme in mmwave communications for high-speed trains," in IEEE Int. Conf. Commun. (ICC), Shanghai, China, 29 April 2019 - 02 May 2019, 1–6.

Chu, E.-H., Lin, C.-Y., Tsai, P.-H., and Liu, J. (2013). Participant selection for crowdsourcing disaster information. *WIT Trans. Built Environ.* 133, 231–240.

Dahrouj, H., and Yu, W. (2010). Coordinated beamforming for the multicell multi-antenna wireless system. *IEEE Trans. Wirel. Commun.* 9, 1748–1759. doi:10.1109/TWC.2010.05.090936

Gendy, M. E., Al-Kabbany, A., and Badran, E. F. (2020). "Maximizing clearance rate by penalizing redundant task assignment in mobile crowdsensing auctions," in Proc. IEEE Wireless Communications and Networking Conference (WCNC), Seoul, Korea (South), 25-28 May 2020, 1–7.

Guo, B., Yu, Z., Zhou, X., and Zhang, D. (2014). "From participatory sensing to mobile crowd sensing," in Proc. IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS), Budapest, Hungary, 24-28 March 2014, 593–598.

Han, K., Zhang, C., and Luo, J. (2016). Taming the uncertainty: Budget limited robust crowdsensing through online learning. *IEEE/ACM Trans. Netw.* 24, 1462–1475. doi:10.1109/tnet.2015.2418191

Han, Y., Zhu, Y., and Yu, J. (2015). "Utility-maximizing data collection in crowd sensing: An optimal scheduling approach," in Proc. IEEE Int. Conf. Sensing, Communication, and Networking (SECON), Seattle, USA, 22-25 June 2015, 345–353.

Hsu, W., Xu, J., Lin, X., and Bell, M. R. (2018). "Integrating online learning and adaptive control in queueing systems with uncertain payoffs," in *Inf. Theory appli. Workshop (ITA)* (San Diego, CA, USA, 1–9.

Lai, T. L., and Robbins, H. (1985). Asymptotically efficient adaptive allocation rules. *Adv. Appl. Math.* 6, 4–22. doi:10.1016/0196-8858(85)90002-8

Lattimore, T., and Szepesvári, C. (2020). *Bandit algorithms*. Cambridge: Cambridge University Press.

Li, F., Liu, J., and Ji, B. (2019). Combinatorial sleeping bandits with fairness constraints. *Proc. IEEE Conf. Comput. Commun.* 7, 1702–1710.

Neely, M. J. (2010). Stochastic network optimization with application to communication and queueing systems. *Synth. Lect. Commun. Netw.* 3, 1–211. doi:10.2200/s00271ed1v01y201006cnt007

Nguyen, T. N., and Zeadally, S. (2021). Mobile crowd-sensing applications: Data redundancies, challenges, and solutions. *ACM Trans. Internet Technol.* 22, 1–15. doi:10.1145/3431502

Sengupta, A., Amuru, S., Tandon, R., Buehrer, R. M., and Clancy, T. C. (2014). "Learning distributed caching strategies in small cell networks," in *Proc. Int. Symp. Wireless commun. Syst. (ISWCS)* (Barcelona, Spain, 917–921.

Sun, Y., Song, J., Zhou, S., Guo, X., and Niu, Z. (2018). "Task replication for vehicular edge computing: A combinatorial multi-armed bandit based approach," in 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 09-13 December 2018, 1–7.

Sun, Y., and Tang, Y. (2019). "Task-oriented data collection strategy in vehicular crowdsensing," in Proc. Int. Conf. Computer Science and Education (ICCSE), Toronto, Canada, 19-21 August 2019, 761–766.

Xia, W., Wen, W., Wong, K.-K., Quek, T. Q., Zhang, J., and Zhu, H. (2021). Federated-learning-based client scheduling for low-latency wireless communications. *IEEE Wirel. Commun.* 28, 32–38. doi:10.1109/mwc.001.2000252

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 03 April 2023
DOI 10.3389/fenrg.2023.1160256

OPEN ACCESS

EDITED BY
Xiao-Kang Liu,
Huazhong University of Science and Technology, China

REVIEWED BY
Yao Weitao,
Nanyang Technological University, Singapore
Rui Wang,
Northeastern University, China

*CORRESPONDENCE
Yunlong Han,
yunlonghan_SIST@126.com

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 07 February 2023
ACCEPTED 23 March 2023
PUBLISHED 03 April 2023

CITATION
Han Y (2023), Wavelet analysis and consensus algorithm-based fault-tolerant control for smart grids. *Front. Energy Res.* 11:1160256. doi: 10.3389/fenrg.2023.1160256

Wavelet analysis and consensus algorithm-based fault-tolerant control for smart grids

Yunlong Han*

Department of Basic and General Knowledge, Shenyang Institute of Science and Technology, Shenyang, China

In this paper, the voltage and frequency regulation problems are investigated for smart grids under the influence of faults. To solve those problem, a wavelet analysis and consensus algorithm-based fault-tolerant control scheme is proposed. Specifically, the wavelet analysis technique is introduced to determine whether there exist faults or not in the smart grids. Then, a distributed fault estimator is designed to estimate the attack signals. Based on this estimator state, a distributed fault-tolerant controller is designed to compensate for the faults. It is theoretically shown that the developed method can achieve the voltage regulation and frequency objectives. Finally, a smart grid with four distributed generations is constructed in MATLAB/Simulink for simulation to validate the effectiveness.

KEYWORDS

secondary control, fault-tolerant control, voltage regulation, distributed control, AC micro grid

1 Introduction

During the past decades, microgrid (MG) has attracted much attention because of their reliability, efficiency, and so on. AC MG consists of various distributed generation (DG), including energy storage, wind farm, and photovoltaic based on the power inverter. To ensure the stability of the MG, it is imminent for the engineers and scientists to design some reasonable control strategies. Please refer to the classical results on AC MG for details, such as (1–3).

For the AC MG, the hierarchical control method is well-used, i.e., the control is divided into the primary control, the secondary control, and the tertiary control. Specifically, the objective of the primary control is to design the droop control method such that the output voltage and frequency of DG track the corresponding reference signal (4). However, if only the droop control is used, then the deviation of the voltage and frequency will happen. To solve the problem, the secondary controller is usually introduced, which is implemented by introducing a consensus-based algorithm to regulate the voltage and frequency of their reference values. Finally, the tertiary control is to achieve optimal scheduling and economic indicators in a larger time range. This paper mainly focuses on the secondary control layer.

In general, the design methods of the secondary controller can be divided into the centralized control method and the distributed control method. The centralized one is to design the control method in a centralized manner (5–7), i.e., the controller of each DG needs the information of all DGs, which leads to this controller containing some disadvantages such as high computational complexity, high communication cost, and so on. To overcome those drawbacks, distributed secondary control method is proposed (8–11), where the

controller of each DG only needs the information from neighbor DGs, which contains the advantages of low communication cost and low computational complexity (12, 13).

Note that as the increase of the number of DGs in AC MG, system faults will inevitably occur, which may lead to the damage of the system performance (14). To ensure the reliability of the AC MG, some fault-tolerant secondary control methods have been developed. The existing fault-tolerant secondary control procedure can be divided into fault detection, fault estimation, and fault-tolerant, which is usually done in the above order. The wavelet analysis method as a fault detection method has been widely used to detect system faults. Based on the detection result, the fault estimation method and the fault-tolerant method will start work. In (16), a robust detector is designed to detect the consistency of measurements through the PMU measurements. In (17), a prototype method is developed to detect faults by using the candidate invariant to compare with the actual invariant. In (18), a vulnerability factor-based distributed fault detection method is proposed to detect the system faults. In addition, some results on fault-tolerant secondary control have been developed. In (19), a distributed fault-tolerant secondary controller is designed to achieve the frequency and voltage restoration by using the adaptive technique. In (20), a fault-tolerant cooperative secondary control scheme is developed for islanded MG with system faults by proposing a hidden layer.

Notice that the results mentioned above focus only on the fault detection or the fault-tolerant of the MGs separately. Besides, different from the faults on the inter-DG communication in MGs, the system faults on the secondary controller are usually more difficult to be detected and eliminated. Although some existing results have been developed to restore the voltage and frequency, the transient performance cannot be ensured and there are large fluctuations which are harmful to MG (22, 23). Based on the above observation, this paper considers the voltage and frequency regulation problems for AC MGs under the influence of faults. To solve this problem, the main method and the main contributions of this paper can be summarized as follows.

- To solve the reliability problem of AC MGs with system faults, a series of methods have been proposed, i.e., a wavelet analysis method is introduced to detect the system faults, the fault estimator is proposed to estimate the unknown faults, and the fault-tolerant controller is designed to eliminate the influence of faults.
- A k-step distributed fault estimator is proposed for each DG to estimate the fault signal. The main feature of this estimator is that with the increase of the number of k, the estimation error will be smaller and smaller until it converges to zero.
- Based on the developed k-step estimator, a distributed fault fault-tolerant secondary controller is designed to compensate for the influence of faults in AC MGs. Besides, the stability of the system is proved by the Lyapunov theorem.

The rest sections are organized as follows. In Section 2, the preliminaries are presented. In Section 3, the fault estimation design method is given. Then, a fault-tolerant controller is designed in Section 4. In Section 5, the simulation test is studied. Finally, the conclusion is given in Section 6.

2 Preliminaries

2.1 Graph theory

In this paper, the considered AC MGs is consisted of networks with DGs and leader nodes. The network topology is fixed and is defined by an undirected $\mathcal{G} = (\mathcal{W}, \mathcal{E}, \mathcal{A})$, where $\mathcal{W} = \{1, 2, \ldots, N\}$ represents the set of nodes, $\mathcal{E} \subset \mathcal{W} \times \mathcal{W}$ means the set of edges, and $\mathcal{A} = [a_{ij}]$ denotes the adjacency matrix. If there exists communication from i to node j, then the edge weight $a_{ij} = 1$; otherwise, $a_{ij} = a_{ji} = 0$. The in-degree matrix and the Laplacian matrix are denoted by $D = \mathrm{diag}(d_1, \ldots, d_N)$ and $\mathcal{L} = D - \mathcal{A}$, where $d_i = \sum_{j=1}^{N} a_{ij}$.

Assumption 1. *The undirected network topology $\mathcal{G}$ is connected.*

2.2 System Model

In this paper, an AC MG is considered. It is assumed that the communication among DGs is ideal and the inductive distribution lines are considered. Besides, the active power and reactive power delivered by DG i at bus i satisfy

$$\begin{cases} P_i = \dfrac{v_{\mathrm{bus},i} v_{\mathrm{mag},i} \sin(\theta_i - \beta_i)}{Z_i}, \\ Q_i = -\dfrac{v_{\mathrm{bus},i}^2}{Z_i} \dfrac{v_{\mathrm{bus},i} v_{\mathrm{mag},i} \cos(\theta_i - \beta_i)}{Z_i}, \end{cases} \tag{1}$$

where P_i and Q_i represent the active power and reactive power, respectively. $v_{\mathrm{mag},i}\angle\theta_i$ and $v_{\mathrm{bus},i}\angle\beta_i$ are the output voltage and bus voltage of DG i. Z_i represents the effective collective impedance of the output filter and the connector between distribution network and DG.

2.3 Control objective

The secondary control objective is to design secondary controllers of frequency and voltage regulation such that i) the frequency restores to its reference value ω_{ref} as frequency reference value; ii) the voltage regulates to $[v_{ref}^1, v_{ref}^2]$, where v_r^1 and v_{ref}^2 are the upper and lower bounds of voltage amplitude.

2.4 Cooperative control solution

Applying the droop control to (1), the following control scheme can be achieved

$$\begin{cases} \omega_i = \omega_{n_i} - n_{P_i} P_i, \\ v_{oi} = V_{n_i} - n_{Q_i} Q_i, \end{cases} \tag{2}$$

where ω_i represents the angular frequency. ω_{n_i} and V_{n_i} represent the set points in the secondary controller. n_{P_i} and n_{Q_i} represent droop coefficients chosen according to the power rating of the corresponding DG(17). By using an appropriate reference coordinate transformation, it is reasonable to adjust $v_{mag,i}$ by v_{oi}.

Driven by the consensus algorithm, the frequency error and voltage error among DGs are defined as follows.

$$e_{\omega,i} = c_\omega \left(\sum_{j\in\mathcal{N}_i} \left(\omega_j - \omega_i\right) + g_i\left(\omega_r - \omega_i\right) + \sum_{j\in\mathcal{N}_i} \left(n_{P_j} P_j - n_{P_i} P_i\right) \right), \tag{3}$$

$$e_{v,i} = c_v \left(\sum_{j\in\mathcal{N}_i} \left(v_{oj} - v_{oi}\right) + g_i\left(v_r - v_{oi}\right) + \sum_{j\in\mathcal{N}_i} \left(n_{Q_j} Q_j - n_{Q_i} Q_i\right) \right), \tag{4}$$

where ω_i and v_{oi} are the frequencies and output voltages of the DGs i-th DG. The positive constants c_ω and c_v are the coupling gains. ω_r and v_r are the frequency and voltage references, respectively.

Together with (2–4), it is shown that

$$e_{\omega,i} = c_\omega \left(\sum_{j\in\mathcal{N}_i} \left(\omega_{n_j} - \omega_{n_i}\right) + \sum_{\gamma=1}^{2} g_i\left(\omega_{n,\mathrm{ref}} - \omega_{n_i}\right) \right), \tag{5}$$

$$e_{v,i} = c_v \left(\sum_{j\in\mathcal{N}_i} \left(V_{n_j} - V_{n_i}\right) + \sum_{\gamma=1}^{2} g_i\left(V^{\gamma}_{n,\mathrm{ref}} - V_{n_i}\right) \right), \tag{6}$$

where $\omega_{n,\mathrm{ref}} = \omega_r + n_{P_i} P_i$ and $V^{\gamma}_{n,\mathrm{ref}} = v^{\gamma}_{\mathrm{ref}} + n_{Q_i} Q_i$[1].

To restore the frequency and voltage of all DGs to the reference values, ω_{ni} and V_{ni} are provided by the secondary controller, which are driven by the information from neighbors and updated according to the frequency and voltage error terms $e_{\omega,i}$and $e_{v,i}$, respectively. From (2), we have

$$\begin{cases} \dot{\omega}_{ni} = \dot{\omega}_i + n_{P_i}\dot{P}_i = e_{\omega,i}, \\ \dot{V}_{ni} = \dot{v}_{\mathrm{oi}} + n_{Q_i}\dot{Q}_i = e_{v,i}. \end{cases} \tag{7}$$

Inspired by (21), we can design ω_{n_i} to stabilize ω_i and $n_{P_i}P_i$ and V_{n_i} to stabilize v_{oi} and $n_{Q_i}Q_i$. Combining (7) with (5, 6), the following distributed secondary controller can be designed

$$\dot{\omega}_{n_i} = c_\omega \left(\sum_{j\in\mathcal{N}_i} \left(\omega_{n_j} - \omega_{n_i}\right) + \sum_{\gamma=1}^{2} g_i\left(\omega_{n,\mathrm{ref}} - \omega_{n_i}\right) \right), \tag{8}$$

$$\dot{V}_{n_i} = c_v \left(\sum_{j\in\mathcal{N}_i} \left(V_{n_j} - V_{n_i}\right) + \sum_{\gamma=1}^{2} g_i\left(V^{\gamma}_{n,\mathrm{ref}} - V_{n_i}\right) \right). \tag{9}$$

Lemma 1. *Under Assumption 1, the designed distributed secondary controllers (8, 9) can regulate the frequency to its reference value and stabilize the voltage amplitude between the upper and lower bounds.*

1 The details of (5) and (6) are given in the Appendix.

3 Fault estimator design

3.1 Fault Model and analysis

Usually, faults may occur with the increase of system operation time in the secondary controller. In this paper, the fault signal is denoted by f_i, and the following conditions are satisfied.

Assumption 2. *The fault is changed slowly, i.e., the fault signal and its derivative are norm bounded as two positive scalars a and b ($\|f\| \le a$ and $\|\dot{f}\| \le b$).*

Under the influence of faults, the secondary controller of voltage regulation comes

$$\begin{aligned} \dot{V}_i &= c_v \left(\sum_{j\in\mathcal{N}_i} \left(V_j - V_i\right) + \sum_{\gamma=1}^{2} g_i\left(V_r - V_i\right) \right) + f_i \\ &= c_v \left(\sum_{\gamma=1}^{2} g_i V_r - \left(\left(d_i + \sum_{\gamma=1}^{2} g_i \right) V_i - \sum_{j\in\mathcal{N}_i} V_j \right) \right) + f_i. \end{aligned} \tag{10}$$

Denote $V = [V_1, \ldots, V_N]^T$, and $f = [f_1, \ldots, f_N]^T$. Then, (10) can be rewritten as follows,

$$\begin{aligned} \dot{V} = c_v \big(& G_1\left(\mathbf{1}_N \otimes v_r\right) + G_2\left(\mathbf{1}_N \otimes V^2_{\mathrm{ref}}\right) \\ & - \left(L + G_1 + G_2\right) V \big) + f. \end{aligned} \tag{11}$$

Define $\Theta_\gamma = \frac{1}{2}L + G_\gamma, \gamma = 1, 2$. Then, it has $L\left(\mathbf{1}_N \otimes V^{\gamma}_{\mathrm{ref}}\right) = (D - A)\left(\mathbf{1}_N \otimes V^{\gamma}_{\mathrm{ref}}\right) = 0, \forall\gamma = 1, 2$. Thus, (10) is rewritten as

$$\begin{aligned} \dot{V} &= c_v \Big(\left(\frac{1}{2}L + G_1 \right)\left(\mathbf{1}_N \otimes v_r\right) + \left(\frac{1}{2}L + G_2 \right) \\ &\quad \times \left(\mathbf{1}_N \otimes V^2_{\mathrm{ref}}\right) - \left(\frac{1}{2}L + G_1 + \frac{1}{2}L + G_2 \right) V \Big) + f \\ &= c_v \left(\sum_{\gamma=1}^{2} \Theta_\gamma \left(\mathbf{1}_N \otimes V_r\right) - \sum_{\gamma=1}^{2} \Theta_\gamma V \right) + f. \end{aligned} \tag{12}$$

It can be seen that the normal secondary controller cannot achieve the voltage regulation objective under the influence of faults.

3.2 Fault estimators design

To estimate the voltage and fault of DG i, the following fault estimators can be designed.

$$\begin{cases} \dot{\hat{V}}_{i,0} = c_v \left(\sum_{j\in\mathcal{N}_i} \left(\hat{V}_{j,0} - \hat{V}_{i,0}\right) + g_i\left(V_r - \hat{V}_{i,0}\right) \right) + b_v \tilde{V}_{i,0} + \hat{f}_{i,0}, \\ \dot{\hat{f}}_{i,0} = -d_v \left(\sum_{j\in\mathcal{N}_i} \left(\tilde{V}_{j,0} - \tilde{V}_{i,0}\right) - g_i \tilde{V}_{i,0} \right), \end{cases} \tag{13}$$

where the positive constants b_v and d_v are coupling gains. $\hat{V}_{i,0}$ and $\hat{f}_{i,0}$ are the initial estimator states of DG i and $\tilde{V}_{i0} = V_i - \hat{V}_{i0}$ is the estimation error of V_i.

During the above estimator, the effect of fault-dependent term $\dot{f}$ is ignored, which may result in estimation bias. To eliminate this influence, the following k-step estimators are introduced.

$$\begin{cases} \dot{\hat{V}}_{i,k} = c_v \left(\sum_{j\in\mathcal{N}_i} \left(\hat{V}_{jk} - \hat{V}_{i,k}\right) + g_i\left(V_r - \hat{V}_{i,k}\right) \right) + b_v \tilde{V}_{i,k} + \hat{f}_{i,k}, \\ \dot{\hat{f}}_{i,k} = -d_v \left(\sum_{j\in\mathcal{N}_i} \left(\tilde{V}_{jk} - \tilde{V}_{i,k}\right) - g_i \tilde{V}_{i,k} \right) + \dot{\hat{f}}_{ik-1}, \end{cases} \tag{14}$$

where $\hat{V}_{i,k}$ and $\hat{f}_{i,k}$ are the k-step estimator states of DG i. $\tilde{V}_{i,k} = V_i - \hat{V}_{i,k}$ represents the estimation error.

Denote $\hat{V}_0 = [\hat{V}_{1,0}, \ldots, \hat{V}_{N,0}]$ and $\hat{f}_0 = [\hat{f}_{1,0}, \ldots, \hat{f}_{N,0}]$. The dynamics of (13) can be rewritten as,

$$\begin{cases} \dot{\hat{V}}_0 = c_v \left(\sum_{\gamma=1}^{2} \Theta_\gamma (\mathbf{1}_N \otimes V_r) - \sum_{\gamma=1}^{2} \Theta_\gamma \hat{V}_0 \right) + b_v \tilde{V}_0 + \hat{f}_0, \\ \dot{\hat{f}}_0 = d_v \sum_{\gamma=1}^{2} \Theta_\gamma \tilde{V}_0. \end{cases} \tag{15}$$

Define $\hat{V}_k = [\hat{V}_{1,k}, \ldots, \hat{V}_{N,k}]$ and $\hat{f}_k = [\hat{f}_{1,k}, \ldots, \hat{f}_{N,k}]$. The dynamic of (14) can be rewritten as

$$\begin{cases} \dot{\hat{V}}_k = c_v \left(\sum_{\gamma=1}^{2} \Theta_\gamma (\mathbf{1}_N \otimes V_r) - \sum_{\gamma=1}^{2} \Theta_\gamma \hat{V}_k \right) + b_v \tilde{V}_k + \hat{f}_k, \\ \dot{\hat{f}}_k = d_v \sum_{\gamma=1}^{2} \Theta_\gamma \tilde{V}_k + \dot{\hat{f}}_{k-1}. \end{cases} \tag{16}$$

Define $\tilde{f}_0 = f - \hat{f}_0$. From (15), it yields

$$\begin{cases} \dot{\tilde{V}}_0 = -\left(c_v \sum_{\gamma=1}^{2} \Theta_\gamma + b_v \right) \tilde{V}_0 + \tilde{f}_0, \\ \dot{\tilde{f}}_0 = -d_v \sum_{\gamma=1}^{2} \Theta_\gamma \tilde{V}_0 + \dot{f}. \end{cases} \tag{17}$$

Define $\tilde{f}_k = f - \hat{f}_k$, for $k = 1, 2, \ldots$. It follows from (16) that

$$\begin{cases} \dot{\tilde{V}}_k = -\left(c_v \sum_{\gamma=1}^{2} \Theta_\gamma + b_v \right) \tilde{V}_k + \tilde{f}_k, \\ \dot{\tilde{f}}_k = -d_v \sum_{\gamma=1}^{2} \Theta_\gamma \tilde{V}_k + \dot{\tilde{f}}_{k-1}. \end{cases} \tag{18}$$

From the above estimators, the following theorem can be achieved.

Theorem 1. *Under Assumptions 1 and 2, the proposed iterative estimators (13, 14) ensure that the estimation errors converge to zero as the iteration step k tends to ∞, i.e.,*

$$\lim_{k\to\infty} \left(f - \frac{1}{k+1} \sum_{s=0}^{k} \hat{f}_s \right) = 0. \tag{19}$$

Proof. For $k = 0, 1, 2, \ldots$, it follows from the estimation errors dynamics (17, 18) that

$$\frac{\tilde{f}_k(t)}{k+1} = -d_v \int_0^t \sum_{\gamma=1}^{2} \Theta_\gamma \frac{\sum_{s=0}^{k} \tilde{V}_s(\tau)}{k+1} d\tau + \frac{\int_0^t \dot{f}(\tau)\, d\tau}{k+1}. \tag{20}$$

According to the fact that $\sum_{\gamma=1}^{2} \Theta_\gamma$ is a positive definite matrix and c_v is a positive constant, it yields that the matrix $H = \begin{bmatrix} -\left(C_v \sum_{\gamma=1}^{2} \Theta_\gamma + b_v\right) & I_N \\ -d_v \sum_{\gamma=1}^{2} \Theta_\gamma & 0_N \end{bmatrix}$ is Hurwitz. Since $\dot{f}$ is bounded, then $\tilde{f}$ and $\tilde{V}$ are bounded, which implies that

$$\lim_{k\to\infty} \frac{\tilde{f}_k(t)}{k+1} = 0. \tag{21}$$

Under Assumption 2, it is shown that f is bounded, which yields

$$\lim_{k\to\infty} \frac{\int_0^t \dot{f}(\tau)\, d\tau}{k+1} = \lim_{k\to\infty} \frac{f(t) - f(0)}{k+1} = 0. \tag{22}$$

Together with (20–22), it yields

$$\lim_{k\to\infty} \int_0^t \frac{1}{k+1} \sum_{s=0}^{k} \sum_{\gamma=1}^{2} \Theta_\gamma \tilde{V}_s(\tau)\, d\tau = 0. \tag{23}$$

Then, it is shown that

$$\lim_{k\to\infty} \frac{1}{k+1} \sum_{s=0}^{k} \sum_{\gamma=1}^{2} \Theta_\gamma \tilde{V}_s = 0, \tag{24}$$

which implies that

$$\lim_{k\to\infty} \frac{1}{k+1} \sum_{s=0}^{k} \tilde{V}_s = 0, \tag{25}$$

during which we have used the fact that $\sum_{\gamma=1}^{2} \Theta_\gamma$ is positive definite. From $\tilde{V}_s = V - \hat{V}_s$, it follows that

$$\lim_{k\to\infty} \sum_{s=0}^{k} \frac{1}{k+1} \left(V - \hat{V}_s \right) = 0. \tag{26}$$

Differentiate the equation as

$$\lim_{k\to\infty} \sum_{i=1}^{k} \frac{1}{k+1} \left(\dot{V} - \dot{\hat{V}}_s \right) = 0. \tag{27}$$

According to (11, 15, 16), one has

$$\lim_{k\to\infty} \frac{1}{k+1} \sum_{s=0}^{k} \left(-d_v \sum_{\gamma=1}^{2} \Theta_\gamma \tilde{V}_s - b_v \tilde{V}_s - \tilde{f}_s \right) + f = 0. \tag{28}$$

It further follows that

$$\lim_{k\to\infty} \frac{1}{k+1} \sum_{s=0}^{k} \hat{f}_s = f. \tag{29}$$

$$\lim_{k\to\infty} \frac{1}{k+1} \sum_{s=0}^{k} \tilde{f}_s = 0, \tag{30}$$

which implies that the mean value of $\tilde{f}_i$, $i = 0, 1, 2, \ldots, k$, can converge to zero when $k \to \infty$. □

4 Fault-tolerant controller design

In this section, a fault-tolerant controller is designed to recover the voltage to its reference value by compensating for the influence of unknown fault f_i. According to the conclusion in Theorem 1, we can define the estimation of the signal of fault as $f_{i,k}$, i.e.,

$$f_{i,k} = \frac{1}{k+1} \sum_{s=0}^{k} \hat{f}_{i,s}. \tag{31}$$

Based on $f_{i,k}$, the following fault-tolerant controller V_i^d is designed,

$$\dot{V}_i^d = \dot{V}_i - m_v f_{i,k}. \tag{32}$$

Define $V^d = \left[V_1^d, \ldots, V_N^d \right]^T$. Based on (31), the following dynamic can be achieved

$$\dot{V}^d = \dot{V} - M_v F_k, \tag{33}$$

where $F_k = [f_{1k}, \ldots, f_{Nk}]^T$ and satisfies $F_k = \frac{1}{k+1}\sum_{s=0}^{k} \hat{f}_s$ according to (30).

Theorem 2. *If the condition in Theorem 1 is satisfied and the step k is large enough, then the controller designed based on iterative mean estimation information in (31) can ensure that the voltage restores to the normal value under the influence of faults.* **Proof.** Denote $\bar{V} = V^d - (\sum_{\gamma=1}^{2}\Theta_\gamma)^{-1}\sum_{\gamma=1}^{2}\Theta_\gamma(\mathbf{1}_N \otimes V_r)$. Under Assumption 1, we know that the matrix $\sum_{\gamma=1}^{2}\Theta_\gamma$ is non-singular and positive-definite. It follows from (32) that the derivative of $\bar{V}$ is

$$\begin{aligned}\dot{\bar{V}} &= c_v\left(\sum_{\gamma=1}^{2}\Theta_\gamma(\mathbf{1}_N \otimes V_r) - \sum_{\gamma=1}^{2}\Theta_\gamma V\right) + f - F_k \\ &= -c_v\sum_{\gamma=1}^{2}\Theta_\gamma \bar{V} + f - F_k. \end{aligned} \tag{34}$$

Define $\bar{\Psi} = \Psi^T + \Psi$ and $\Psi = -c_v\sum_{\gamma=1}^{2}\Theta_\gamma$. Then, formula (33) can be rewritten as

$$\dot{\bar{V}} = \Psi\bar{V} + f - F_k. \tag{35}$$

Construct the Lyapunov function as $W = \bar{V}^T\bar{V}$. The time derivative of W along (34) is

$$\dot{W} = \bar{V}^T\bar{\Psi}\bar{V} + 2\bar{V}^T(f - F_k). \tag{36}$$

Using Young's inequality leads to

$$\begin{aligned} 2\bar{V}^T(f - F_k) &\le \bar{V}^T\left(-\frac{1}{2}\bar{\Psi}\right)\bar{V} + (f - F_k)^T(-2\bar{\Psi})(f - F_k) \\ &\le -\frac{1}{2}\bar{V}^T\bar{\Psi}\bar{V} + 2\|\bar{\Psi}\| \cdot \|f - F_k\|^2. \end{aligned} \tag{37}$$

TABLE 1 The AC MG parameters and the fault-tolerant controller parameters.

Lines impedance	$R_1 = R_3 = 0.5\Omega, R_2 = R_4 = 0.6\Omega$
Primary Controller	$m_i = n_i = 2e-4$
Secondary Fault-Tolerant Controller	$c_{vi} = c_{wi} = 3, g_i^y = 3$
Reference Signals	$w_{ref} = 50$ Hz, $v_{ref}^1 = 310$ V, $v_{ref}^2 = 311$ V
Load	$R_L = 2\Omega$

Substituting (36) into (35) yields

$$\dot{W} \le \frac{1}{2}\bar{V}^T\bar{\Psi}\bar{V} + 2\|\bar{\Psi}\| \cdot \|f - F_k\|^2. \tag{38}$$

From $\bar{\Psi} < 0$ and $f - F_k \to 0$ as $k \to \infty$, it shows that $\bar{V} \to 0$ as $k \to \infty$. Thus, it is shown that the voltage belongs to the set $[v_{ref}^1, v_{ref}^2]$. □

Remark 1. *Similarly, the developed can be also used to solve the frequency restore problem of MG under the influence of faults. Specifically, the frequency controller can be designed as follows,*

$$\begin{cases} \dot{\hat{\omega}}_{i,l} = c_\omega\left(\sum\limits_{j\in\mathcal{N}_i}\left(\hat{\omega}_{j,l} - \hat{\omega}_{i,l}\right) + g_i\left(\omega_r^1 - \hat{\omega}_{i,l}\right) + g_i^2\left(\omega_r^2 - \hat{\omega}_{i,l}\right)\right) + b_i\tilde{\omega}_{i,l} + m_w\hat{\dot{f}}_{i,l}, \\ \hat{\dot{f}}_{i,l} = -d_w\left(\sum\limits_{j\in\mathcal{N}_i}\left(\tilde{\omega}_{j,l} - \tilde{\omega}_{i,l}\right) - \sum\limits_{\gamma=1}^{2} g_i\tilde{\omega}_{i,l}\right) + \hat{\dot{f}}_{i,l-1}, \end{cases} \tag{39}$$

where $\hat{\dot{f}}_{i-1} = 0$. *The corresponding conclusions and stability analysis process are similar to that of voltage regulation and they are omitted here.*

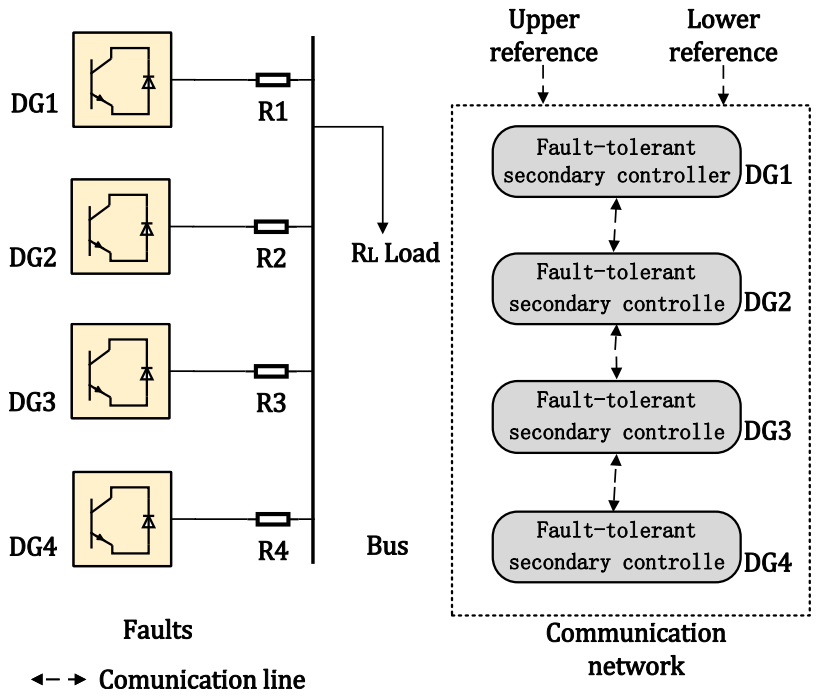


FIGURE 1
Simulation test system.

TABLE 2 The considered faults.

Time(s)	0–5	5–9	9–12	12–16
$f_{v,1}(t)$	0	$5t$	$5t$	$5t$
$f_{v,2}(t)$	0	0	40 sin t	40 sin t
$f_{v,3}(t)$	0	0	0	0
$f_{v,4}(t)$	0	0	0	40
$f_{w,1}(t)$	0	$4t$	$4t$	$4t$
$f_{w,2}(t)$	0	0	30 sin t	30 sin t
$f_{w,3}(t)$	0	0	0	0
$f_{w,4}(t)$	0	0	0	40

TABLE 3 Parameters of the designed fault-tolerant controller.

Fault-tolerant controller (31)	$b_v = 28, d_v = 13, m_v = 7$
Fault-tolerant controller (38)	$b_w = 22, d_w = 16, m_w = 8$
Iterative Number	$k = 6$

5 Simulation study

In this simulation, AC MG systems are considered, which consist of 4 DGs and 2 leaders as shown in Figure 1. The parameters of all DGs are shown in Table 1. To show the effectiveness of our developed fault-tolerant controller, the faults considered are displayed in Table 2, where $f_{v,i}$ and $f_{w,i}$ represent the attack signal on the voltage and frequency controller of DG i, respectively. Then, the following two cases are introduced to show the effectiveness of the developed method.

Case I. (Voltage and frequency regulation with distributed secondary controllers (Zhang et al., 2021) and (Anand et al., 2012)) In this case, the secondary controllers are added at $t = 2$ s. For the secondary controllers (Zhang et al., 2021) and (Anand et al., 2012), the trajectories of voltage and frequency under faults are shown in Figure 2. The simulation result can be summarized from the following stages.

1) Stage 1 (0–2 s): Only primary control is actuated. The frequency ω_i is deviated from its reference value 50 Hz and the voltage V_i is also deviated from its reference range, i.e., [310,311].
2) Stage 2 (2–5 s): The secondary controllers (Zhang et al., 2021) and (Anand et al., 2012) are added at $t = 2$ s, which can restore the frequency ω_i to its reference value and can also control the voltage V_i to [310,311].
3) Stage 3 (5–9 s): The trajectories of frequency ω_1 deviate from its reference value and the trajectories of voltage V_1 cannot stay within its scope under faults signal $f_{v,1}$ and $f_{w,1}$ as shown in Table 2.
4) Stage 4 (9–12 s): The trajectories of frequency ω_1 and ω_2 deviate from its reference value and the trajectories of voltages V_1 and V_2 cannot stay within its scope with the addition of fault signals $f_{v,2}$ and $f_{w,2}$.
5) Stage 5 (12- s): All trajectories of frequency ω_i deviate from its reference value and the trajectories of voltage V_i cannot stay within its scope with the addition of fault signals $f_{v,2}$ and $f_{w,2}$.

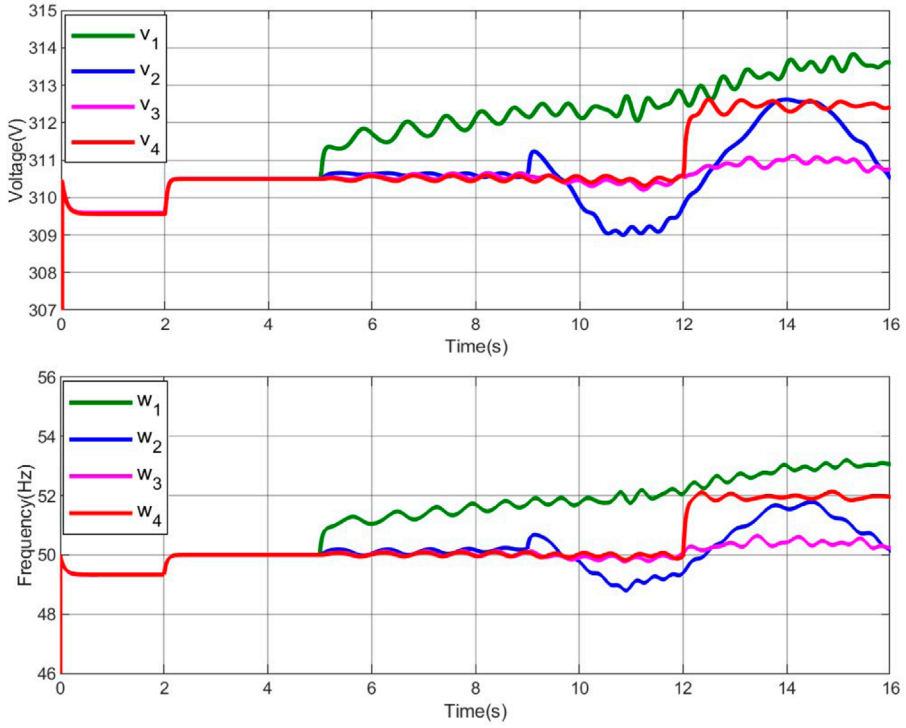


FIGURE 2
Trajectories of voltage and frequency with the distributed secondary controller (8)–(9) under faults.

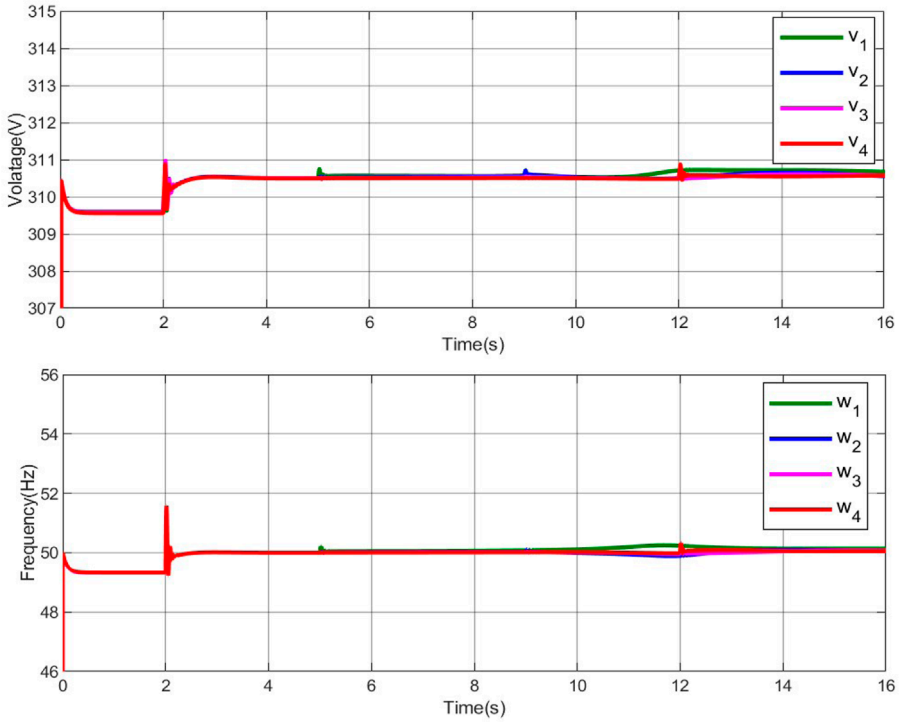


FIGURE 3
Trajectories of voltage and frequency with the secondary fault-tolerant controller (31) and (38) under faults.

In summary, under the distributed secondary controllers (Zhang et al., 2021) and (Anand et al., 2012)), the frequencies and voltages cannot be adjusted to the proper values under the influence of faults.

Case II. (Voltage and frequency regulation using distributed fault-tolerant controllers (31) and (38))

First, the parameters in the fault-tolerant controllers (31) and (38) are displayed in Table 3. Based on the distributed fault-tolerant controllers (31) and (38), the trajectories of voltage and frequency under the influence of faults are shown in Figure 3. Similar to case I, the following stages are shown.

1) Stage 1 (0–5 s): The conclusions are same to those in Stage 1 of Case I.
2) Stage 2 (5–16 s): The distributed fault-tolerant controllers (31) and (38) ensure that the frequency voltage are restored to their range.

Under the influence of the considered faults, the developed fault-tolerant control scheme can achieve the frequency recovery and voltage regulation effectively.

6 Conclusion

The distributed secondary fault-tolerant control problems of voltage and frequency regulation have been studied for smart grids under faults. A wavelet analysis and consensus algorithm-based fault-tolerant control method has been developed to solve the considered problems. In particular, the wavelet analysis technique is firstly introduced to detect the considered faults. Then, distributed secondary fault-tolerant controllers are designed to compensate for the influence of faults by using a newly introduced distributed fault estimate method. It is shown that the designed scheme ensures that both frequency regulation and voltage recovery objectives can be achieved. Finally, AC MGs with four distributed generations are used to validate the effectiveness of the developed fault-tolerant schemes.

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 the paper has been written by YH.

Funding

The work was supported by the basic scientific research project of Department of Education in

2022 of Liaoning Provincial of China under Grant LJKMZ20221993.

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

References

Anand, S., Fernandes, B., and Guerrero, J. (2012). Distributed control to ensure proportional load sharing and improve voltage regulation in low-voltage DC microgrids. *IEEE Trans. Power Electron.* 28 (4), 1900–1913. doi:10.1109/tpel.2012.2215055

Beg, O., Johnson, T., and Davoudi, A. (2017). Detection of false-data injection attacks in cyber-physical DC microgrids. *IEEE Trans. Ind. Inf.* 13 (5), 2693–2703. doi:10.1109/tii.2017.2656905

Bidram, A., and Davoudi, A. (2012). Hierarchical structure of microgrids control system. *IEEE Trans. Smart Grid* 3 (4), 1963–1976. doi:10.1109/tsg.2012.2197425

Chen, Y., Qi, D., Dong, H., Li, C., Li, Z., and Zhang, J. (2020). A FDI attack-resilient distributed secondary control strategy for islanded microgrids. *IEEE Trans. Smart Grid* 12 (3), 1929–1938. doi:10.1109/tsg.2020.3047949

Deng, C., WangWen, Y. C., Xu, Y., and Lin, P. (2020). Distributed resilient control for energy storage systems in cyber–physical microgrids. *IEEE Trans. Ind. Inf.* 17 (2), 1331–1341. doi:10.1109/tii.2020.2981549

Deng, C., Wen, C., Zou, Y., Wang, W., and Li, X. (2022). A hierarchical security control framework of nonlinear CPSs against DoS attacks with application to power sharing of AC microgrids. *IEEE Trans. Cybern.* 52 (6), 5255–5266. doi:10.1109/tcyb.2020.3029045

Ding, L., Han, Q., Ning, N., and Yue, D. (2019). Distributed resilient finite-time secondary control for heterogeneous battery energy storage systems under denial-of-service attacks. *IEEE Trans. Ind. Inf.* 16 (7), 4909–4919. doi:10.1109/tii.2019.2955739

Dou, C., Yue, D., and Guerrero, J. (2016). Multiagent system-based event-triggered hybrid controls for high-security hybrid energy generation systems. *IEEE Trans. Ind. Inf.* 13 (2), 584–594. doi:10.1109/tii.2016.2618754

Dragičević, T., Lu, X., Vasquez, J., and Guerrero, J. (2015). DC microgrids—Part I: A review of control strategies and stabilization techniques. *IEEE Trans. Power Electron.* 31 (7), 1–4891. doi:10.1109/tpel.2015.2478859

Fu, W., Ma, Q., Qin, J., and Kang, Y. (2021). Resilient consensus-based distributed optimization under deception attacks. *Int. J. Robust Nonlinear Control* 31 (6), 1803–1816. doi:10.1002/rnc.5026

Liu, Y., Ning, P., and Reiter, M. (2011). False data injection attacks against state estimation in electric power grids. *ACM Trans. Inf. Syst. Sec.* 14, 1–33. doi:10.1145/1952982.1952995

Lu, X., Guerrero, J., Sun, K., and Vasquez, J. (2013). An improved droop control method for DC microgrids based on low bandwidth communication with DC bus voltage restoration and enhanced current sharing accuracy. *IEEE Trans. Power Electron.* 29 (4), 1800–1812. doi:10.1109/tpel.2013.2266419

Mehrizi-Sani, A., and Iravani, R. (2012). Constrained potential function—based control of microgrids for improved dynamic performance. *IEEE Trans. Smart Grid* 3, 1885–1892. doi:10.1109/tsg.2012.2197424,

Modares, H., Kiumarsi, B., Lewis, F., Ferrese, F., and Davoudi, A. (2019). Resilient and robust synchronization of multiagent systems under attacks on sensors and actuators. *IEEE Trans. Cybern.* 50 (3), 1240–1250. doi:10.1109/tcyb.2019.2903411

Morstyn, T., Hredzak, B., and Agelidis, V. (2016). Control strategies for microgrids with distributed energy storage systems: An overview. *IEEE Trans. Smart Grid* 9 (4), 3652–3666. doi:10.1109/tsg.2016.2637958

Ge, P., Chen, B., and Teng, F., " Event-triggered distributed model predictive control for resilient voltage control of an islanded microgrid," *Int. J. Robust Nonlinear Control*, 31, 2001, 1979, 2000. doi:10.1002/rnc.5238

Sahoo, S., Mishra, S., Peng, J., and Dragičević, T. (2018). A stealth cyber-attack detection strategy for DC microgrids. *IEEE Trans. Power Electron.* 34 (8), 8162–8174. doi:10.1109/tpel.2018.2879886

Sánchez, H., Rotondo, D., Escobet, T., Puig, V., and Quevedo, J. (2019). Bibliographical review on cyber attacks from a control oriented perspective. *Annu. Rev. Control* 48, 103–128. doi:10.1016/j.arcontrol.2019.08.002

Savaghebi, M., Jalilian, A., Vasquez, J., and Guerrero, J. (2012). Secondary control scheme for voltage unbalance compensation in an islanded droop-controlled microgrid. *IEEE Trans. Smart Grid* 3 (2), 797–807. doi:10.1109/tsg.2011.2181432

Valenciaga, F., and Puleston, P. (2005). Supervisor control for a stand-alone hybrid generation system using wind and photovoltaic energy. *IEEE Trans. Energy Convers.* 20 (2), 398–405. doi:10.1109/tec.2005.845524

Wang, Z., and Passino, K. (2016). Stable reactive power balancing strategies of grid-connected photovoltaic inverter network. *Int. J. Robust Nonlinear Control* 26 (9), 2023–2046. doi:10.1002/rnc.3408

Zhang, H., Yue, D., Dou, C., Xie, X., Li, K., and Hancke, G. (2021). Resilient optimal defensive strategy of tSK fuzzy-model-based microgrids' system via a novel reinforcement learning approach. *IEEE Trans. Neural Netw. Learn. Syst.* doi:10.1109/TNNLS.2021.3105668

Zhao, J., Mili, L., and Wang, M. (2018). A generalized false data injection attacks against power system nonlinear state estimator and countermeasures. *IEEE Trans. Power Syst.* 33 (5), 4868–4877. doi:10.1109/tpwrs.2018.2794468

Zhou, Q., Shahidehpour, M., Alabdulwahab, A., and Abusorrah, A. (2020). A cyber-attack resilient distributed control strategy in islanded microgrids. *IEEE Trans. Smart Grid* 11 (5), 3690–3701. doi:10.1109/tsg.2020.2979160

Zuo, S., Beg, O., Lewis, F., and Davoudi, A. (2020). Resilient networked AC microgrids under unbounded cyber attacks. *IEEE Trans. Smart Grid* 11 (5), 3785–3794. doi:10.1109/tsg.2020.2984266

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 13 April 2023
DOI 10.3389/fenrg.2023.1119461

OPEN ACCESS

EDITED BY
Chao Deng,
Nanjing University of Posts and Telecommunications, China

REVIEWED BY
Yao Weitao,
Nanyang Technological University, Singapore
Zhanjie Li,
Nanjing University of Posts and Telecommunications, China

*CORRESPONDENCE
Yanping Wang,
wangyp@dlpu.edu.cn

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 08 December 2022
ACCEPTED 29 March 2023
PUBLISHED 13 April 2023

CITATION
Shi S, Wang Y and Jin J (2023), Multi-agent-based control strategy for centerless energy management in microgrid clusters.
Front. Energy Res. 11:1119461.
doi: 10.3389/fenrg.2023.1119461

Multi-agent-based control strategy for centerless energy management in microgrid clusters

Shuang Shi, Yanping Wang* and Jiyu Jin

School of Information Science and Engineering, Dalian Polytechnic University, Dalian, China

Interconnecting microgrids with similar geographical environment and related characteristics electrically and communicatively, this constitutes a microgrid cluster, which is a higher-level distributed power structure and an effective way to improve the utilization rate of distributed energy and local absorption level. However, the stable operation of microgrid clusters is determined by the cluster structure and its control strategy, which is also the focus and difficulty of current research. To address this challenge, this paper proposes a ring-based multi-agent microgrid cluster energy management strategy, which realizes the centerless coordinated autonomous operation of microgrid clusters with high stability. In addition, based on the multi-agent control strategy, this paper designs a variety of control strategies that can be switched autonomously for different control objectives of the microgrid cluster to realize smooth grid connection and off-grid of the microgrid cluster. Finally, based on Matlab platform simulation, the experimental results show that the control structure and energy management strategy proposed in this paper can realize the centerless coordination and autonomy of the microgrid cluster on the basis of stable operation, minimize the energy interaction between the cluster and the distribution network, and greatly improve the utilization rate and local absorption level of distributed energy.

KEYWORDS

microgrid clustering, energy management, multi-agent, coordinated autonomy, centerless

1 Introduction

Microgrid (MG) is the basic component unit of microgrid cluster, so the stability, efficiency and control flexibility of microgrid operation are the basis for realizing microgrid cluster operation. At present, the development of microgrids has been relatively mature, and people have done a lot of work on microgrid operation, control, and protection Joos et al., 2000; Chiradeja and Ramakumar, 2004; Quezada et al., 2006; Hegazy et al., 2003; Hoff and Shugar., 1995; Greatbanks et al., 2003. Which has accelerated the development of microgrid technology. However, the large-scale access of microgrids leads to a large number of bidirectional currents in the distribution network with scattered access points, which brings new problems to voltage regulation, protection coordination and energy optimization Patel et al., 2008. Exploring and developing active distribution technologies for distribution networks is an effective way to solve this problem. On the other hand, how to effectively integrate a large number of microgrids to make them operate autonomously and reduce the impact on the distribution network is the hope and requirement of the distribution network for microgrids and one of the goals of microgrid clustering operation

Yao et al., 2021. Microgrid cluster (MGC) is a new structure of distributed power supply in recent years, and relatively little research has been done on MGC at home and abroad, but as a further development trend of coordinated operation of multiple microgrids, its superiority in distributed power integration, flexible control mode, energy utilization efficiency, and support to the distribution microgrid has led to the research on MGC at home and abroad Yao et al., 2023. More and more attention has been paid to the research of MGC both at home and abroad. For example, Qingzhu Wan et al. from the School of Electrical and Control Engineering of Northern Polytechnic University wrote a paper to explain the concept, structure and composition of microgrid clusters Qingzhu and Xue, 2017. The general design of microgrid cluster operation was studied by Jiangmen Power Supply Bureau of Guangdong Power Grid Company and the School of Electronic Information and Electrical Engineering of Shanghai Jiaotong University, which compared the characteristics of microgrid cluster operation with related methods such as coordinated operation of multiple microgrids and single microgrid operation, and discussed the control methods of microgrid cluster operation Cai Dehua et al., 2015. Zhou Xiaoqian et al. from the School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, proposed local centralized optimization and global optimization methods. A framework of local centralized optimization and global distributed iterative distributed optimization based on consistency and model predictive control theory was proposed, which well solves the large-scale MGC distributed optimization solution problem Zhou et al., 2018a. In addition, the Microgrid Laboratory of Aalborg University, Denmark, has conducted research on integrated modeling of microgrid clusters, multi-timescale energy management of multilayer microgrid clusters, distributed coordinated control and optimization based on coordinated operation of multiple microgrids, and virtual and semi-physical simulations to verify the correctness and effectiveness of the proposed scheme. The stable operation of microgrid clusters depends on the correct design and selection of control strategies, and currently, the main control methods are centralized, decentralized and distributed. Centralized control cannot meet the robustness and flexibility requirements of microgrid clusters due to its own characteristics Karimi et al., 2017; Wei and Xianwei., 2011; Minchala-Avila et al., 2016. Decentralized control achieves stable control of the unit itself through local control methods, and power balancing within and between microgrids can be achieved through reasonably designed control algorithms Nutkani et al., 2016; Divshali et al., 2012. Subsequently, the shortcomings of traditional sag control have been optimized and corresponding control methods for autonomous operation of microgrids have been proposed, providing directions and ideas for the control of microgrid clusters Jin et al., 2010; M. D. Cook et al., 2017; Dou et al., 2015; Gu et al., 2014. However, the decentralized control systems of microgrids are "fragmented", which makes it difficult for microgrid clusters to support large microgrids. In recent years, the application of distributed control theory in power systems has gradually gained attention, and the resulting distributed control of microgrids has been gradually developed. The distributed control approach not only has high flexibility, immunity and ability to resist cyber security threats, but also can provide strong technical support to the microgrid, so distributed control is particularly suitable for microgrid clusters Han et al., 2019. Under the distributed control architecture, various optimal control methods have been proposed for different control objects and optimization objectives, such as autonomous coordinated control strategies for clusters with multiple AC-multiple DC microgrids Zhou et al. (2018b), hierarchical cluster control schemes for islanded DC microgrid clusters Luo et al., 2018, and distributed hierarchical cooperative control strategies for islanded microgrid clusters with intermittent communication Li et al., 2019, and simulation verified the effectiveness of the control strategy. During the research of microgrid cluster control strategies, the time scale of cluster energy management, dynamic characteristics of cluster operation, communication delays, the impact of microgrid plugging and unplugging on communication links, line loss and economy, and topology optimization have also been discussed and solutions have been proposed Xing et al., 2017; Z. Zhao et al., 2017; Han et al., 2018; Han et al., 2017; Zhou et al., 2017. However, the existing cluster topologies and their distributed algorithms are still largely based on the instructions issued by the high-level controller or the dominant node, and do not operate autonomously without a control center in the real sense.

The shortage of energy resources and ecological degradation have led to the development of photovoltaic and wind power generation, but a single microgrid is no longer able to meet the needs of users. Therefore, an effective solution to directly dispatch and manage a large number of distributed generators in the distribution network is to centralize distributed power sources into a microgrid cluster. A microgrid cluster is formed by interconnecting microgrids with similar geography and related characteristics electrically and communicatively and interacting with the distribution microgrid through a unique access point, which is a higher-level distributed power architecture. In this paper, a ring-shaped distributed microgrid cluster structure is proposed. In this structure, distributed control systems with integrated communication modules are distributed in each microgrid unit in a ring link, and the microgrid only interacts with neighboring microgrids to obtain current status information and user information of other microgrids in the cluster. When the cluster is operating, each distributed control system (agent) can simultaneously solve the next operation plan of the cluster, in addition to maintaining the stable operation within the microgrid. Therefore, each distributed control system is both the control center of each microgrid unit and the control center at the cluster level, completely eliminating the need for a centralized dispatch center and dominant node. This fully distributed control strategy can greatly improve the stability and robustness of cluster operation and finally realize autonomous operation without central coordination, which has important theoretical research significance and application promotion value. Through simulations and experiments, we demonstrate that our approach can improve the reliability, stability, and efficiency of microgrid operation while reducing the dependence on the main power grid. Our approach utilizes real-time dynamic compensation control, which can adapt to changing conditions in the microgrid more quickly. We also show that our approach can effectively balance the power generation and demand of the microgrid by integrating a smart agent-based control system and using energy storage devices such as batteries and supercapacitors. Our results suggest that our approach could be a promising solution for the large-scale deployment of renewable energy microgrids in the future.

2 Materials and methods

As a more complex form of distributed power structure, microgrid clusters will present the following new features: each micro power source within the cluster can independently meet the power demand of each microgrid load in the islanded state; when there is a microgrid that cannot meet the load power demand, the cluster networking state can be carried out to share the surplus power of the respective microgrid and the power in energy storage through energy interaction. When the overall power of the cluster is insufficient or the energy storage equipment is surplus, flexible automatic load shedding can be carried out at the level of meeting the self-sufficiency rate of the cluster, or feedback or request for power supply to the distribution network through the public coupling point (PCC). The above features greatly enhance the flexibility of cluster operation. The increased degree of autonomy of microgrid clusters will allow more active control, since the topology of the cluster will be determined to some extent by the interactions between the microgrids. We present the cluster operation of microgrids and its superiority. The cluster structure and its operational control strategy determine the stability of the system and the utilization of distributed energy, while the realization of the autonomous operation goal determines the magnitude of the support role for the distribution microgrid. In the existing microgrid secondary control, the main control objectives are voltage and frequency recovery, SOC equalization, etc., Ma et al., 2022; Deng et al., 2022. In the control proposed in this paper, the control objective of the multi-agent system is to adjust the control mode and static switching state of the energy storage unit according to the operating status of each microgrid to ensure that all microgrids can operate stably. For example, when a microgrid cannot operate autonomously because the battery SOC reaches the constraint condition, another microgrid is selected to interconnect with this microgrid through the two-layer multi-agent system and the energy storage units of other microgrids are used to maintain the power balance and ensure stable operation.

2.1 Microgrid cluster overview

In this paper, a ring distributed autonomous system structure and its control strategy are proposed, i.e., a ring structure is used to connect microgrid clusters to the microgrid, a solution for smooth microgrid connection of microgrid clusters and coordinated autonomy of each microgrid is proposed, and a multi-agent system (MAS) based energy management method for microgrid clusters is proposed. The schematic diagram of the microgrid cluster model is shown in Figure 1.

The multi-agent system used in this paper is a distributed control system composed of agents in which the agents communicate with each other to coordinate and control the microgrid cluster and the state of each microgrid according to certain rules, using a distributed approach to problem solving. Each agent integrates a communication module and forms a ring link in which it is connected to the microgrid and in which distributed control algorithms are embedded. Typically, each agent is interconnected with its neighboring agents and interacts with information through a control policy developed within the agent. The proposed cyclic agent architecture allows neighboring agents to communicate with each other, and the agent is used for energy management of the microgrid cluster, and the cyclic structure of the agent is used to reasonably allocate energy storage and achieve dynamic energy compensation among sub-microgrids. Second, starting from the actual microgrid connection, the control strategy of the sub-microgrids is converted to include pre-synchronization control to enable smooth microgrid connection of the microgrid clusters at the PCC.

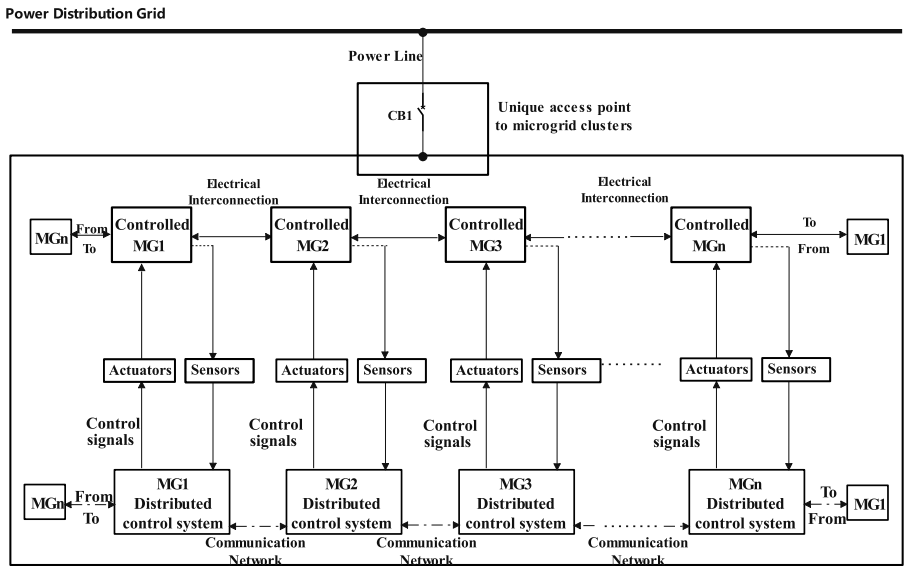


FIGURE 1
Schematic diagram of microgrid cluster model.

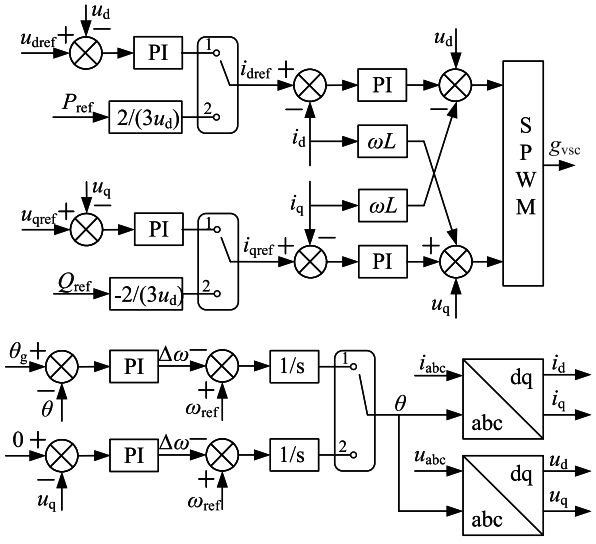


FIGURE 2
The control structure diagram of microgrid master control unit.

2.2 Control strategy of agent system in a single microgrid

Before studying the control of microgrid clusters, the control of single microgrid should be considered first, then study the control of single microgrid using agent system, which is important to study the stability of microgrid clusters.

In this paper, the master-slave control strategy is used for a single microgrid. As power terminals, the PV and wind generation units are mainly controlled with maximum power tracking to maximize the utilization of wind and solar energy. Similarly, in grid-connected operation mode, energy storage also acts as a power terminal, receiving power commands from the upper-level dispatch system for constant power control (PQ) charging and discharging. However, as a balancing terminal, the energy storage device uses constant voltage and constant frequency (V/f) control to maintain stable AC bus voltage and frequency under island operation mode conditions. The operation mode and static switching of each microgrid is controlled by a multi-agent system. The control structure diagram of the microgrid master-slave inverter is shown in Figure 2.

As shown in Figure 2, u_{dref} and u_{qref} are reference values of the d-axis and q-axis PCC voltage, respectively; i_{dref} and i_{qref} are reference values of the d-axis and q-axis grid-connected current, respectively; u_{abc} is three-phase PCC voltage; u_{abc} is three-phase grid-connected current; ω_{ref} is rated reference value of fundamental angular frequency; θ is phase information of PCC voltage; θ_g is phase information of grid-connected voltage; $\triangle\omega$ is variation of fundamental angular frequency.

As shown in Figure 2, in grid-connected operation mode, the control objective of the energy storage system is to track the power command, so PQ control is used. d-axis and q-axis current references are generated from the power command and d-axis voltage, and phase information is obtained from the phase-locked loop (PLL). In islanded operation, the microgrid loses the support of grid voltage and frequency, and if constant power control is still used, when the inverter output power does not match the load power. It will cause fluctuation of system voltage and frequency, resulting in system instability. Therefore, in the islanding mode, the main inverter of the microgrid needs to assume the role of voltage and frequency regulation and adjust its own output power according to the load demand. In the islanding mode of operation, the current reference values for the d and q axes are generated by the PI regulator of the outer voltage loop. In the inner current loop, decoupling and voltage feedforward control are also considered. In addition, pre-synchronization control needs to be considered when the microgrid is switched from islanding mode to grid-connected mode. The purpose is to achieve phase alignment on both sides of the static switch before switching, thus reducing the voltage/current impact caused during the switching process.

The topology structure of voltage source grid connected inverter is shown in Figure 3. In dq two-phase synchronous rotating reference frame, the mathematical model of grid connected inverter can be expressed as follows

$$L\frac{di_d}{dt} + Ri_d - \omega Li_q = v_d - u_d \tag{1}$$

$$L\frac{di_q}{dt} + Ri_q + \omega Li_d = v_q - u_q \tag{2}$$

Where L is filter inductance; R is the parasitic resistance of the filter inductance; ω is fundamental angular frequency; i_d and i_q

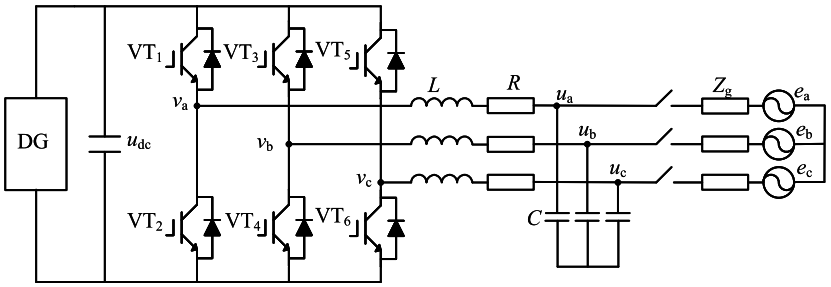


FIGURE 3
The topology structure of voltage source grid connected inverter.

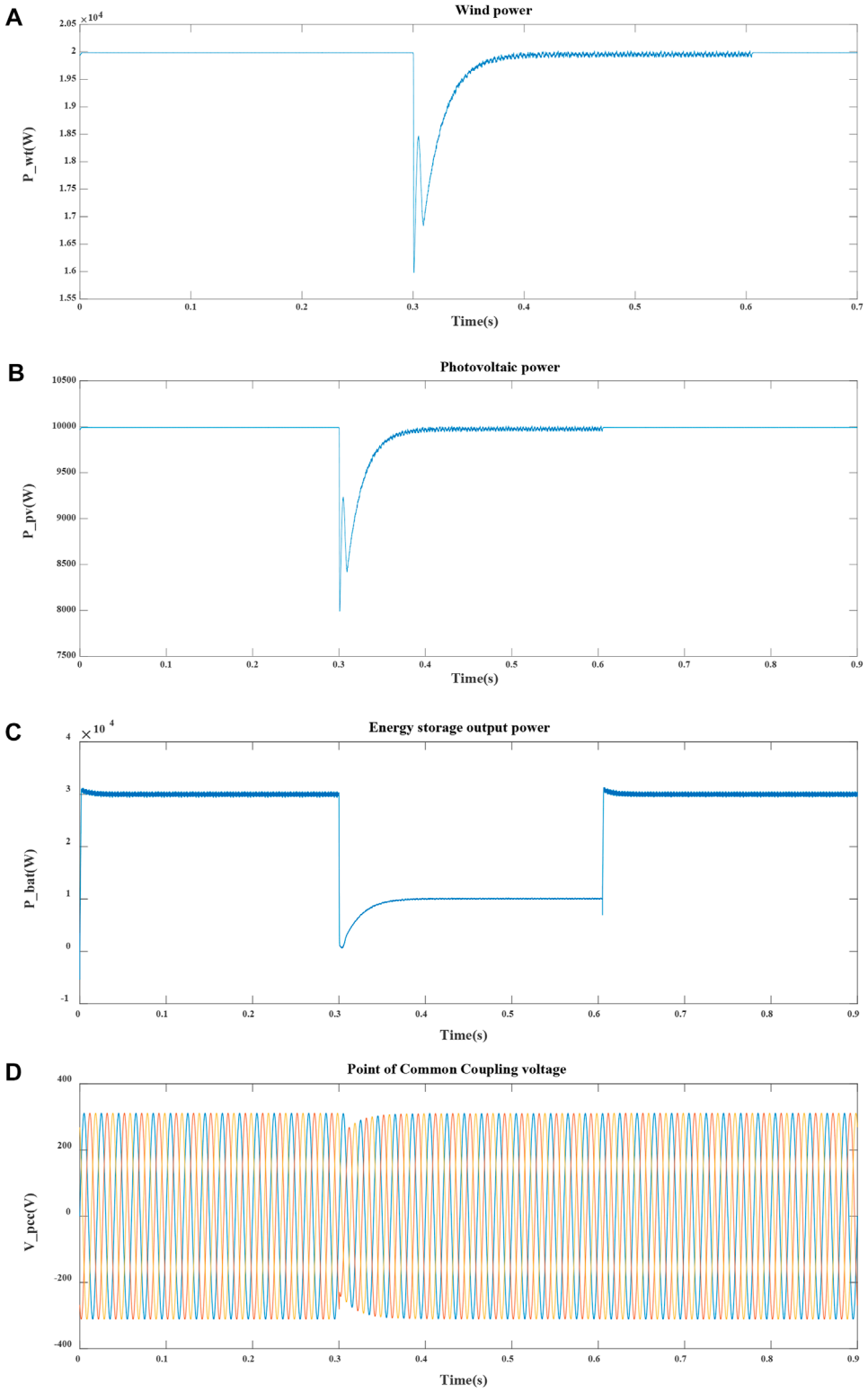


FIGURE 4
Single microgrid simulation results.

are the d-axis and q-axis components of grid connected current; u_d and u_q are the d-axis and q-axis components of point of common coupling (PCC) voltage; v_d and v_q are the d-axis and q-axis components of converter terminal voltage. The active power and reactive power can be expressed as follows

$$P = \frac{3}{2} i_{\mathrm{d}} u_{\mathrm{d}} \tag{3}$$

$$Q = -\frac{3}{2} i_{\mathrm{q}} u_{\mathrm{d}} \tag{4}$$

In grid-connected operation mode, PCC voltage is considered constant, so the current reference value can be expressed as follows

$$i_{\mathrm{dref}} = \frac{2P_{\mathrm{ref}}}{3u_{\mathrm{d}}} \tag{5}$$

$$i_{\mathrm{qref}} = -\frac{2Q_{\mathrm{ref}}}{3u_{\mathrm{d}}} \tag{6}$$

Where Pref and Qref are reference values of active and reactive power, respectively. In island operation mode, voltage closed-loop control is used to maintain constant voltage and frequency of AC bus, so the current reference value can be expressed as follows

$$i_{\mathrm{dref}} = \left(u_{\mathrm{dref}} - u_{\mathrm{d}}\right)\left(k_{\mathrm{vp}} + \frac{k_{\mathrm{vi}}}{s}\right) \tag{7}$$

$$i_{\mathrm{qref}} = \left(u_{\mathrm{qref}} - u_{\mathrm{q}}\right)\left(k_{\mathrm{vp}} + \frac{k_{\mathrm{vi}}}{s}\right) \tag{8}$$

where u_{dref} and u_{qref} are reference values of the d-axis and q-axis voltage, respectively; k_{vp} and k_{vi} are proportional and integral coefficients of PI regulator, respectively.

The single microgrid control strategy is verified by Matlab platform and the simulation results are shown in Figure 4.

The results show that: the microgrid operates grid-connected at 0~0.3*s*, when the wind turbine power, PV power and energy storage power are 20 KW, 10 KW and 30 KW respectively; When a microgrid fault is detected at 0.3*s*, the microgrid switches from grid-connected operation to islanded operation. The agent system switches the control strategy from PQ control to V/f control, at which time the energy storage alone bears the power difference between the scenic generation and the load. Therefore, the energy storage discharges 10 KW.When the microgrid switches from islanded operation to grid operation at 0.6*s*. The agent system switches the control strategy from V/f control to PQ control and starts pre-synchronous control. The simulation results in Figure 4 show that the PCC has a smooth voltage transition, which proves that the single network can operate stably and the agent system can complete the switchover from grid to grid independently and can be connected to the grid smoothly without the need for a cluster coordination control strategy.

2.3 Microgrid cluster ring control structure

Several microgrids are electrically interconnected according to geography or type to form microgrid clusters. The transmission of microgrid status information and control signals depends on an efficient and reliable communication network, while effective prediction of uncertainty in distributed power sources and loads is a prerequisite for inter-microgrid energy dispatch. First, individual microgrids are established, and then all microgrids of the same type are electrically and communicatively interconnected according to the microgrid type to form a microgrid cluster. When the cluster is operating, each distributed control system synchronizes to solve the next operation plan of the cluster, in addition to maintaining the stable operation within the microgrid. Therefore, each distributed control system (agent) is both the control center of each microgrid unit and the control center at the cluster level, completely eliminating the need for a centralized dispatch center and a dominant node. Based on the idea of multi-agent controller, an agent-based control strategy for microgrid cluster ring network is proposed. This control strategy mainly consists of the communication network of agents and multiple microgrids, in which agents communicate with each other to form a communication network. The microgrid consists of distributed photovoltaic power generation, wind power generation, energy storage modules and loads. The overall structure is shown in Figure 5. The line between two agents in the Figure indicates the communication interconnection, and the direction of arrow indicates the direction of information transfer; The dotted line between agent and MG indicates their connection, mainly the information upload of MG and the command sending of agent system; The solid line between MG and MG indicates the electrical interconnection, which can carry out energy transfer, and the arrow indicates its direction.Each microgrid is equipped with an intelligent body (agent). It is responsible for collecting the current state of

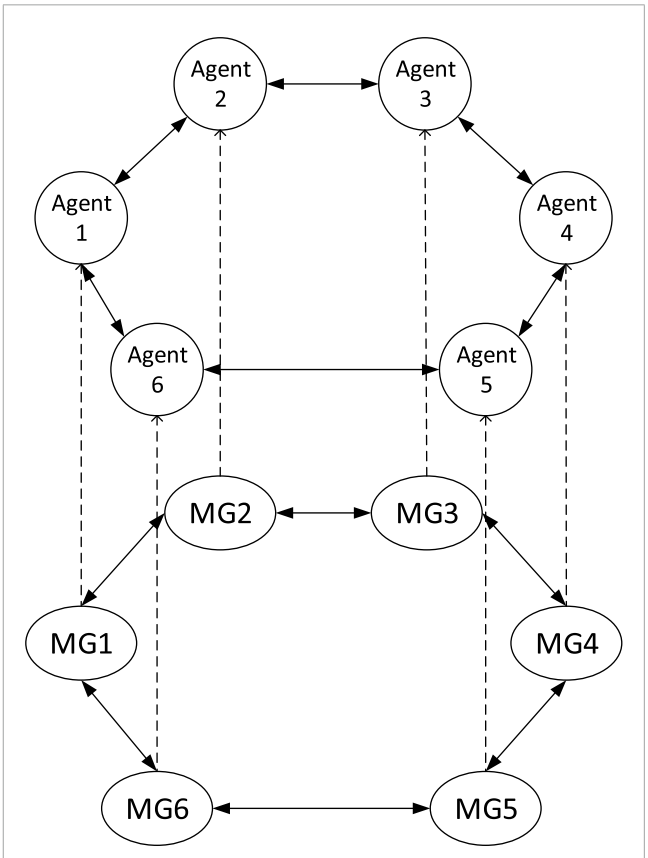


FIGURE 5
Microgrid cluster control model for ring networks.

charge (SoC) and calculating the difference between the airflow and the power required by the user. It is also responsible for exchanging the current power error values with the connected agents and comparing the upper and lower limits of the SoC in order to issue the best control commands according to the control rules. Each agent interacts only with connected agents, and this ring network structure minimizes communication delays.

Remark 1, This ring distributed architecture proposed in this paper achieves the goal of centerless energy coordination and autonomy by eliminating the influence of the central controller and the dominant agent at the cluster level, and reduces the impact of communication failures on the system. This distributed ring structure greatly improves the robustness of microgrid clusters.

2.4 Distributed energy management control strategy for microgrid clusters

a) Reasonable design of control strategy. The stability of operation is built on the basis of intra-cluster power balance, so the power shortage and excess of each microgrid needs to be reasonably matched and scheduled. Under the fully distributed control architecture, we propose a real-time dynamic compensation control strategy: The control system calculates the power difference of the microgrid in the next operation step based on the power generation measurement and customer power consumption information, and performs inter-grid power scheduling on the basis of ensuring its own stable operation. In addition, the control objective of the autonomous operation of this project requires that when the total power difference of the cluster is less than zero, the decision to connect to the grid is made according to the self-sufficiency rate set by the cluster.

b) For the control strategy adopted in this project, autonomous operation without central coordination requires that when the distributed control system is given global information, it is able to synchronize the real-time dynamic compensation schemes of the clusters and then, after the schemes are completed, verify their consistency. Therefore, in addition to synchronizing the clocks, the start and end times of the solving process should be set in the system. If the solution results are inconsistent, the above solution and verification process should be repeated until convergence, and the number of repetitions should be adjusted according to the cluster operation parameters. Under the fully distributed control architecture, the above control strategy and synchronous solving algorithm will eliminate the need for dominant nodes and central controllers to realize the centerless autonomous operation of the microgrid cluster.

In order to maintain the energy balance of microgrid clusters for coordination and autonomy, it is necessary to eliminate the prediction errors. In this paper, a real-time dynamic compensation

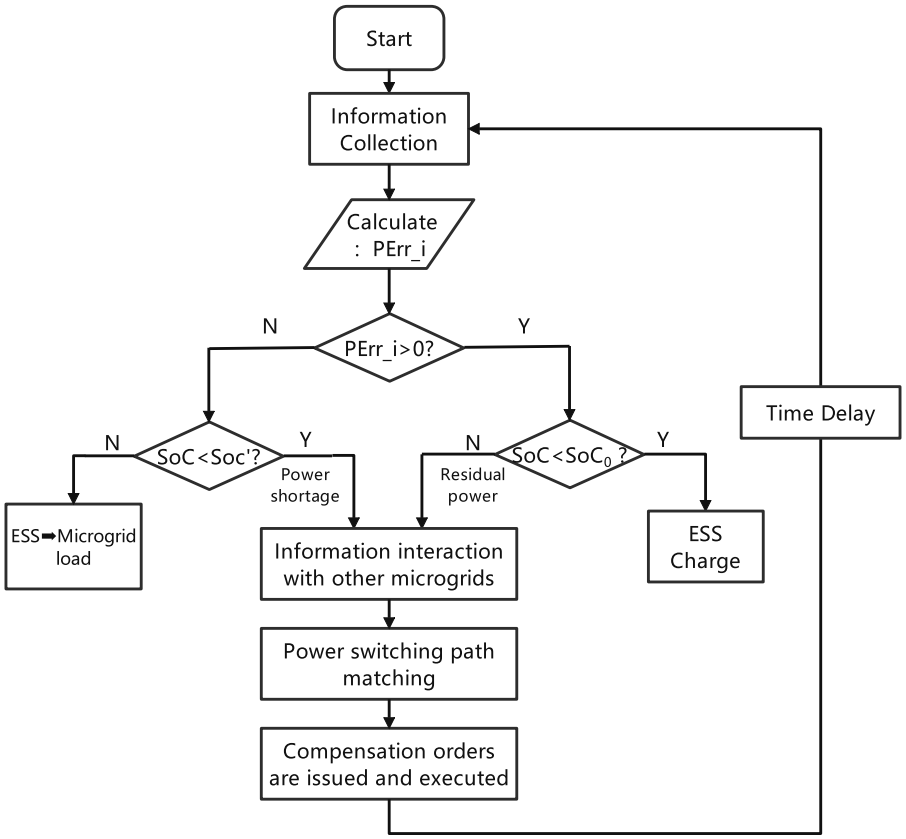


FIGURE 6
Real-time dynamic compensation control strategy flow chart.

control strategy is proposed for the error between the actual data of each microgrid unit and the customer demand data. The power error is represented as follows

$$P_{error_i} = P_{wt\_real_i} + P_{pv\_real_i} - P_L \tag{9}$$

where, i represents the i microgrid, $P_{pv\_real_i}$ and $P_{wt\_real_i}$ are the measured power of Photovoltaic and wind, respectively, and P_L is the demand power of the user, and P_{error_i} is the error power. The specific algorithm flow chart is shown in Figure 6.

The detailed steps of the real-time dynamic compensation control strategy are as follows.

Step 1: Calculate all the P_{error_i} ($P_{error_1} \sim P_{error_6}$), and get 6 output states ($P_{state1} \sim P_{state6}$) according to the hysteresis loop comparator judgment, at this time, when P_{error_i} is greater than 0, make $P_{state1} = 1$, which proves that the actual power of scenery power generation is greater than the user demand power, then the remaining power is stored in energy storage system (ESS) or interact with other microgrid energy, at this time, it represents that the energy storage is charging. When P_{error_i} is less than 0, that is, the actual power of scenic power generation is less than the user's demand power, so $P_{state1} = 0$, which means that the storage energy is being discharged or the cluster is interacting with the electric energy.

Step 2: Based on the collected SoC value, determine whether the energy storage has reached the constrained state (SoC' is the lower limit, SoC_0 is the upper limit), and output the state ($SW_{BES1} \sim SW_{BES6}$, $MG1_{state} \sim MG1_{state}$) as 1 if the constraint is reached, and 0 if the constraint is not reached.

Step 3: The agent system calculates the states of $P_{state1} \sim P_{state6}$ in step 1 and $SW_{BES1} \sim SW_{BES6}$ in step 2 to obtain the state ordering of the microgrid sequence numbers and SoC values that require energy mutualization.

Step 4: Energy exchange path matching is performed based on the sequence number and the ranking of SoC states of the microgrids that need energy interaction obtained in step 3. Microgrids that do not participate in energy interaction continue to operate in isolation. If the cluster cannot operate autonomously, they enter the grid smoothly through PCC.

Remark 2, in order to enable the cluster to meet the load demand value through energy management, to achieve avoid the ESS not to meet the user demand when the actual output power of Photovoltaic and wind is small. The real-time dynamic compensation control strategy is proposed in this paper. It is worth stating that the dynamic compensation scheme includes two aspects. 1) Energy scheduling between distributed power sources, energy storage devices and loads within the microgrid. 2) Energy scheduling between microgrids. In the energy scheduling of microgrid clusters, the power deficit and surplus of each microgrid should be correctly matched.

3 Simulation results at each working condition

The following is an example of six microgrids, MG1, MG2, MG3, MG4, MG5, and MG6, whose microgrid structures all

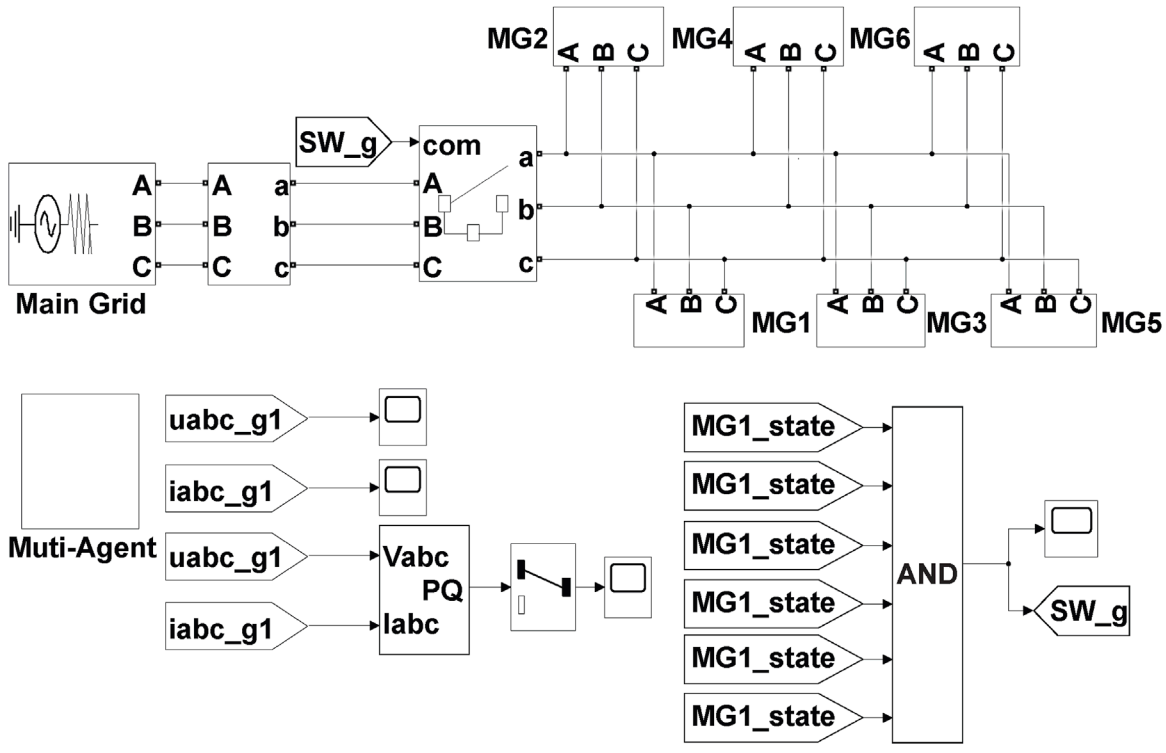


FIGURE 7
Microgrid cluster overall simulation diagram.

choose PV power generation, wind turbines, load and energy storage systems, and the busbar chooses AC busbar. When the ring-type microgrid cluster control strategy is adopted, the power coordination autonomy of the cluster can be realized, and the overall simulation diagram is shown in Figure 7. The following is the simulation model built with Matlab, and the results show the power generation, SoC value, energy storage output power, and grid connection current of each microgrid. It proves the effectiveness of the control strategy in this paper.

3.1 Cluster operation simulation results (condition 1: Excess power)

The simulation time is set to 2.3s, and the initial SoC values of all six microgrids are close to the upper limit, which is specified to be 80%, and the SoC value of MG5 is the smallest, and the P_{error_i} of MG5 is less than 0. The cluster operation is observed. Table 1 shows the parameters of each microgrid. Figures 8–10 shows the results of the cluster simulation.

From the above simulation results, under the initial conditions, the energy storage SoC values of MG1-MG6 are close to the upper limit, when the power excess comes to verify whether the cluster can share by itself. From the data in the table, we can see that only the P_{error_i} of MG5 is less than 0, and the energy storage SoC value of MG5 is at a high level, so the P_{error_i} of MG5 is borne by the energy storage of MG5 alone; The P_{error_i} of the other five microgrids is greater than 0, showing a state of excess power, and their own energy storage is at a high level. This condition is used to verify whether the cluster can share the excess power by itself. First, the energy storage of MG1 reaches the upper limit first. At this time, by detecting the microgrid that meets the transmission condition as MG5, the grid connection switch of MG1 and MG5 is closed, and the excess

TABLE 1 Margin specifications.

Unit/power	MG1	MG2	MG3	MG4	MG5	MG6
Photovoltaic units	10 KW	20 KW	20 KW	15 KW	10 KW	20 KW
Wind power unit	20 KW	20 KW	20 KW	20 KW	20 KW	20 KW
Energy storage unit	40 KWh	40 KWh	40 KWh	40 KWh	40 KWh	40 KWh
Load unit	20 KW	20 KW	20 KW	20 KW	40 KW	20 KW
Error power	10 KW	20 KW	20 KW	15 KW	−10 KW	20 KW
Busbar voltage	380 V	380 V	380 V	380 V	380 V	380 V

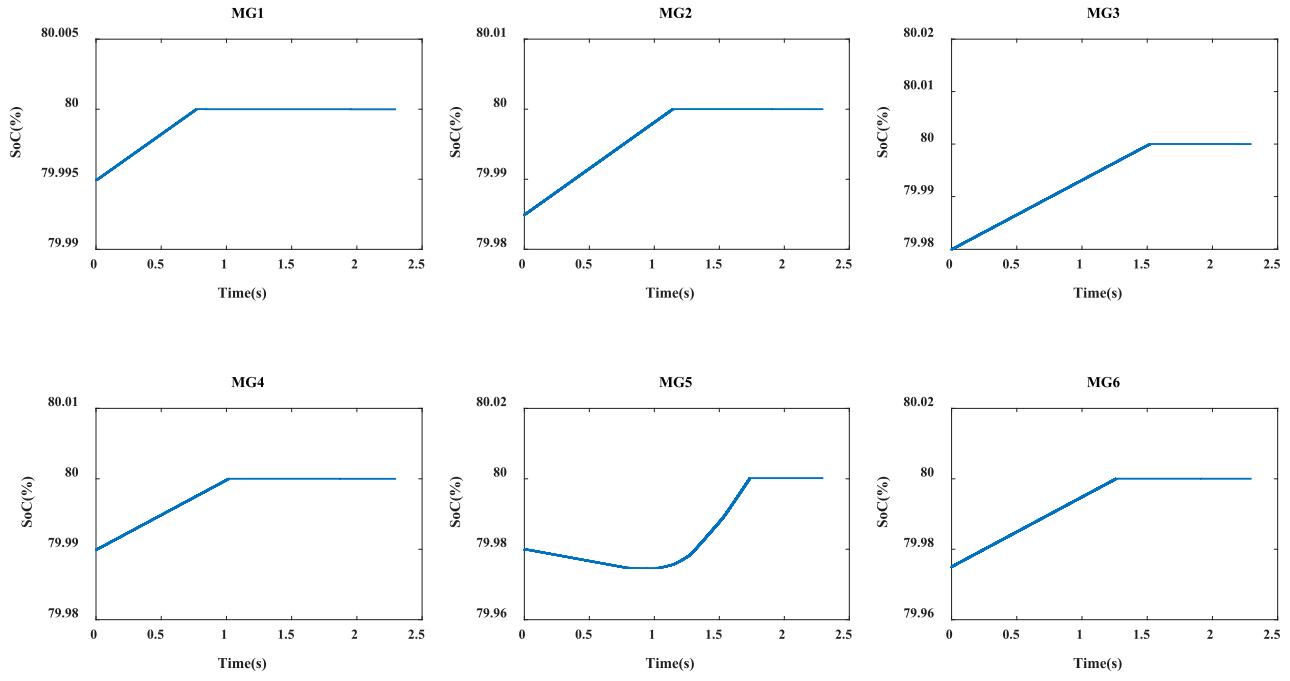


FIGURE 8
Soc value of each microgrid.

power of MG1 is borne by MG5; Then, the energy storage of MG2, MG3, MG4 and MG6 reach the upper limit in turn, and their grid connection switch is closed in turn, and the excess power is borne by MG5. The excess power is borne by MG5; Finally, the energy storage SoC value of MG5 also reaches the upper limit, the grid-connected main switch (PCC) is closed, the excess power is balanced by the distribution network, and the energy storage is all on standby. The simulation results in Figures 8–10 prove the autonomy and high robustness of the energy management control strategy proposed in this paper under the condition of cluster overpower. This also proves that the control strategy proposed in this paper can maintain the stability of the system under the condition of excess power generation or relatively low user load, which has good practical significance.

3.2 Cluster operation simulation results (case 2: Power shortage)

The initial SoC values for all six microgrids were close to the lower limit, specified as 80%, and the maximum SoC value was observed for MG2, with an P_{error_i} greater than 0 for clustered operation. Table 2 shows the parameters of each microgrid. The energy storage SoC values for MG1-MG6 are all close to the lower limit under initial conditions to verify that the clusters can operate autonomously and connect to the grid stably without autonomy in case of power shortage. Table 2 shows the parameters of each microgrid. Figures 11–13 show the cluster operation results.

From the above simulation results, only the P_{error_i} of MG2 is greater than 0, and the energy storage SoC value of MG2 is at a

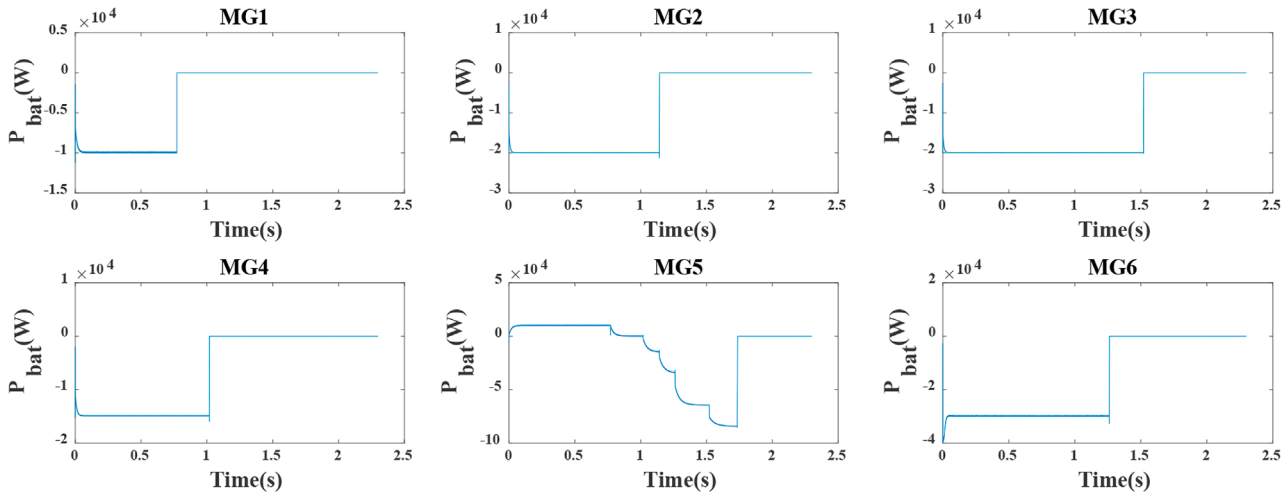


FIGURE 9
Energy storage output power of each microgrid.

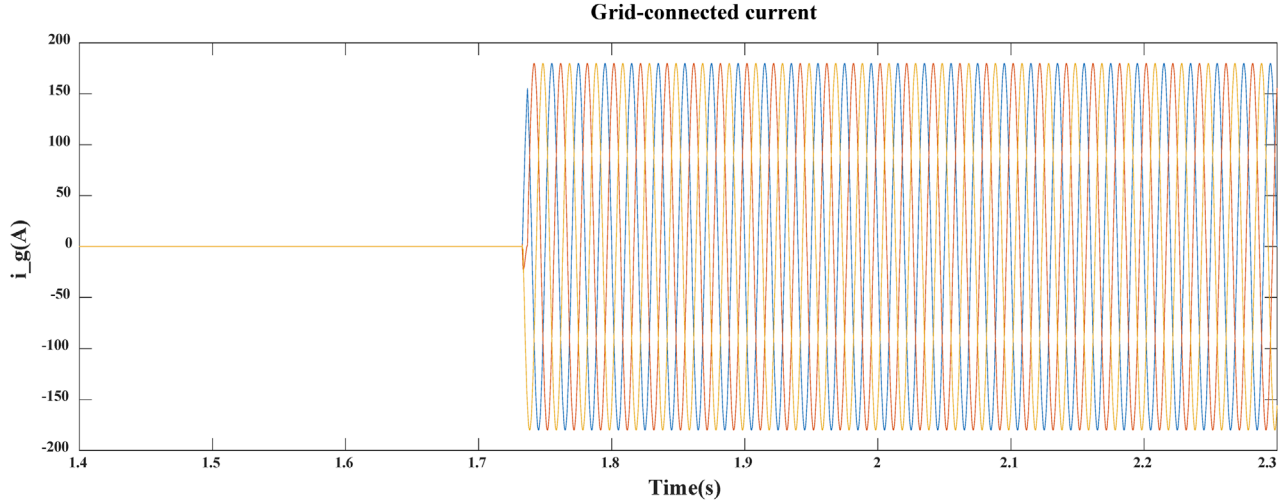


FIGURE 10
Cluster grid-connected current.

TABLE 2 Power of each unit.

Unit/power	MG1	MG2	MG3	MG4	MG5	MG6
Photovoltaic units	10 KW	10 KW	10 KW	10 KW	10 KW	10 KW
Wind power unit	20 KW	20 KW	10 KW	20 KW	20 KW	10 KW
Energy storage unit	40 KWh	40 KWh	40 KWh	40 KWh	40 KWh	40 KWh
Load unit	40 KW	10 KW	35 KW	40 KW	40 KW	40 KW
Error power	−10 KW	20 KW	−15 KW	−10 KW	−10 KW	−20 KW
Busbar voltage	380 V	380 V	380 V	380 V	380 V	380 V

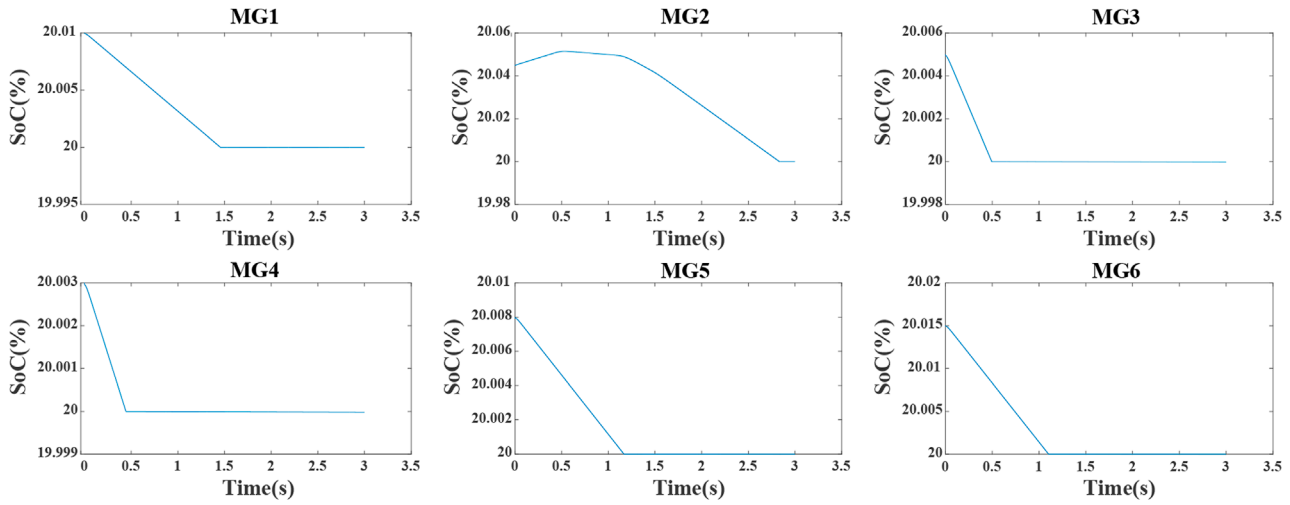


FIGURE 11
SoC value of each microgrid.

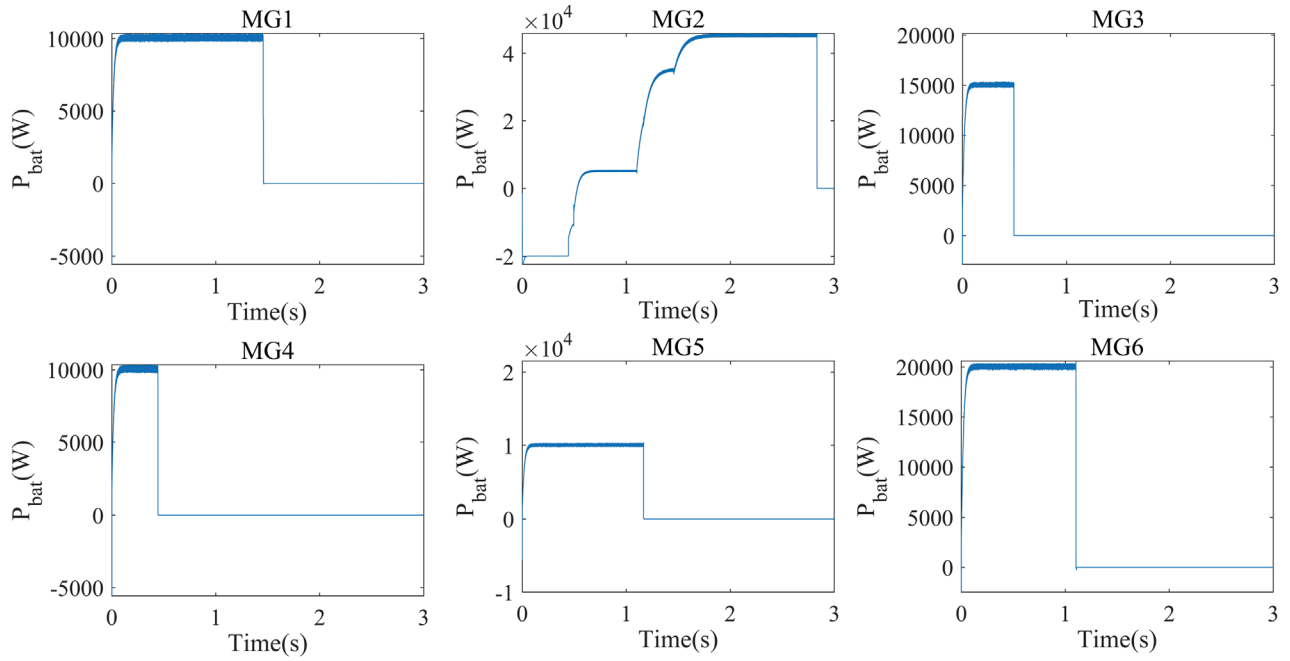


FIGURE 12
Energy storage output power of each microgrid.

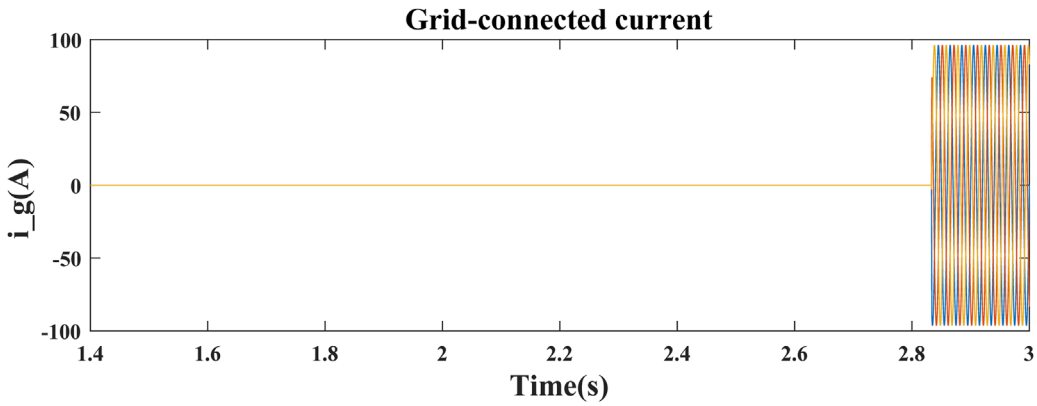


FIGURE 13
Cluster grid-connected current.

high level; The P_{error_i} of the other five microgrids is less than 0, showing a power shortage state, and their own energy storage is at a low level, so as to verify whether the power shortage cluster can carry out mutual energy assistance. First, each microgrid meets the autonomous state within 0~0.5*s*, and its own P_{error_i} is borne by itself; When it reaches close to 0.5*s*, the energy storage of MG3 and MG4 reaches the lower limit first. At this time, by detecting the microgrid that meets the power transfer condition as MG2, the microgrid connection switch of MG2, MG3, and MG4 is closed. MG2 transfers the remaining power to MG3 and MG4 to make up for its power shortage; Then, the energy storage of MG1, MG5, and MG6 reach the lower limit in turn and close the grid connection switch in turn, and the shortage of power is borne by MG2. Finally, the energy storage SoC value of MG2 also reaches the lower limit, the grid connection switch (PCC) is closed, the power shortage is compensated by the distribution network, and the energy storage is all in the standby state. The simulation results in Figures 11–13 demonstrate the operation under the above conditions, and the simulation results prove that the energy management strategy proposed in this paper is autonomous and highly robust in the case of cluster power shortage. It also proves that the control strategy proposed in this paper can still maintain the stability of the system in the case of insufficient power generation or relatively high user load, which has good practical significance.

4 Conclusion

This paper presents a centerless energy management control strategy for microgrid clusters. Centerless control is not the absence of a control center, but a control center for each microgrid distributed control system, which can simultaneously obtain information about the current status of all microgrids in the cluster and simultaneously resolve the next operation plan of the cluster. This greatly reduces the possibility of system collapse due to control center, dominant node or individual communication line failure, thus improving the robustness and autonomous operation of the system. Selection of PQ control and V/f control strategies in individual microgrids. In this paper, different operation modes are used for different operating conditions, based on the values of P_{error_i} and SoC, to achieve the coordination and autonomy of cluster energy. In addition, simulations are performed in this paper under 2 different operating conditions, and the obtained simulation results verify the effectiveness and high robustness of the proposed control strategy. And the effectiveness and feasibility of the proposed multi-agent based microgrid cluster energy management strategy is demonstrated to achieve microgrid cluster autonomy without center. It also proves that the control strategy can greatly reduce the interference to the distribution network while satisfying its own self-sufficiency rate, which has high application value both for microgrid and distribution network. Based on this, the subsequent work will simulate the control strategy under various cases, and optimize and improve the control strategy through the simulation results so that the control strategy still has high robustness in various cases.

Data availability statement

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

Author contributions

Authors SS and YW came up with this idea. The author SS developed this theory and calculations and simulations. Author SS was helpful in validating the analytical method. YW and JJ supervised and reviewed this work. All authors discussed the results and contributed to the final manuscript.

Funding

The authors acknowledge financial support from the Natural Science Foundation of Educational Department of Liaoning Province (Grant: J2020053), and Technology Innovation Fund (Grant: 2020JJ26GX029) and would like to express many thanks

to the support of Dalian Key Laboratory of Smart Micro-grid and Green Recycling Industry.

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

References

Cai Dehua, W. L. L. H. W. P., He, X., and Qian, A. (2015). "Design and Realization of cluster operation model for multi-microgrids in interconnection network," in *Electrical and energy efficiency management technology*, 64–71.

Chiradeja, P., and Ramakumar, R. (2004). An approach to quantify the technical benefits of distributed generation. *IEEE Trans. Energy Convers.* 19, 764–773. doi:10.1109/TEC.2004.827704

Cook, M. D., Robinett, R. D., Parker, G. G., and Weaver, W. W. (2017). Decentralized mode-adaptive guidance and control for dc microgrid. *IEEE Trans. Power Deliv.* 32, 263–271. doi:10.1109/TPWRD.2016.2583384

Deng, C., Guo, F., Wen, C., Yue, D., and Wang, Y. (2022). Distributed resilient secondary control for dc microgrids against heterogeneous communication delays and dos attacks. *IEEE Trans. Industrial Electron.* 69, 11560–11568. doi:10.1109/TIE.2021.3120492

Divshali, P. H., Alimardani, A., Hosseinian, S. H., and Abedi, M. (2012). Decentralized cooperative control strategy of microsources for stabilizing autonomous vsc-based microgrids. *IEEE Trans. Power Syst.* 27, 1949–1959. doi:10.1109/TPWRS.2012.2188914

Dou, C., Lv, M., Zhao, T., Ji, Y., and Li, H. (2015). Decentralised coordinated control of microgrid based on multi-agent system. *IET Generation, Transm. Distribution* 9, 2474–2484. doi:10.1049/iet-gtd.2015.0397

Greatbanks, J., Popovic, D., Begovic, M., Pregelj, A., and Green, T. (2003). "On optimization for security and reliability of power systems with distributed generation," in *2003 IEEE bologna power tech conference proceedings (IEEE)*, 1, 8.

Gu, Y., Xiang, X., Li, W., and He, X. (2014). Mode-adaptive decentralized control for renewable dc microgrid with enhanced reliability and flexibility. *IEEE Trans. Power Electron.* 29, 5072–5080. doi:10.1109/TPEL.2013.2294204

Han, R., Tucci, M., Martinelli, A., Guerrero, J. M., and Ferrari-Trecate, G. (2019). Stability analysis of primary plug-and-play and secondary leader-based controllers for dc microgrid clusters. *IEEE Trans. Power Syst.* 34, 1780–1800. doi:10.1109/TPWRS.2018.2884876

Han, Y., Li, H., Coeldo, E. A. A., Zhang, K., and Guerrero, J. M. (2017). Plug-in and plug-out dispatch optimization in microgrid clusters based on flexible communication. *J. Mod. Power Syst. Clean Energy* 5, 663–670. doi:10.1007/s40565-016-0235-2

Han, Y., Zhang, K., Coelho, E. A. A., Li, H., and Guerrero, J. M. (2018). Mas-based distributed coordinated control and optimization in microgrid and microgrid clusters: A comprehensive overview. *IEEE Trans. Power Electron.* 33, 6488–6508. doi:10.1109/TPEL.2017.2761438

Hegazy, Y., Salama, M., and Chikhani, A. (2003). Adequacy assessment of distributed generation systems using Monte Carlo simulation. *IEEE Trans. Power Syst.* 18, 48–52. doi:10.1109/TPWRS.2002.807044

Hoff, T., and Shugar, D. S. (1995). The value of grid-support photovoltaics in reducing distribution system losses. *IEEE Trans. Energy Convers.* 10, 569–576. doi:10.1109/60.464884

Jin, C., Loh, P. C., Wang, P., Mi, Y., and Blaabjerg, F. (2010). "Autonomous operation of hybrid ac-dc microgrids," in *2010 IEEE international conference on sustainable energy technologies (ICSET)* (IEEE), 1–7.

Joos, G., Ooi, B., McGillis, D., Galiana, F., and Marceau, R. (2000). "The potential of distributed generation to provide ancillary services," in *2000 power engineering society summer meeting (cat. no. 00ch37134)* (IEEE), 1762–1767.

Karimi, M., Wall, P., Mokhlis, H., and Terzija, V. (2017). A new centralized adaptive underfrequency load shedding controller for microgrids based on a distribution state estimator. *IEEE Trans. Power Deliv.* 32, 370–380. doi:10.1109/TPWRD.2016.2594866

Li, Y., Dong, P., Liu, M., and Yang, G. (2019). A distributed coordination control based on finite-time consensus algorithm for a cluster of dc microgrids. *IEEE Trans. Power Syst.* 34, 2205–2215. doi:10.1109/TPWRS.2018.2878769

Luo, Z., Geng, H., and Zhu, G. (2018). "Hierarchical cooperative control for islanded dc microgrid cluster - note: Sub-titles are not captured in xplore and should not be used," in *2018 IEEE international power electronics and application conference and exposition (PEAC)* (IEEE), 1–5.

Ma, Y.-S., Che, W.-W., Deng, C., and Wu, Z.-G. (2022). Observer-based fully distributed containment control for mass subject to dos attacks. *IEEE Trans. Syst. Man, Cybern. Syst.* 1, 897–907. doi:10.1109/TSMC.2022.3189092

Minchala-Avila, L. I., Garza-Castañon, L., Zhang, Y., and Ferrer, H. J. A. (2016). Optimal energy management for stable operation of an islanded microgrid. *IEEE Trans. Industrial Inf.* 12, 1361–1370. doi:10.1109/TII.2016.2569525

Nutkani, I. U., Loh, P. C., Wang, P., and Blaabjerg, F. (2016). Linear decentralized power sharing schemes for economic operation of ac microgrids. *IEEE Trans. Industrial Electron.* 63, 225–234. doi:10.1109/TIE.2015.2472361

Patel, D., Varma, R. K., and Seethapathy, R. (2008). *Impact assessment of distributed generation on medium voltage distribution systems of hydro one networks inc*.

Qingzhu, W., and Xue, C. (2017). The new progress in technology of regional microgrids cluster. *J. Electr. Eng.* 12, 53–59.

Quezada, V., Abbad, J. R., and Roman, T. G. S. (2006). Assessment of energy distribution losses for increasing penetration of distributed generation. *IEEE Trans. Power Syst.* 21, 533–540. doi:10.1109/TPWRS.2006.873115

Wei, Z. Y., and Xianwei, L. (2011). Research on micro-grid centralized control strategy based on active power vacancy. *Power Syst. Prot. Control* 39, 106–111.

Xing, X., Meng, H., Xie, L., Li, P., Toledo, S., Zhang, Y., et al. (2017). "Multi-time-scales energy management for grid-on multi-layer microgrids cluster," in *2017 IEEE southern power electronics conference (SPEC)* (IEEE), 1–6.

Yao, W., Wang, Y., Xu, Y., and Dong, C. (2023). Small-signal stability analysis and lead-lag compensation control for dc networked-microgrid under multiple time delays. *IEEE Trans. Power Syst.* 38, 921–933. doi:10.1109/TPWRS.2022.3169821

Yao, W., Wang, Y., Xu, Y., Lin, P., Qi, Y., and Wu, Q. (2021). Distributed layered control and stability analysis of islanded networked-microgrids. *Int. J. Electr. Power & Energy Syst.* 129, 106889. doi:10.1016/j.ijepes.2021.106889

Zhao, Z., Yuewu, W., Xu, Z., Yang, P., and Guerrero, J. M. (2017). Dynamic characteristics analysis and stabilization of pv-based multiple microgrid clusters. *IEEE Trans. Smart Grid* 10, 805–818. doi:10.1109/TSG.2017.2752640

Zhou, X., Ai, Q., and Hao, W. (2018a). Distributed optimal microgrid cluster scheduling with plug and play function. *Automation Electr. Power Syst.* 42, 106–113. doi:10.7500/AEPS20170919005

Zhou, X., Ai, Q., and Wang, H. (2017). Adaptive marginal costs-based distributed economic control of microgrid clusters considering line loss. *Energies* 10, 2071. doi:10.3390/en10122071

Zhou, X., Zhou, L., Chen, Y., Guerrero, J. M., Luo, A., Wu, W., et al. (2018b). A microgrid cluster structure and its autonomous coordination control strategy. *Int. J. Electr. Power & Energy Syst.* 100, 69–80. doi:10.1016/j.ijepes.2018.02.031

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 09 May 2023
DOI 10.3389/fenrg.2023.1181340

OPEN ACCESS

EDITED BY
Zhi-Wei Liu,
Huazhong University of Science and Technology, China

REVIEWED BY
Ghulam Hafeez,
University of Engineering and Technology, Mardan, Pakistan
Peiyuan Guan,
University of Oslo, Norway
Tiedong Ma,
Chongqing University, China

*CORRESPONDENCE
Dawei Gong,
pzhzhx@126.com

RECEIVED 07 March 2023
ACCEPTED 27 April 2023
PUBLISHED 09 May 2023

CITATION
Shen S, Chen J, Wang C, Wang C and Gong D (2023), Design of AUVs based on 3D coils positioning and distributed base station control for nuclear spent fuel pools.
Front. Energy Res. 11:1181340.
doi: 10.3389/fenrg.2023.1181340

Design of AUVs based on 3D coils positioning and distributed base station control for nuclear spent fuel pools

Shijun Shen[1], Jiaoyuan Chen[1], Chaofan Wang[1], Chenyang Wang[2] and Dawei Gong[1]*

[1]University of Electronic Science and Technology of China, Chengdu, China, [2]Key Laboratory of Thermal Power Technology, Wuhan Second Ship Design Institute, Wuhan, China

An underwater magnetic induction positioning and communication system and an energy-efficient distributed control algorithm for underwater base stations are proposed in this paper. The positioning and communication system consists of a number of base stations and an Autonomous Underwater Vehicle (AUV) equipped with three-axis source coils respectively. The AUV receives this signal and based on its amplitude and phase information is able to locate the AUV and communicate with the base station. Due to the short positioning distance of magnetic induction positioning technology, a large number of base stations need to be installed underwater, which puts high demands on the control of the base stations. In this paper, an energy-efficient distributed control algorithm for underwater base stations is proposed to enable the AUV to meet the operational requirements while minimizing the total energy consumption of the base station. According to the simulation results, the design solves the problem that traditional underwater positioning and communication equipment cannot work stably for long periods of time in a high radiated environment, with a positioning error of no more than 10 cm within a preset operating range, and the algorithm proposed in this paper is able to reduce energy wastage by about 20%.

KEYWORDS

autonomous underwater vehicles, distributed control, positioning, 3D coils, magnetic induction, nuclear power

1 Introduction

With the increasing scarcity of traditional energy sources, nuclear energy, as a new type of energy source with high energy density, has been greatly developed in recent years (Hisham and Carmine, 2016). Spent fuel generated during the power generation process is stored in pools, and leaks from spent fuel pools can cause very serious accidents if they are not detected and stopped in time (Hirano et al., 2012; Gu, 2018). Due to the high temperature, high humidity and radiation environment of nuclear pools, traditional manual inspections can have a significant impact on the health of workers. Using robots for nuclear pool operations can free workers from harsh environments (Bakari et al., 2007). Autonomous Underwater Vehicles (AUVs) are also exposed to high temperatures, humidity and radiation environment when inspecting and placing nuclear batteries, which places high demands on communication and positioning components.

In terms of communications, the complex environment underwater causes electromagnetic waves to be severely attenuated as the frequency rises (Zhang et al., 2011; Domingo, 2012). As a result, electromagnetic wave communication technologies that can be used in underwater environments often require large size coils to provide low frequency electromagnetic waves. For example, EM wave communication systems on submarines typically require towed antennas that are hundreds of meters long, which is not clearly suitable for limited water applications.

Compared to the strong attenuation of electromagnetic waves, sound waves are less attenuated in water and can travel over distances of several tens of kilometers. However, due to its narrow bandwidth, large propagation delay and the multipath effects it produces in confined spaces (Liu et al., 2012). This makes acoustic communications also have many limitations in practical applications.

The channel of MI communication is less influenced by the medium and is more stable (Akyildiz et al., 2015). In addition, MI communication has the advantage of a small coil size. It has been widely used in areas such as wireless earth-permeable communication (Sun and Akyildiz, 2010; Lin et al., 2015). The most important thing is that MI coils are made of ordinary copper wire and are not affected by nuclear radiation.

In terms of positioning, the MI positioning system works by constructing a magnetic dipole model. The three coils of the source coil are orthogonal and are excited in turn by an AC signal of the same frequency. The amplitude and phase information of the signals is obtained using the receiving coils, which can then be algorithmically obtained (Hu et al., 2012).

In terms of base station control, traditional magnetic induction positioning application scenarios are often small and do not need to consider base station control methods (Hu et al., 2005), but in nuclear pools, due to the huge area, multiple base stations need to be laid for cooperative positioning, opening them all will waste a lot of energy, so the control of base stations not only need to consider the conditions of the AUV, but also need to reduce energy consumption as much as possible.

A wireless robotic capsule endoscope for digestive tract examination using magnetic induction positioning was displayed in (Hu et al., 2005). A small permanent magnet was added to the capsule to create a permanent magnetic field in the patient and the capsule position is then determined by an external sensor. C. Hu, Z et al. proposed a two-dimensional magnetic induction positioning system for a domestic robot with base station coils mounted on the ceiling, and because the height was fixed in this scenario, only one set of coils had to be used for the base station (Hu et al., 2015). T. Li and L. Chen investigated evolutionary clustering to improve the quality of clusters for moving objects (Li et al., 2022).H. G. Zhang et al. proposed distributed algorithms to obtain the most available solution for the energy management system through local communication and computation (Zhang et al., 2017). S.S. Ge et al. proposed a magnetic induction positioning application for the positioning of ships towing anchors, which solves the problem of short magnetic induction positioning distances by means of a recursive algorithm (Ge et al., 2014). T. Li and L. Chen optimize uncertain trajectories in the road network (Li et al., 2020; Li et al., 2021). H. Xu et al. designed an underwater MI communication system based on Quasi-cyclic LDPC codes (Xu et al., 2023). G. C et al. presented a wireless communication system based on magnetic coils for underwater vehicles and verified its performance in different water conditions (Canales-Gomez et al., 2022). Malik et al. conducted a study for underground magnetic induction communication, comparing both EM and MI methods and verifying the superiority of MI communication in some scenarios (Malik et al., 2022).

This paper consists of five parts, the first part is the introduction, the second part will introduce the framework and theoretical basis of the magnetic induction positioning and communication system, the third part will describe the energy efficient distributed control algorithm for underwater base stations. The fourth part will verify the performance of the system through simulation, and the fifth part is the conclusion.

2 Design of underwater magnetic induction positioning and communication system

2.1 Working environment

The pool of a nuclear power plant is a typical water structure with limited space. Its dimensions are usually rectangular pools of about 10–30 m each in length, width and depth. The bottom of the pool is fitted with storage tanks for nuclear fuel. As a result, the water in the pools has a high level of nuclear radiation. The temperature of the pools is generally in the range of 40°C–80°C due to the large amount of heat released during the decay of the nuclear fuel.

Hence, the communication and positioning system needs to be designed with the two characteristics of radiation resistance and high temperature resistance. For high temperature conditions, although the resistivity of copper wire varies with temperature, the pattern of its variation is known and therefore the calculated parameters can be adjusted by measuring the ambient temperature, thus avoiding the effects of temperature. For radiation environments, the positioning and communication system designed in this paper can all be made radiation-proof, with only the coils being exposed to radiation, and the coils are made of ordinary copper wire and are not affected by nuclear radiation. In summary, underwater MI communication is promising in this environment.

2.2 System topology

As shown in Figure 1, the system is composed of the transmitting part and receiving part. For transmitting part, this design uses Frequency-Shift Keying (FSK), where the data needs to be passed through a numerically controlled oscillator and digital-to-analogue converter to generate a high-frequency signal, which is enlarged by a power amplifier and passed to the source coil.

For receiving part, the received voltage signal is first amplified and filtered through a band-pass filter to remove noise. After pre-processing the signal is split into two parts and passed into the position section and the communication section respectively. The positioning section samples the voltage after adjusting its magnitude

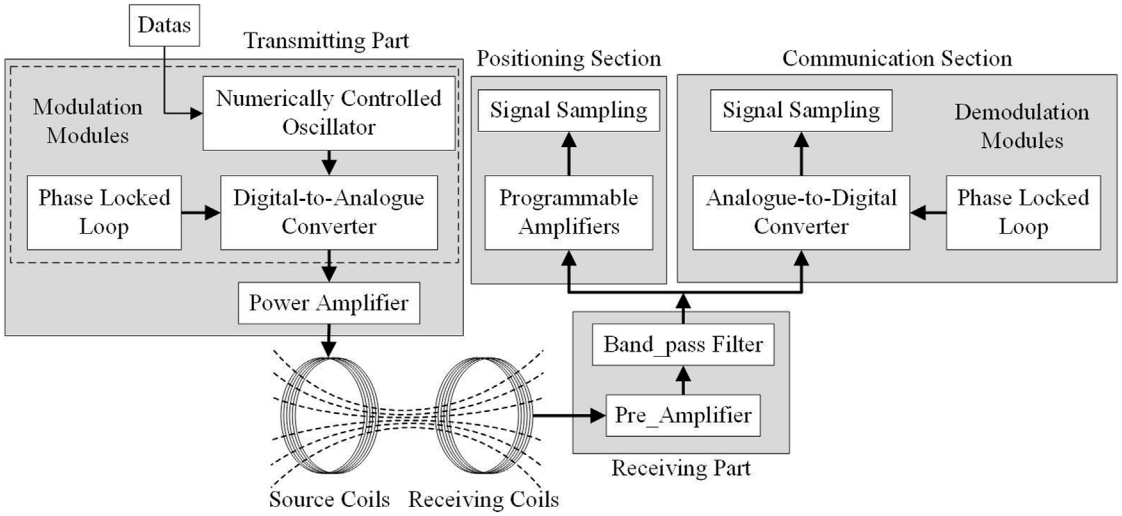


FIGURE 1
Circuit system composition.

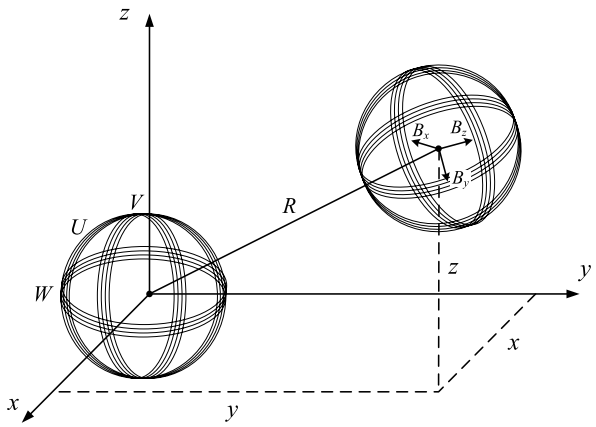


FIGURE 2
Underwater positioning model based on 3D omnidirectional magnetic induction.

with a programmable amplifier and then calculates the position of the AUV by means of a positioning algorithm.

2.3 Positioning method

The positioning system is composed of two sets of three-dimensional omnidirectional coils, with three mutually orthogonal magnetic induction coils of the source coil acting as signal sources, which in turn generate mutually orthogonal magnetic fields in space.

As shown in Figure 2, U, V, W represent the three source coils perpendicular to the *x*, *y*, *z*-axes; B_x, B_y, B_z represent the magnitude of the magnetic field induced by the three receiving coils; *R* represents the distance between the source coil and the induction coil.

A coil fed with current forms a magnetic dipole. Based on Magnetic dipole models for magnetic field distribution in three-dimensional space. The magnetic induction intensity generated by the three sets of coils U, V, W at any point p is defined as

$$\begin{cases} B_{up} = B'_{ux}\boldsymbol{i} + B'_{uy}\boldsymbol{j} + B'_{uz}\boldsymbol{k} \\ \quad = \dfrac{B_T(3x^2 - R^2)}{R^5}\boldsymbol{i} + \dfrac{3B_T xy}{R^5}\boldsymbol{j} + \dfrac{3B_T xz}{R^5}\boldsymbol{k} \\ B_{vp} = B'_{vx}\boldsymbol{i} + B'_{vy}\boldsymbol{j} + B'_{vz}\boldsymbol{k} \\ \quad = \dfrac{3B_T xy}{R^5}\boldsymbol{i} + \dfrac{B_T(3y^2 - R^2)}{R^5}\boldsymbol{j} + \dfrac{3B_T yz}{R^5}\boldsymbol{k} \\ B_{wp} = B'_{wx}\boldsymbol{i} + B'_{wy}\boldsymbol{j} + B'_{wz}\boldsymbol{k} \\ \quad = \dfrac{3B_T zx}{R^5}\boldsymbol{i} + \dfrac{3B_T zy}{R^5}\boldsymbol{j} + \dfrac{B_T(3z^2 - R^2)}{R^5}\boldsymbol{k} \end{cases} \tag{1}$$

where $B_T = \mu NIr^2/4$, μ is the magnetic permeability, $\mu = \mu_0 \cdot \mu_r$, μ_0 is the magnetic constant, μ_r is the relative permeability of water, *N* is the number of turns of coils, *I* is the current in the coil, *r* is the coil radius, *R* is the distance from the geometric center of the source coil to point p, and $\vec{i}, \vec{j}, \vec{k}$ is a unit vector in three directions of the coordinate axes.

Since the orthogonal rotation matrix does not change the magnitude of the magnetic induction intensity, we have

$$B_x^2 + B_y^2 + B_z^2 = B'^2_{ux} + B'^2_{uy} + B'^2_{uz} \tag{2}$$

Hence, by superimposing the magnetic field we have

$$\begin{aligned} B_u^2 &= B_{ux}^2 + B_{uy}^2 + B_{uz}^2 = B'^2_{ux} + B'^2_{uy} + B'^2_{uz} \\ &= B_T^2\left(\frac{3x^2}{R^8} + \frac{1}{R^6}\right) \end{aligned} \tag{3}$$

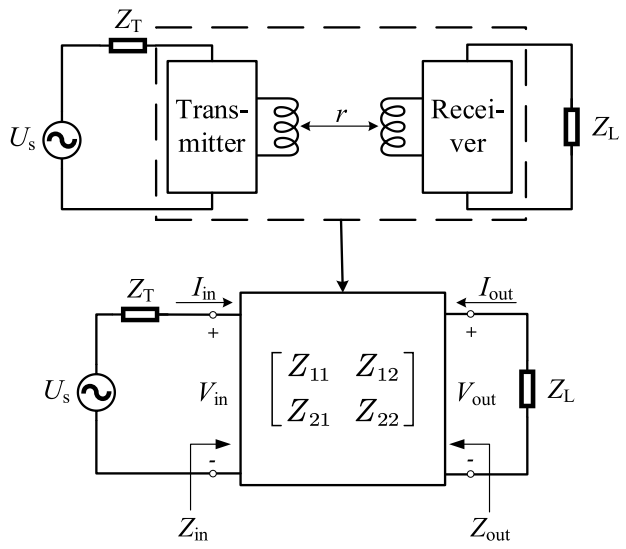


FIGURE 3
Two-port equivalent model for MI transceiver system.

Similarly, B_v and B_u of coil V and W. Therefore, combined magnetic field is defined as

$$B^2 = B_u^2 + B_v^2 + B_w^2 \tag{4}$$

Hence, the distance R is defined as

$$R = \sqrt[6]{\frac{6B_T^2}{B^2}} \tag{5}$$

Using (5) in (3), we have

$$x = \pm \frac{\sqrt{B_u^2/B_T^2 - B^2/(6B_T^2)}}{\sqrt{3}\left[B^2/(6B_T^2)\right]^{2/3}} \tag{6}$$

The process of solving for y and z is similar to that for x and will not be repeated in this paper. At this point, the position of the induction coil has been fully determined.

2.4 Communication model building and operating frequency selection

The MI communication carries data through a pair of transceiver coils, the data information is carried by a time varying magnetic field generated by a modulated sinusoidal current along the transmitter coil antenna, and the receiver coil gets the signal from this time varying field and demodulates it to obtain the information.

The MI transceiver system needed to be modeled to do a quantitative study. The transmitter part is fed by a source with an internal impedance $Z_T = R_T + X_T$ and the receiver part is terminated by a load impedance $Z_L = R_L + X_L$. The MI transceiver system can be transformed into a two-port network as depicted in Figure 3. It is assumed that the input voltage V_{in}, output voltage V_{out}, input current I_{in} and output current I_{out} of the dual antenna system are the same for the two-port network. Hence, we can use the impedance matrix to represent the relationship between the input and output of the system as

$$\begin{pmatrix} V_{in} \\ V_{out} \end{pmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{pmatrix} I_{in} \\ I_{out} \end{pmatrix} \tag{7}$$

The path loss of the MI transceiver system consists of two main components, one caused by eddy-current-loss caused by electrical conductivity in media and the other from the circuitry in the transceiver system.

For eddy-current-loss, water media tend to have a certain conductivity and coils working in water can produce eddy-current-effects which increase transmission losses dramatically. The attenuation is defined as

$$\alpha = \sqrt{\pi f \mu \sigma} \tag{8}$$

where μ is the magnetic permeability, $\mu = \mu_0 \cdot \mu_r$, μ_0 is the magnetic constant, μ_r is the relative permeability of water and σ is the electrical conductivity of sea water.

We notice from (8) that a lower operating frequency results in lower losses for defined environmental condition.

Therefore, the loss caused by eddy-current-loss in media can be defined as

$$P_{L\_e} = 20\log\left(e^{\alpha r}\right) = 8.69\alpha r \tag{9}$$

Next, without considering the effect of media on the transceiver system and based on the two-port network model of the transceiver system, we can obtain the path loss by comparing the transmit power and the receive power.

Therefore, P_{L_ts} is defined as

$$P_{L\_ts} = -10\log\frac{P_R}{P_T} \tag{10}$$

The transmitted and received powers are given by

$$\begin{aligned} P_T &= \mathrm{Re}(Z_{in})|I_1|^2 \\ &= \mathrm{Re}\left(Z_{11} - \frac{Z_{12}^2}{Z_L + Z_{22}}\right)|I_1|^2 \end{aligned} \tag{11}$$

$$\begin{aligned} P_R &= \mathrm{Re}(Z_L)\left|I_2^2\right| \\ &= \mathrm{Re}(Z_L)\frac{|Z_{12}|^2}{|Z_L + Z_{12}|^2}|I_1|^2 \end{aligned} \tag{12}$$

During the operation of the MI transceiver system, the power at the transmitter side is not actually lost, but not fully absorbed by the receiver coil, so the influence of the receiver side should be ignored in the power calculation of the transmitter coil, the transmitted power of the MI system is redefined as follows:

$$P_T = \mathrm{Re}(Z_{11})|I_1|^2 \tag{13}$$

Using (11)-(13) in (10), P_{L_ts} is given as

$$P_{L\_ts} = -10\log\frac{R_L\left(R_{12}^2 + X_{12}^2\right)}{R_{11}\left[\left(R_L + R_{22}\right)^2 + \left(X_L + X_{22}\right)^2\right]} \tag{14}$$

We notice from (12) that the received power depends on the load impedance Z_L, using the principle of impedance matching, the

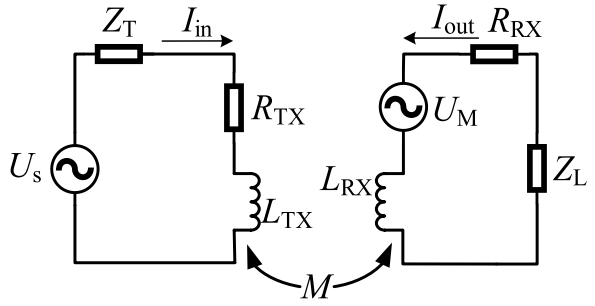


FIGURE 4
Equivalent circuit for coupled coils.

load receives the most power when $Z_L = \bar{Z}_{out}$ so the receiver side load is defined as

$$Z_L = Z_{out} = Z_{22} - \frac{Z_{12}^2}{Z_T + Z_{11}} \tag{15}$$

Next the impedance matrix of the two-port network needs to be calculated. As shown in Figure 4, we equate the MI transceiver system to an inductively coupled circuit, L_{TX} and L_{RX} represent the self-inductance of the transmitter and receiver antennas. R_{TX} and R_{RX} represent the coil resistance of the transmitter and receiver antennas. U_M is caused by the time varying flux in the coil L_{TX} through the mutual inductance M.

Hence, we can obtain the parameters of the impedance matrix by solving the model in Figure 4.

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} = \begin{bmatrix} R_{TX} + j\omega L_{TX} & j\omega M \\ j\omega M & R_{RX} + j\omega L_{RX} \end{bmatrix} \tag{16}$$

where $w = 2\pi f$ is the angular frequency of the transmitting signal.

The self-inductance L of a coil is obtained as

$$L = \frac{\mu \cdot N^2 \cdot A}{l} \tag{17}$$

where N is the number of turns of coil, l is the length of the solenoid and A is the cross-sectional area of the wire.

The mutual inductance M between two coils can be calculated as

$$M = \frac{\pi \cdot \mu \cdot N_{TX} \cdot a_{TX}^2 \cdot N_{RX} \cdot a_{RX}^2}{2\sqrt{(a_{TX}^2 + r^2)^3}} \tag{18}$$

where a_{TX} and a_{RX} are the radii of the transmitter and receiver coils, N_{TX} and N_{RX} are the number of turns of the transmitter and receiver coils and r is the distance between the transmitter and receiver coils.

Based on application scenario considerations, it is assumed that the transmitter and receiver coils are completely same, assuming $a_{TX} = a_{RX} = a$, $N_{TX} = N_{RX} = N$, the mutual inductance M between two coils is redefined as follows:

$$M = \frac{\pi \cdot \mu \cdot N^2 \cdot a^4}{2\sqrt{(a^2 + r^2)^3}} \tag{19}$$

The coil resistance is given by

$$R = \frac{N \cdot 2\pi \cdot a \cdot \rho}{A} \tag{20}$$

where ρ is the electrical resistivity.

Assuming there is no impedance in the transmitter circuit and using (15)-(20) in (14), P_{L_ts} is defined as

$$P_{L\_ts} = -10\log\frac{R_L \cdot \omega^2 \cdot M^2}{R_{TX}(R_L + R_{RX})^2 + R_{TX}(X_L + \omega L_{RX})^2} \tag{21}$$

Therefore, the path loss P_L is defined as

$$P_L = P_{L\_ts} + P_{L\_e} \tag{22}$$

The path loss as a function of the frequency and the transmission distance caused by eddy-current-effects and transceiver system are shown in Figure 5. The number of turns is 200, the radius of the transmitter and receiver is 0.2 m and the diameter of the wire is 0.3 mm.

We observe that the path loss caused by eddy-current-effects in Figure 5A is increased dramatically with the frequency and the transmission distance. The attenuation in formulas (8) is increased with the frequency and the eddy-current-loss defined in formulas (9) is increased with the transmission distance. Therefore, the higher the electrical conductivity of the system's operating environment, the greater the effect of frequency on eddy current losses.

In Figure 5B, we observe that the path loss caused by transceiver system is increased with the transmission distance. Within 5 m, the path loss increases sharply, and the upward trend is slower out of 5 m. However, as the frequency increases, the path loss shows a decreasing trend at the same distance. Therefore, designing for higher frequencies will reduce the path loss of the system.

Hence, eddy-current-losses increase with frequency and transceiver-system-losses decreases with frequency, so we can obtain the optimum operating frequency based on the parameters of the coil and the data of the operating environment. Combining Figures 5A, B we can obtain the total losses of the system for this operating condition as shown in Figure 5C, where the optimum frequency is set to 2 MHz.

3 Energy efficient distributed control algorithm for underwater base stations

3.1 Control models

From the above derivation it is clear that $\propto (1/R^3)$, the positioning accuracy and communication quality in this design is very sensitive to distance, and therefore multiple base stations are required to ensure safe and stable operation in a wide range of applications such as nuclear pools.

In traditional indoor multi-base station positioning applications, multiple base stations are often switched on to improve positioning accuracy, while in magnetic induction positioning and communication systems, the magnetic field decreases rapidly with distance, resulting in a large amount of energy wastage. As shown in Figure 6, this paper proposes a distributed control base station switching method, as well as three operating modes based on the different positioning accuracy requirements of underwater robots

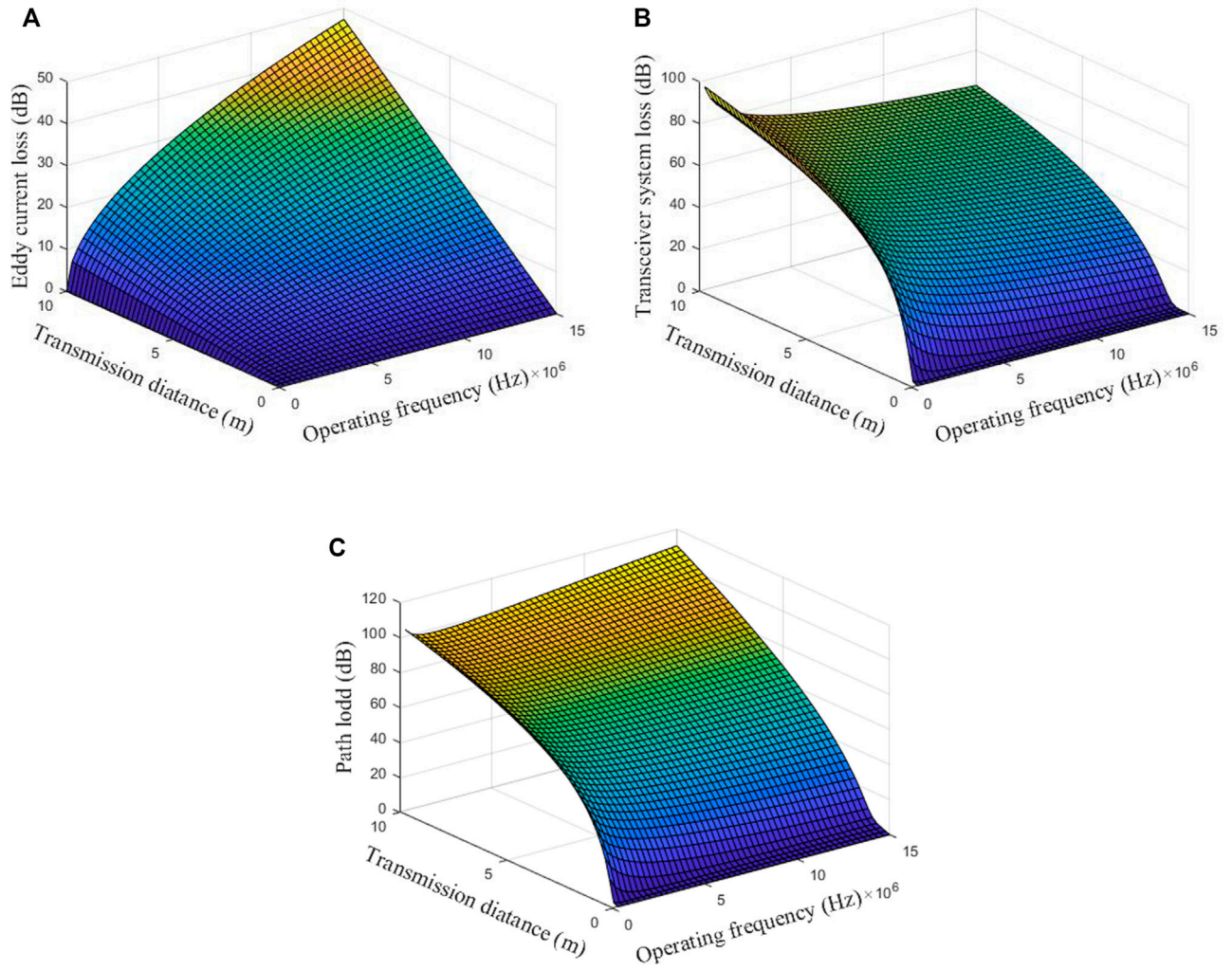


FIGURE 5
Path loss of the underwater MI communication channel. **(A)** Path loss caused by eddy-current-effects. **(B)** Path loss caused by transceiver system. **(C)** The total path loss.

under different operating conditions, to ensure the normal operation of the robot and to significantly reduce energy wastage.

3.2 Analysis of base station operating modes and AUV operating modes

This paper divides the base station operating modes into two.

(1) Single base station operating mode

When low positioning accuracy is required, the base stations operate in this mode and only the nearest base station to the AUV is switched on at the same point in time for the whole system.

(2) Multi-base station operation mode

When the positioning accuracy is high, the base station operates in this mode, as shown in Figure 7. At point a, as it is closer to base station 1, base station 1 can meet its positioning and communication needs, so only base station 1 can be turned on at this location; at point b, as it is a certain distance away from both base station 1 and base station 2, so base station 1 and base station 2 are turned on to ensure the positioning accuracy and communication quality; at point c, as it is farther away from both base station 1 and base station 2, so base station 3 is turned on at the same time to further ensure its positioning accuracy and communication quality.

This paper sets out three operating modes for AUVs.

(1) Normal marching mode.

In this mode, the AUV requires low positioning accuracy and therefore uses a single base station operating mode. As shown in Figure 8A, In phase A, the AUV is closer to base station 2 and therefore turns on base station 2. In a similar way, phase B corresponds to base station 3, phase C to base station 1.

(2) Fixed-point working mode.

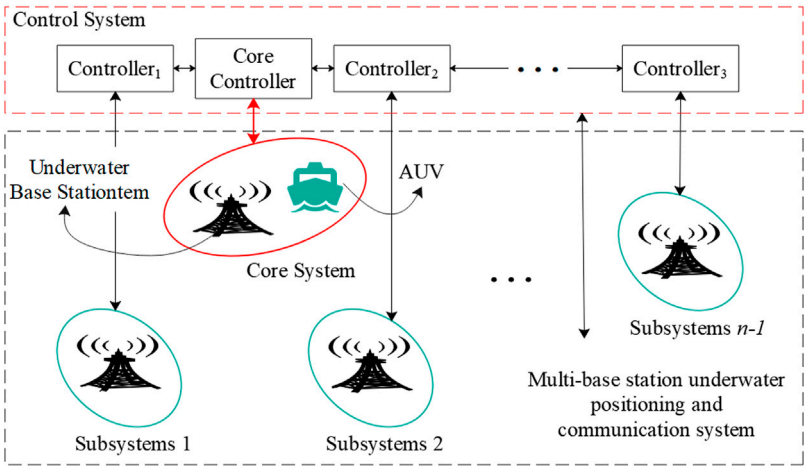


FIGURE 6
Distributed control of a magnetic induction positioning and communication system.

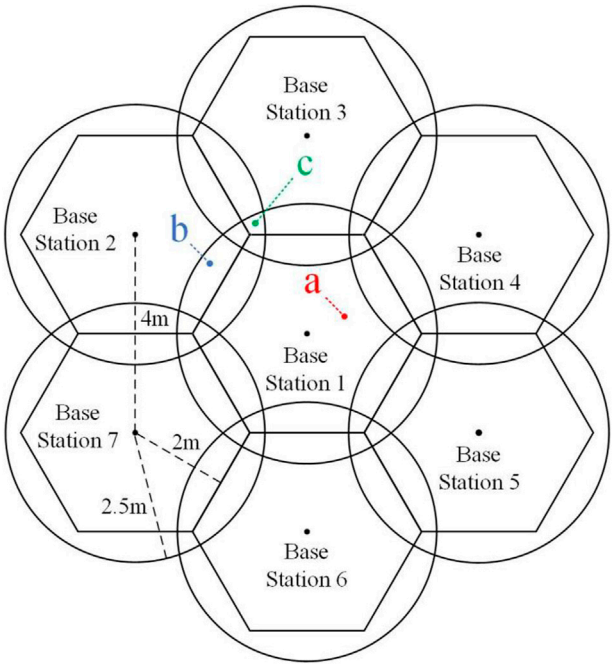


FIGURE 7
Multi-base Underwater base station distribution.

As shown in Figure 8B, when the AUV needs to operate at a fixed-point, single base station operating mode may not be able to meet its required positioning accuracy. This mode uses Multi-base station operation mode once the AUV arrives at the operating point, improving its positioning accuracy and communication quality to meet the system's requirements.

(3) Wrap-around detection mode.

In this mode, the AUV requires high positioning accuracy and therefore uses a multi-base station operating mode. The AUV starts from the starting point and stops when it reaches the detection radius of the object to be detected, then the AUV starts to lower to the set detection height and loop around the object. After completing the detection the AUV rises above the height of this object and proceeds to the next object to be detected. During this time the system is in multi-base station operating mode, to avoid any accidental position deviations that could cause the AUV to collide with the object. The path of the AUV run is shown in Figure 8C, green line indicates an open base station, yellow line means two and red line means three.

The basic working logic of AUV is as follows: After the AUV has been activated, it will first confirm its current position, receive the current job assignment and analyze the required base station operating modes. Next the AUV will perform a heading calculation based on the current position and the target position and proceed as calculated. During the course of the voyage the base station will be switched on and off in ways as described above. The AUV may deviate from the desired course due to uncertain underwater disturbances. The positioning system therefore operates continuously at regular intervals to calculate the latest position of the AUV and to correct the heading. After completing the expected work the AUV will wait underwater for a new work order to be given.

4 Simulation analysis

4.1 Settings

The experimental results of the underwater magnetic induction positioning and communication system will be presented in terms of positioning accuracy of the 3D coil and operating condition simulation. The various parameters of the coil are shown in Table 1.

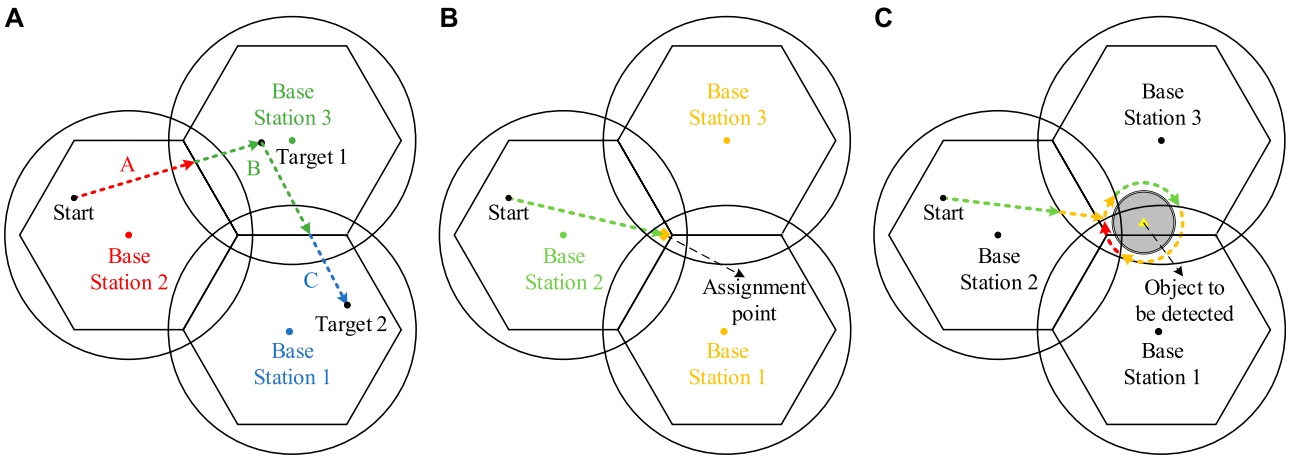


FIGURE 8
Working paths for different operating modes **(A)** Normal marching mode. **(B)** Fixed-point working mode.**(C)** Wrap-around detection mode.

TABLE 1 Parameters of coils.

Parameters	Data
Number of turns	200
Radius of coil [m]	0.2
Width of coil [m]	0.03
Diameter of wire [mm]	0.3
Operating frequency [MHz]	2
Electrical resistivity [Ω·mm²/m]	0.0185

4.2 Normal marching mode simulation

In the normal marching mode, the starting point is (2, 2, 4) and the target point is (10, 6, 6), and the trajectories of AUVs in single base station operating mode and multi-base station operating mode are simulated respectively.

As shown in Figure 9A, when the AUV runs near the area where the working range of the two base stations meet, the positioning accuracy drops due to the opening of only one base station, and the trajectory of the AUV shows a more obvious deviation. Although the AUV can correct its own course to reach the target point, the course deviation makes the AUV's running time longer. Its total elapsed time is 23.3 s, the same as the base station turn-up time.

When using the multi-base station operating mode, as shown in Figure 9B, the AUV is still able to maintain a high level of positioning accuracy in the area where the working range of the two base stations meet, resulting in more accurate navigation of the AUV compared to the former and therefore a shorter running time in this mode. Its time was only 22.9 s. However, due to multiple base stations operating together over a period of time, the total turn-up time of its base stations reached 26.1 s. Total base station turn-up time improved by 2.8 s compared to single base station operating mode, in non-emergency situations, the multi-base station operation mode can waste a lot of energy.

When the AUV crosses the area where the working range of the three base stations meet, this wastefulness will be exacerbated. As shown in Figure 10, after changing the target position to (9, 7, 1), the AUV will cross several triple base stations crossover areas. In single

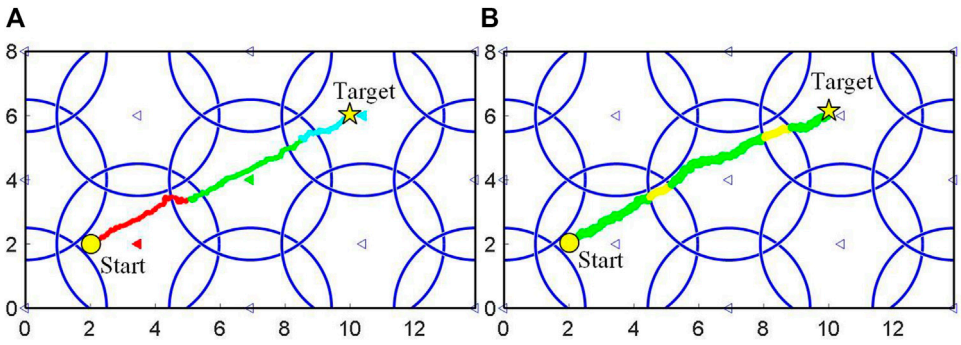


FIGURE 9
AUV's trajectories in the normal marching mode [target point (10,6,6)]:**(A)** single base station operating mode;**(B)** multi-base station operating mode.

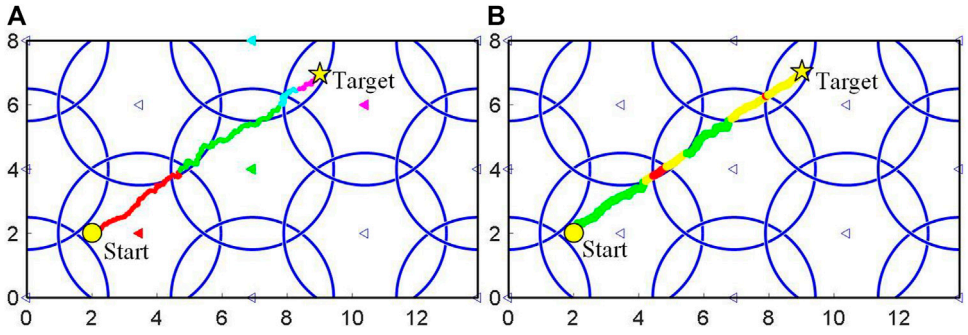


FIGURE 10
AUV's trajectories in the normal marching mode [target point (9,7,1)]:**(A)** single base station operating mode;**(B)** multi-base station operating mode.

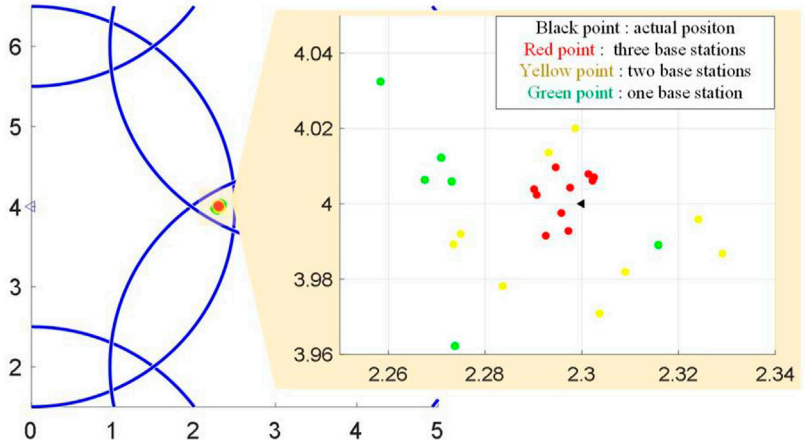


FIGURE 11
Continuous positioning results of AUV with different number of base stations (xy-projection plane).

base station operating mode, the total base station turn-on time is only 23.7 s, while in multi-base station operation mode it is 27.5 s.

Simulations were carried out using the two base station operating modes by replacing different starting and target points. Experimental results showed that both allowed the robot to reach the target location, with an acceptable average increase in robot run time of 4.3% for the single base station mode compared to the multiple base station mode, and an average reduction in total base station turn-on time of 15.6%, with a maximum reduction of 24.2%. The algorithm keeps the increase in robot runtime within acceptable limits, while significantly reducing base station turn-up time and saving energy.

4.3 Fixed-point working mode simulaiton

As the point (2.3,4,2) is located at the junction of three base stations and the results are more representative, it is chosen as the operation point to verify the fixed-point operation mode in this paper.

As shown in Figure 11, When three base stations are on, the results are closer to the actual position of the AUV; when two base stations are on, the next closest, and the worst when a single base station is on.

In this paper, the squared heel of the deviation in the three directions is defined as the localization error

$$E_p = \sqrt{(x_a - x_c)^2 + (y_a - y_c)^2 + (z_a - z_c)^2} \tag{23}$$

where (x_a, y_a, z_a) is the actual position of the AUV and (x_c, y_c, z_c) is the calculated position of the AUV.

Therefor, the positioning deviation of AUV with different number of base stations is presented in Table 2. From Table 2, we can see that with the increase in the number of base stations, the positioning effect is significantly improved, and when three base stations are working simultaneously, the positioning error of the system can be controlled to within 5 cm.

TABLE 2 Positioning deviation of AUV with different number of base stations.

	One base station	Two base stations	Three bass stations
Positioning deviation in *x*-direction [m]	0.0729	0.0510	0.0286
Positioning deviation in *y*-direction [m]	0.0804	0.0501	0.0356
Positioning deviation in *z*-direction [m]	0.0375	0.0257	0.0193
Positioning deviation [m]	0.1149	0.0759	0.00496

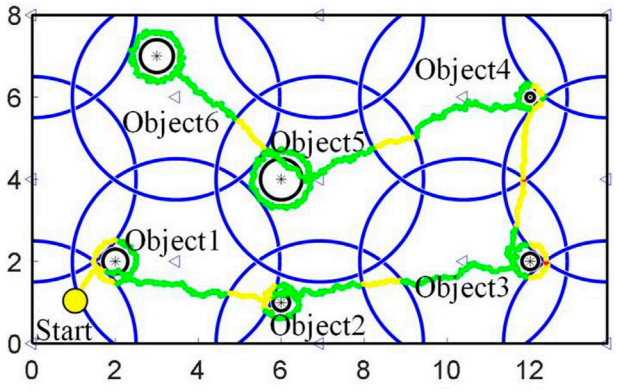


FIGURE 12
AUV's trajectory in the wrap-around detection mode.

4.4 Wrap-around detection mode simulation

In the wrap-around detection mode, the starting point is (1,1,4) and the coordinates of the objects to be detected are (2,2,1), (6,1,1), (12,2,1), (12,6,1), (6,4,1), (3,7,1); the radius of the objects to be detected are 3,4,1,2,5,3; the height of the objects to be detected are 0.3,0.2,0.2,0.1,0.5,0.4. (All units above are in meters).

The trajectory of the AUV is shown in Figure 12, the total time for the rise and fall of the AUV was 57 s, during which the base station was turned on in 78.9 s; the time for the AUV to move horizontally was 47.2 s, during which time the base station was turned on for 63.5 s; the time for AUV wrap-around detection was 36.7 s, during which time the base station was turned on for 45.8 s. The whole process took 140.9 s and 188.2 s for the base station turn-on. The wheel path planning algorithm enables the AUV to perform underwater surround detection tasks with a significant reduction in base station turn-up time and energy savings.

The simulation results prove that the proposed energy saving control algorithm for distributed underwater base station ensures that the AUV is able to perform its tasks under a variety of set conditions while saving energy wastage due to the opening of redundant base stations.

5 Conclusion

An underwater magnetic induction positioning and communication system and an energy-efficient distributed control algorithm for underwater base stations are proposed in this paper. The different requirements for positioning accuracy and communication quality under different operating conditions of AUVs are analyzed and two different base station control modes and a reasonable distributed control strategy for the three working states of AUVs. This system solves the problem of wireless positioning and communication in high radiation underwater environments, while the energy consumption of the base station is significantly reduced under the premise of ensuring the normal operation of the AUV, but the problem of how to improve the positioning and communication distance of this system is urgently needed by the serious attenuation of the magnetic field strength as the distance becomes larger. Experimental results show that with this control algorithm, the AUV can perform its tasks under three pre-defined operating conditions; in addition, the algorithm can reduce the energy loss by about 20% when moving over a large area and passing through more base station nodes. Compared to previous control algorithms, this paper reduces the total energy consumption of the system operation while satisfying the AUV operating conditions. In the future the system could be used in the daily work of nuclear spent fuel pools, freeing workers from the harsh working conditions.

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

Methodology, SJ; software, JC; formal analysis, SS and JC; investigation, CFW; data curation, CYW and CFW; writing-original draft preparation, JC; visualization, JC; supervision, DG; project administration, SS; funding acquisition, JC. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.All authors contributed to the article and approved the submitted version.

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

References

Akyildiz, I. F., Wang, P., and Sun, Z. (2015). Realizing underwater communication through magnetic induction. *IEEE Commun. Mag.* 53 (11), 42–48. doi:10.1109/MCOM.2015.7321970

Bakari, M. J., Zied, K. M., and Seward, D. W. (2007). Development of a multi-arm mobile robot for nuclear decommissioning tasks. *SAGE Publ.* 4 (4), 51–406. doi:10.5772/5665

Canales-Gomez, G., Leon-Gonzalez, G., Jorge-Munoz, N., Arroyo-Nunez, J. H., Antonio-Yanez, E. D., and Nunez-Cruz, R. S. (2022). Communication system based on magnetic coils for underwater vehicles. *Sensors* 22, 8183. doi:10.3390/s22218183

Domingo, M. C. (2012). Magnetic induction for underwater wireless communication networks. *IEEE Trans. Antennas Propag.* 60 (6), 2929–2939. doi:10.1109/TAP.2012.2194670

Ge, S. S., Zhao, Z., He, W., and Choo, Y. S. (2014). Localization of drag anchor in mooring systems via magnetic induction and acoustic wireless communication network. *IEEE J. Ocean. Eng.* 39 (3), 515–525. doi:10.1109/JOE.2013.2271957

Gu, Z. (2018). History review of nuclear reactor safety. *Ann. Nucl. Energy* 120, 682–690. doi:10.1016/j.anucene.2018.06.023

Hirano, M., Yonomoto, T., Ishigaki, M., Watanabe, N., Maruyama, Y., Sibamoto, Y., et al. (2012). Insights from review and analysis of the fukushima dai-ichi accident. *J. Nucl. Sci. Technol.* 49 (1-2), 1–17. doi:10.1080/18811248.2011.636538

Hisham, K., and Carmine, D. (2016). Economics of nuclear and renewables. *Energy Policy* 96, 740–750. doi:10.1016/j.enpol.2016.04.013

Hu, C., Feng, Z. Q., Ren, Y. P., Chen, Y., and Bao, J. (2015). "An efficient magnetic localization system for indoor planar mobile robot," in 34th Chinese Control Conference (CCC), Hangzhou, 28-30 July 2015, 4899–4904.

Hu, C., Meng, M. Q., and Mandal, M. (2005). "Efficient magnetic localization and orientation technique for capsule endoscopy," in IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, 02-06 August 2005, 3365–3370.

Hu, C., Song, S., Wang, X. J., Meng, M. Q. H., and Li, B. (2012). A novel positioning and orientation system based on three-Axis magnetic coils. *IEEE Trans. Magnetics* 48 (7), 2211–2219. doi:10.1109/TMAG.2012.2188537

Li, T., Chen, L., Jensen, C. S., Pedersen, T. B., Gao, Y. J., and Hu, J. L. (2022). "Evolutionary clustering of moving objects," in Proc. 38th IEEE Int. Conf. Data Mining, Kuala Lumpur, 09-12 May 2022, 2399–2411.

Li, T., Chen, L., Jensen, C. S., and Pedersen, T. B. (2021). Trace: Real-Time compression of streaming trajectories in road networks. *Proc. 47th Int. Conf. Very Large Data Bases* 13 (7), 1175–1187. doi:10.14778/3450980.3450987

Li, T., Huang, R., Chen, L., and Pedersen, T. B. (2020). Compression of uncertain trajectories in road networks. *Proc. 46th Int. Conf. Very Large Data Bases* 13 (7), 1050–1063. doi:10.14778/3384345.3384353

Lin, S. C., Akyildiz, I. F., Wang, P., and Sun, Z. (2015). Distributed cross-layer protocol design for magnetic induction communication in wireless underground sensor networks. *IEEE Trans. Wirel. Commun.* 14 (7), 4006–4019. doi:10.1109/TWC.2015.2415812

Liu, L. B., Zhou, S. L., and Cui, J. H. (2012). Prospects and problems of wireless communication for underwater sensor networks. *Wirel. Commun. Mob. Comput.* 8 (8), 977–994. doi:10.1002/wcm.654

Malik, P. S., Abouhawwash, M., Almutairi, A., Singh, R. P., and Singh, Y. (2022). Comparative analysis of magnetic induction based communication techniques for wireless underground sensor networks. *Peerj Comput. Sci.* 8, e789. doi:10.7717/peerj-cs.789

Sun, Z., and Akyildiz, I. F. (2010). Magnetic induction communications for wireless underground sensor networks. *IEEE Trans. Antennas Propag.* 58 (7), 2426–2435. doi:10.1109/TAP.2010.2048858

Xu, H., Shi, W. J., and Sun, Y. J. (2023). Performance analysis and design of quasi-cyclic LDPC codes for underwater magnetic induction communications. *Phys. Commun.* 56, 101950. doi:10.1016/j.phycom.2022.101950

Zhang, H., Geng, D. W., Zhang, G. P., and Gulliver, T. A. (2011). "The impact of antenna design and frequency on underwater wireless communications," in IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Victoria, Canada, 23-26 August 2011, 868–872.

Zhang, H. G., Li, Y. S., Gao, D. W., and Zhou, J. (2017). Distributed optimal energy management for energy internet. *IEEE Trans. Industrial Inf.* 13 (6), 3081–3097. doi:10.1109/TII.2017.2714199

frontiers | Frontiers in Energy Research

TYPE Methods
PUBLISHED 16 May 2023
DOI 10.3389/fenrg.2023.1152032

OPEN ACCESS

EDITED BY
Zhi-Wei Liu,
Huazhong University of Science and Technology, China

REVIEWED BY
Bowen Zhou,
Northeastern University, China
Ruizhuo Song,
University of Science and Technology Beijing, China

*CORRESPONDENCE
Tianfeng Chu,
15524182863@126.com

RECEIVED 27 January 2023
ACCEPTED 26 April 2023
PUBLISHED 16 May 2023

CITATION
Chu T, An X, Zhu Y, Wang X and Deng L (2023), Multiple game trading strategy of multiple virtual power plants considering carbon trading.
Front. Energy Res. 11:1152032.
doi: 10.3389/fenrg.2023.1152032

Multiple game trading strategy of multiple virtual power plants considering carbon trading

Tianfeng Chu[1,2]*, Xingchen An[1], Yu Zhu[2], Xueying Wang[2] and Limin Deng[1]

[1]College of Information Science and Engineering, Northeastern University, Shenyang, China, [2]State Grid Liaoning Electric Power Co., Ltd., Electric Power Research Institute, Shenyang, China

The large-scale access of distributed energy resources has a certain impact on the power grid, so distributed energy resources cannot participate in the power market transactions alone. The concept of the virtual power plant (VPP) has thus emerged, which can aggregate distributed power sources and controllable loads in a region for coordinated regulation. The trading of VPPs should not only consider the economy but also its degree of low carbon. Therefore, this paper constructs a unified bidding strategy for multi-VPPs that considers carbon–electricity integration trading. We design a multi-game trading strategy among multi-VPPs to achieve unified trading, after each VPP determines its internal trading strategy. Finally, through simulation, we verify that the multiple game strategy between multi-VPPs that considers carbon trading here proposed can effectively improve the efficiency and trading income of VPPs and promote the consumption of new energy.

KEYWORDS

integrated carbon–electricity trading, virtual power plant trading, multiple game strategy, stepped carbon trading, distributed optimization

1 Introduction

Due to the fluctuating power output and high uncertainty of each distributed energy source, direct participation in grid dispatch and power trading is not possible (Huang et al., 2019). VPP can aggregate generation and consumption units in the region to form a "self-generating and self-consuming" whole (Liu et al., 1109) which can be self-sufficient and conduct two-way power trading with external interests (Wang et al., 2022a). It can conduct direct trading between VPPs or choose to trade with the grid, reducing power supply pressure on the main network and ensuring that renewable energy can be consumed locally (Shayegan-Rad et al., 2017).

At present, research on the trading of individual VPP is relatively mature (Feng et al., 2018). Improving the uncertainty that the system will face in ensuring that VPPs can participate in the medium and long-term electricity market can improve its robustness. In the current electricity market environment, the number of subjects that can participate in electricity trading is gradually increasing, and the competition is more diversified. Therefore, the influence of the bidding strategies of other competitors amongst multi-VPPs must be considered (Zhi et al., 2021). Aggregated electric vehicles with dual characteristics of energy storage and load use market clearing to characterize the relationship between other VPPs and their own strategies and use non-cooperative game method to conduct day-ahead transactions.

Existing trading models for VPPs tend to focus only on economic benefits and ignore the additional environmental costs incurred in the generation of electricity by the internal units of VPPs, and on the process of purchasing electricity, which can inhibit the carbon reducing potential of VPPs. Some literature has also studied the low-carbon operation of VPPs (Chu et al., 2023). Some has considered the coordination and optimization of the environment and economy in a single VPP trading model and reducing pollution caused by VPP operations (Zhao et al., 2020). Other literature has considered the uncertainty faced by the power system and the impact of carbon emissions on the conservatism of VPP bidding, building a low-risk VPP bidding model to analyze the impact of carbon emissions (Liu et al., 2023). A VPP low-carbon economic dispatch has been proposed that accounts for the user electricity consumption behavior considering carbon trading (Huang et al., 2022). The impact of carbon consumption constraints on VPP participation in electricity market bidding has been considered; the results demonstrate that carbon emission constraints can change the bidding strategies of VPPs.

Considering these problems and challenges, this paper will limit the carbon utilization in VPP operations based on stepped carbon trading so as to improve renewable energy consumption. First, it establishes a multi-VPPs model consisting of wind turbines (WTs), photovoltaic power generation (PV), energy storage systems (ESSs), micro-turbines (MTs), and controllable loads (CL) to study how they participate in integrated carbon–electricity trading. In the trading process, each VPP should consider not only its own trading strategy but also the impact of other participants' trading strategies on itself. This paper designs a multiple game mechanism for multi-VPPs and uses a distributed solution method to determine the final game equilibrium solution. Finally, by setting different scenarios, the accuracy and rationality of the proposed method are verified. The following principal contributions we make are:

1) A stepped carbon trading mechanism is introduced into the power trading system of VPPs to form integrated carbon–electricity trading of multi-VPPs, thus reducing the system's carbon emissions. At the same time, the introduction of a ladder-type carbon emission transaction strategy can improve the level of clean energy consumption.
2) We propose a multiple game trading mechanism among multi-VPPs which no longer uses a one-game approach with unstable game equilibrium points, as in previous studies. Non-cooperative and evolutionary games are introduced in the framework of a master–slave game.
3) For coordinated optimization problems between multiple regions, the use of distributed solution algorithms can greatly accelerate the iteration speed and improve the solution's efficiency.

This article is organized thus: Section 2 focuses on the integrated carbon–electricity trading market for VPPs; Section 3 introduces a proposed multiple game model; Section 4 is the solution method; Section 5 performs the simulation analysis; Section 6 draws conclusions.

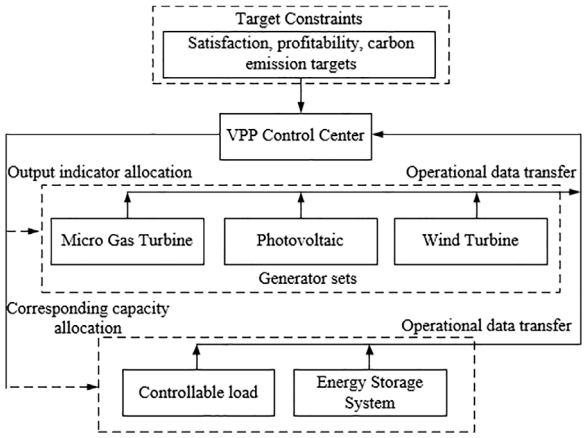


FIGURE 1
Operation structure of typical VPP.

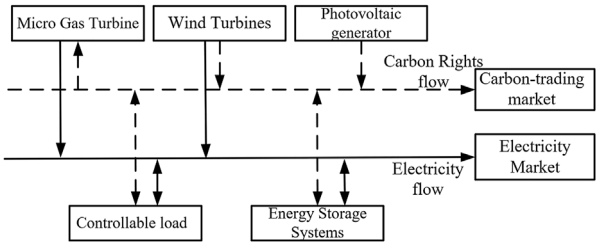


FIGURE 2
Integrated carbon–electricity trading volume.

2 Integrated carbon–electricity trading model

2.1 Virtual power plant operation structure analysis

In traditional power systems, wind power, photovoltaic, and various types of loads participate independently in the market, and the grid lacks real-time awareness of each unit's operating status (Li et al., 2020). The system cannot come to a unified dispatch and trading method due to distributed power resources, the power characteristics of differentiated loads, and the carbon emission level of each unit (Wang et al., 2022b). Building an integrated carbon–electricity market needs to rely on VPP to aggregate distributed energy, load, and energy storage systems in a certain region (Zhu et al., 2022). The structure of VPP operation is shown in Figure 1.

At the data level, each unit in a VPP needs to report internal data to the VPP control center (Dou et al., 2022), which then allocates the output and load of each generator unit and the response capacity of ESS according to the constraints of current user satisfaction, income, carbon emissions, and other objectives.

2.2 Characterization of integrated carbon–electricity trading market

Carbon trading is the trading of carbon credits as commodities. It is crucial to measure and count the carbon credits and also to monitor their non-reuse. Carbon credits flow together with electricity trading, and the credits and corresponding benefits are transferred together with the electricity (Zhang et al., 2017). The flow of an integrated carbon–electricity transaction in a VPP is shown in Figure 2.

The trading features are.

1) No carbon emission is generated while WT and PV operate, and the VPP centralized control center can allocate the corresponding carbon emission to MT or sell them to obtain additional economic benefits (Liu et al., 2021);
2) If VPP invokes CL for demand response in the dispatch, it needs to give corresponding carbon credits to users as economic compensation.

2.3 Calculation of VPP stepped carbon trading cost

Regulatory authorities will first allocate a certain amount of free carbon emissions to each carbon emission source and ensure that total amounts decrease annually (Liu, 2022). Each VPP will formulate and adjust the scheduling and bidding plan according to the allocated amount. Therefore, there are two situations in the process of carbon trading by VPPs (Chen et al., 2021b): the first is that the actual amount of carbon emissions generated in the VPP scheduling is more than the free quota allocated by the regulatory authority, so the VPP needs to pay for the excess in the market; in the second case, when the actual amount of carbon emissions generated in the VPP scheduling is less than the given free quota, the manufacturer can sell the remaining quota on the carbon trading market and obtain the corresponding income according to the current carbon trading price (Zhang et al., 2020).

Unlike the traditional carbon trading approach, a stepped carbon trading mechanism further motivates power producers to reduce emissions by linearizing the carbon price in segments. The cost of carbon emissions exceeding a certain amount of carbon quotas will be penalized according to the amount exceeded, thus prompting them to adjust their power generation strategies or introduce low-carbon technologies. The stepped carbon trading mentioned in this study principally comprises three parts: initial carbon emission quota, actual carbon emissions of operators, and carbon trading costs.

2.3.1 Initial carbon emission quota model

Generally speaking, the electricity purchased by a VPP from the external network comes from thermal units (Guotao et al., 2021). This makes two major sources of carbon emissions in the operation of VPP: MT and the electricity purchased from the external network. At present, China's carbon trading policy is mainly based on the free quota of the actual power generation of the manufacturer, so the determination by the regulatory authority of the carbon emission quota of the VPP operator by two sources is expressed thus:

$$E_F = \delta_3 \sum_{t=1}^{T} \left(P_t^{PV} \Delta t + P_t^{WT} \Delta t + P_t^{grids} \Delta t + P_t^{MT} \Delta t \right). \tag{1}$$

Here, T is the dispatch period; P_t^{PV}, P_t^{WT}, and P_t^{MT} are the power output of the corresponding unit at time t; E_F is the free carbon emission allowances for VPP operators set by the regulator; δ_3 is the carbon emission allowances of the production of per unit electricity; and P_t^{grids} is the purchased power by the VPP from the main network at time t.

2.3.2 Actual carbon emissions model

According to the power network carbon emission calculation method, the actual carbon emissions generated by a VPP during the dispatch process are determined as

$$E_A = \sum_{t=1}^{T} \delta_1 P_{k,t}^{grids} + \sum_{t=1}^{T} \delta_2 P_{k,t}^{MT}. \tag{2}$$

Here, E_A is the amout of actual carbon emission by the VPP; δ_1 and δ_2 are the carbon emission factor of per unit active output of thermal power units and MT, respectively.

2.3.3 Stepped carbon trading cost model

Compared with the traditional unified pricing mechanism, this paper adopts a stepped carbon trading pricing model to achieve further control carbon emission reduction. Under this mechanism, carbon emissions are divided into multiple intervals (Liu et al., 2022). For intervals with higher carbon emissions, the unit carbon emission price and the carbon compensation cost will be higher. The specific calculation formula is as follows:

$$C_T = \begin{cases} \mu (E_A - E_F), E_A \le E_F + l \\ \mu l + (1+\alpha)\mu (E_A - E_F - l), E_F + l \le E_A \le E_F + 2l \\ (2+\alpha)\mu l + (1+2\alpha)\mu (E_A - E_F - 2l), E_F + 2l \le E_A \le E_F + 3l \\ (3+3\alpha)\mu l + (1+3\alpha)\mu (E_A - E_F - 3l), E_F + 3l \le E_A \le E_F + 4l \\ (4+6\alpha)\mu l + (1+4\alpha)\mu (E_A - E_F - 4l), E_A \ge E_F + 4l. \end{cases} \tag{3}$$

Here, C_T is the carbon trading cost to be paid by VPPs; μ is the market benchmark price of carbon emissions; l is the interval length of each carbon emission interval; and α is the growth rate of the stepped carbon trading price.

2.4 Objective function

Costs include two items: VPP operating costs f_k and carbon trading costs C_T.

$$F_k = \min (f_k + C_T), \tag{4}$$

$$f_k = C_k^{WT} + C_k^{PV} + C_k^{ESS} + C_k^{ld} + C_k^{MT} + C_k^{VPP} + C_k^{grid}. \tag{5}$$

Here, $C_k^{WT}, C_k^{PV}, C_k^{ESS}, C_k^{MT}$, respectively, denote the operating costs of WT, PV, ESS, and MT in VPPk; C_k^{VPP}, C_k^{grid} denote the direct power trading costs between VPPk and other VPPs, and between VPPk and the main grid; and C_k^{ld} denotes the compensation cost of VPPk calling CL. Among these, the operation cost function of each unit is the same as that in the traditional way and will not be repeated.

2.5 Constraint condition

2.5.1 Power balance constraint

$$P_{k,t}^{ld} + P_{k,t}^{gridp} = P_{k,t}^{WT} + P_{k,t}^{PV} + P_{k,t}^{ESS} + P_{k,t}^{MT} + P_{k,t}^{grids}. \quad (6)$$

Here, $P_{k,t}^{ld}$ is the load after VPPk scheduling optimization at moment t; $P_{k,t}^{ESS}$ is positive when the energy storage system is charging and negative when it is discharging; and P_t^{gridp} is the power sold by VPP to the main network at time t.

2.5.2 Renewable energy output constraints

$$\begin{cases} 0 \le P_{k,t}^{WT} \le P_k^{WT,max}, \\ 0 \le P_{k,t}^{PV} \le P_k^{PV,max}. \end{cases} \quad (7)$$

Here, $P_k^{WT,max}$ and $P_k^{PV,max}$ are the upper limits of WT and PV output of VPPk, respectively.

2.5.3 Energy storage system constraints

For the ESS, its first priority is to meet the charging and discharging power without exceeding the limit:

$$P_k^{ESS,min} \le P_{k,t}^{ESS} \le P_k^{ESS,max}. \quad (8)$$

Here, $P_k^{ESS,min}$ and $P_k^{ESS,max}$ are the lower limit and upper limits of ESS power of VPPk, respectively.

The internal storage power of ESSk is

$$SOC_{k,t} = SOC_{k,t-1} + \omega_c (1-b) P_{k,t}^{ESS} \Delta t + \frac{1}{\omega_d} b P_{k,t}^{ESS} \Delta t. \quad (9)$$

At the same time, its state of charge must also meet certain constraints:

$$\begin{cases} SOC_{k,0} = SOC_{k,T} \\ SOC_k^{\ min} < SOC_{k,t}^{ESS} \le SOC_k^{max}. \end{cases} \quad (10)$$

Here, ω_c and ω_d denote the charging and discharging efficiency of the ESS, respectively; b denotes charging and discharging state: 0 when charging, 1 when discharging; and $SOC_{k,0}$ *and* $SOC_{k,T}$ denote the initial and termination power of the ESS in one dispatching cycle, respectively.

2.5.4 Interaction power constraint

This equation includes direct interactive power constraints between VPPs and interactive power constraints between VPPs and the grid:

$$\begin{cases} -\mathrm{P}_{\mathrm{k}}^{\mathrm{VPP,max}} \le \mathrm{P}_{\mathrm{k,t}}^{\mathrm{VPP}} \le \mathrm{P}_{\mathrm{k}}^{\mathrm{VPP,max}} \\ 0 \le \mathrm{P}_{\mathrm{k,t}}^{\mathrm{gridp}} \le \mathrm{P}_{\mathrm{max}}^{\mathrm{grid}} \\ 0 \le \mathrm{P}_{\mathrm{k,t}}^{\mathrm{grids}} \le \mathrm{P}_{\mathrm{max}}^{\mathrm{grids}}. \end{cases} \quad (11)$$

2.5.5 MT output and climbing constraint

$$0 \le e_{MT}^t \le P_{MT}^{max}, \quad (12)$$

$$P_{MT}^{down} \le \left| e_{MT}^t - e_{MT}^{t-1} \right| \le P_{MT}^{up}. \quad (13)$$

Here, P_{MT}^{max}, P_{MT}^{min} are the upper and lower limits of MT unit output, respectively; P_{MT}^{down}, P_{MT}^{down} are the downward and upward climbing power of the MT unit, respectively.

3 Multi-game model for multi-VPPs

Based on the optimal results of each VPP in the first stage, the supply–demand ratio (SDR) coefficient is introduced to distinguish between power-sale VPP and power-purchase VPP.

$$SDR_{k,t} = \frac{P_{k,t}^S}{P_{k,t}^D}, \quad (14)$$

$$P_{k,t}^S = P_{k,t}^G + P_{k,t}^{ESS} + P_{k,t}^{MT}. \quad (15)$$

Here, $P_{k,t}^S$ is the electricity supply of VPPk in time period t; $P_{k,t}^D$ is the whole electricity demand of VPPk in time period t; and $P_{k,t}^G$ is the renewable energy output of VPPk in time period t; during period t: if $SDR_{k,t} < 1$, define the VPPk as power-purchase VPP, if $SDR_{k,t} > 1$, define the VPPk as power-sale VPP.

3.1 VPP utility function

3.1.1 Power-sale VPP utility function

The power-sale VPP obtains economic benefits by selling excess power to the load side and the power-purchase VPP, whose utility function can be expressed as

$$U_{j,t}^s = \alpha_{j,t} P_{j,t}^{ld} - \frac{\beta_t}{2} \left(P_{j,t}^{ld} \right)^2 + \gamma_{j,t} P_{j,t}^{VPP}. \quad (16)$$

Here, $\alpha_{j,t}$ is the satisfaction coefficient of the demand side of the power-sale VPP; $\gamma_{j,t}$ is the power trading strategy of the power-sale VPP—electricity price in t period; and $P_{j,t}^{VPP}$ is the interactive electricity of VPPj participating in VPP direct trading.

3.1.2 Power-purchase VPP utility function

The internal power supply of the power-purchase VPP is insufficient to meet its own internal demand, so it needs to purchase the corresponding demand from the power-sale VPP. The utility function of the power-purchase VPPi choosing the power-sale VPPj to purchase power at time t can be expressed as

$$U_{ij,t}^s = \alpha_{i,t} P_{i,t}^{ld} - \frac{\beta_t}{2} \left(P_{i,t}^{ld} \right)^2 - \gamma_{j,t} P_{i,t}^{VPP}. \quad (17)$$

Here, $\alpha_{i,t}$ is the satisfaction coefficient on the demand side of the power-purchase VPP.

3.2 The evolutionary game between power-purchase VPP

After comprehensive evaluation of its electricity price information, the power-purchase VPP selects the best source for electricity trading. Power-purchase VPPs must constantly adjust their selection strategies to form an evolutionary game, which takes the following form.

The selection strategy of a power-purchase VPP at the time of power purchase can be expressed by the probability $r_{j,t}\ (0 \le r_{j,t} \le 1, \sum_{j=1}^{J} r_{j,t} = 1)$ of power-purchase VPPi choosing power-sale VPPj to purchase power at time t. The ultimate goal is to maximize the utility function of power-purchase VPPi.

In the evolution process, a correction factor $v_{m,j}(U_{i,t}^{P})$ is introduced to denote the proportion of power-purchase VPPi shifted from the selection of power-sale VPPm to power-sale VPPj in time period t. The modified protocol of the dynamic evolutionary game can be formulated as

$$v_{m,j}\left(U_{i,t}^{P}\right) = \frac{exp\left(U_{i,j,t}^{P}\right)}{\sum_{m=1}^{J} exp\left(U_{i,m,t}^{P}\right)}. \tag{18}$$

The dynamic evolution equation of power-purchase VPP is

$$\frac{\partial r_{j,t}}{\partial t} = \frac{exp\left(U_{i,j,t}^{P}\right)}{\sum_{m=1}^{J} exp\left(U_{i,m,t}^{P}\right)} - r_{j,t} = v_{m,j}\left(U_{i,t}^{P}\right) - r_{j,t}. \tag{19}$$

The power-purchase VPP group reaches evolutionary equilibrium when $\frac{\partial r_{j,t}}{\partial t} = 0$; that is, $v_{m,j}(U_{i,t}^{P}) - r_{j,t} = 0$.

3.3 The non-cooperative game among power-sale VPPs

When the trading between the power-sale VPP and the power-purchasing VPP begins, the utility of each power-sale VPP depends not only on its own reported trading strategy and the power-purchasing VPP's selection strategy but also on the published electricity prices of other power-sale VPPs—the process of publishing electricity selling prices by the power-sale VPP can be described by the non-cooperative game model.

The power-sale VPP chooses the power sales price $\gamma_{j,t}$ at time t as its game strategy when playing the power sales game, with the final goal of optimizing the utility function of the power-sale VPPj. From Eq. 17, the optimal purchased electricity for power-purchase VPPi in time period t is

$$P_{ij,t} = \frac{\alpha_{i,t} - \gamma_{i,t}}{\beta_t}. \tag{20}$$

The total amount of electricity purchased by power-purchase VPPi from power-sale VPPj in time period t is

$$Q_{j,t} = \sum_{i=1}^{I} r_{j,t} P_{ij,t}. \tag{21}$$

Therefore, the utility function of the power-sale VPP can be expressed as

$$U_{j,t}^{s} = \begin{cases} \alpha_{i,t} P_{j,t}^{sld} - \frac{\beta_t}{2}\left(P_{j,t}^{sld}\right)^2 + \gamma_{j,t} P_{ij,t}, P_{ij,t} \le Q_{j,t}, \\ \alpha_{i,t} P_{j,t}^{sld} - \frac{\beta_t}{2}\left(P_{j,t}^{sld}\right)^2 + \gamma_{j,t} r_{j,t} \sum_{i=1}^{I} \frac{\alpha_{i,t} - \gamma_{i,t}}{\beta_t}, P_{ij,t} > Q_{j,t}. \end{cases} \tag{22}$$

In the process of the game, each VPP changes its pricing strategy according to the aforementioned utility function until the utility is optimal; the game reaches equilibrium when all power-sale VPPs cannot benefit from changing their strategies.

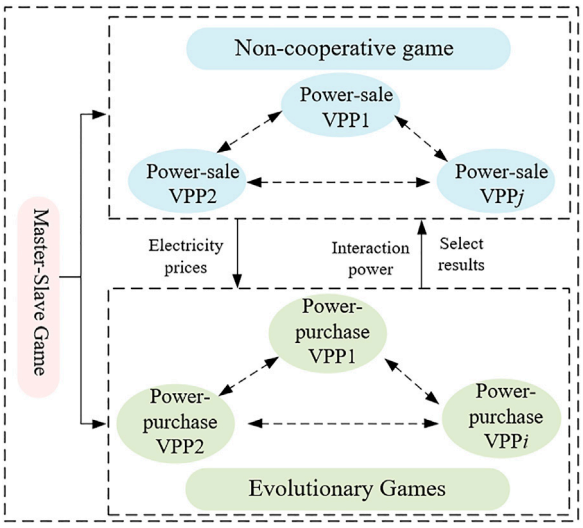


FIGURE 3
Game structure diagram of multi-VPP system.

3.4 The master–slave game between power-sale VPP and power-purchase VPP

Figure 3 shows the overall game framework for multi-VPPs. The power-sale VPP changes its pricing according to the market situation, so as to guide the power-purchase VPP to make trading choices and feed them back to the power-sale VPP, which updates its pricing according to its latest utility function and the power-purchasing VPP's strategy, thus ultimately maximizing the benefits of both parties. In this case, there is a certain sequence of participants in the electricity market when making corresponding decisions, and this behavior can be described by a follower–leader game model, with the power-sale VPP as the leader and the power-purchase VPP as the follower.

Both the power-purchase VPP and the power-sale VPP need to participate in the follower–leader game, and the strategy sets of both are still the power sale price of the power sale-VPP and the power purchase choice state of the power-purchase VPP. The objective of the game is still to optimize the utility function of both players.

The strategy set of the leader power-sale VPP can be expressed as

$$\gamma_j = \left[\gamma_{j,1}, \gamma_{j,2}, \ldots, \gamma_{j,t}, \ldots, \gamma_{j,T}\right]. \tag{23}$$

Here, γ_j is the vector of electricity price strategies published by the power-sale VPPj.

The strategy set of the follower power-purchase VPP can be expressed as

$$r_i = \left[r_{j,1}, r_{j,2}, \ldots, r_{j,t}, \ldots, r_{j,T}\right]. \tag{24}$$

Here, r_i is the probability vector that power-purchase VPPi chooses power-sale VPPj.

The objective of the master–slave game model is to optimize the utility of both the power-sale VPP and the power-purchase VPP. After the power-sale VPP releases selling price information γ_j, the power-purchase VPP can find an equilibrium strategy to optimize its

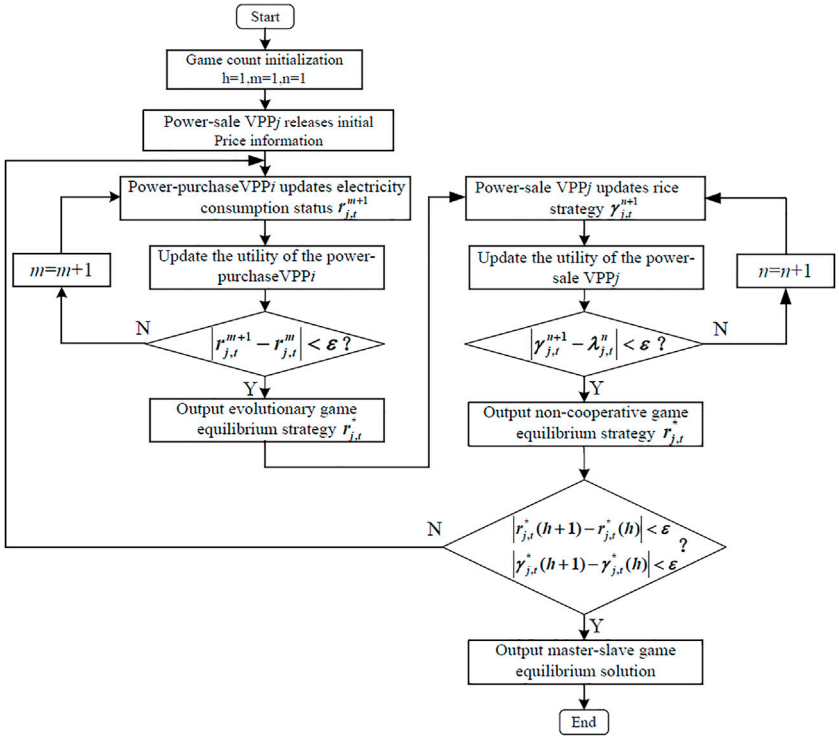


FIGURE 4
Flowchart of multiple game solution.

own utility, which is expressed by r_i^*, and feed it back to the power-sale VPP. The power-sale VPP determines its optimal electricity price strategy set γ_j^* according to its own utility function and the optimal response strategy of the power-purchase VPP, so the Nash equilibrium solution of this master–slave game can be expressed by (γ_j^*, r_i^*).

4 Multiple game solving process

For the proposed multiple game model, this paper adopts a parallel distributed solution method; the solution process is shown in Figure 4.

Step 1: As the leader of the master–slave game, the power-sale VPP first transmits initial price information to the follower power-purchase VPP.

Step 2: The power-purchase VPP makes decisions on its own power purchasing selection status according to the leader's strategy and continuously updates it until the evolutionary game reaches equilibrium to obtain its optimal utility. The evolutionary game equilibrium solution is solved as follows:

1) Power-purchase VPPi selects a power-sale VPP according to the leader's strategy and calculates the optimal power purchase and power consumption utility of the power-purchase VPP according to Eq. 20 and Eq. 17, respectively.
2) Discretize Eq. 19 as follows:

$$r_{j,t}^{m+1} = r_{j,t}^{m} + \tau_1 \left[v_{mj}\left(U_{i,t}^{P}\right) - r_{j,t}^{m} \right]. \tag{25}$$

Here, m is the number of iterations of the evolutionary game; τ_1 is the iteration step size.

The power-purchase VPP updates the power purchase status by Eq. 25;

3) Judge whether the power purchase strategy reaches the evolutionary equilibrium. If it reaches equilibrium state, go to Step 3, otherwise return to 1); Step 3: According to the results of Step 2, each power-sale VPP obtains the optimal power selling price through a non-cooperative game. The non-cooperative game solving process is as follows:

1) Calculate the total amount of power purchased by power-purchase VPPi from power-sale VPPj and the own utility of power-sale VPPj in time period t according to Eqs 21, 22.
2) The power-sale VPPj updates its own price strategy through Eq. 26.

$$\gamma_{j,t}^{n+1} = \gamma_{j,t}^{n} + \tau_2\left(Q_{j,t}^{n} - P_{ij,t}\right). \tag{26}$$

Here, n is the number of iterations of the non-cooperative game; τ_2 is the iteration step size.

3) Judge whether the power-sale VPP strategy reaches equilibrium. If it reaches equilibrium state, go to Step 4, otherwise return to 1); Step 4: Judge whether the system master–slave game reaches

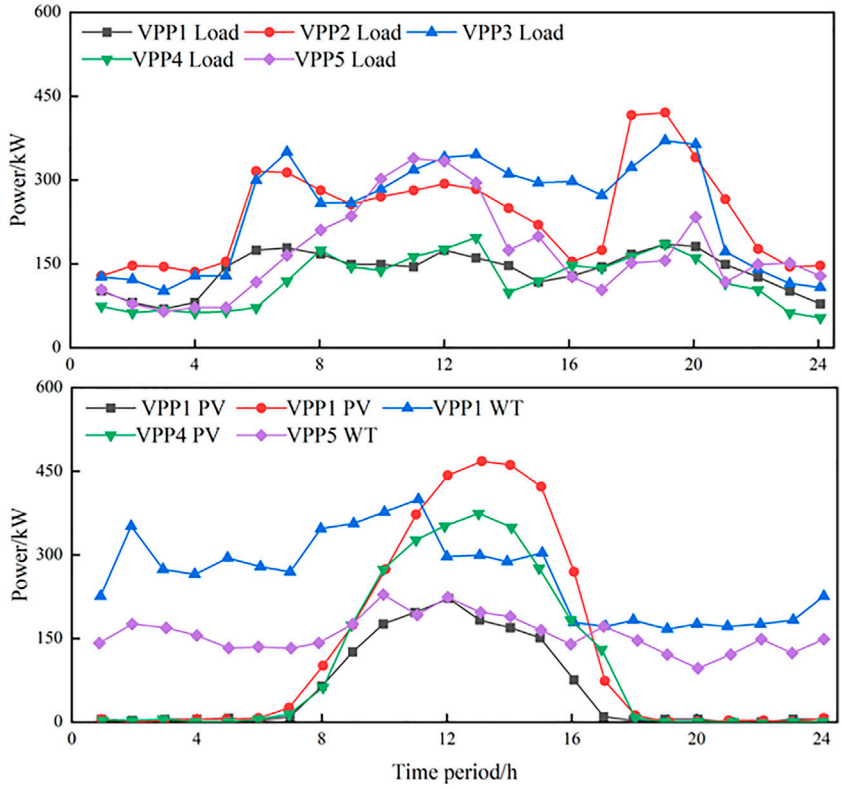


FIGURE 5
Forecasting curves of renewable energy output and load demand.

TABLE 1 Electricity purchase and sale price from VPP to main network.

Trading form	Time period	Price (yuan/kW)
Electricity purchase	9:00–11:00; 18:00–22:00	1.289
	7:00–8:00; 12:00–17:00	1.289
	1:00–6:00; 23:00–24:00	1.289
Electricity sale	1:00–24:00	1.289

equilibrium. If so, stop the iterative solution and output the result, otherwise return to Step 2.

5 Case studies

5.1 Basic data

In this paper, we established a multi-VPPs system consisting of five VPPs. The total output of WT and PV in each VPP, load situation, and the unit purchase price of electricity from the VPP to the main network is shown in Figure 5. The price range of VPP electricity sales specified by the electricity sales market is [0.3,1.2]; the price of electricity trading from VPP to main grid is shown in Table 1. In the stepped carbon trading model, μ is set at 280 yuan/t, l is set at 2t, and α is set at 25%.

5.2 Model simulation results

In order to verify the effectiveness of the stepped-type carbon trading model designed in this paper in the operation of VPP, we set up four different models. Model 1 is the traditional carbon trading model in which the carbon trading range is not divided, and the CO_2 exceeding the carbon emission quota is directly settled at the initial carbon trading price. Model 2 is still the traditional carbon trading model, but the benchmark price is set at 320 yuan/t. Model 3 is a stepped carbon trading model without considering the carbon emission cost. In the objective function, only the operating costs and power purchase costs of each unit in the VPP are considered. Model 4 comprehensively considers the stepped carbon emission cost and VPP operation cost for economic and low-carbon enhancement. Taking VPP1 as an example, the operation results of these four scenarios are shown in Table 2.

Model 1 has the lowest total cost among the four models because it adopts the traditional carbon trading mechanism with a uniform benchmark price for carbon emissions; however, its energy saving and emission reduction potential is not fully exploited, and there is a certain amount of new energy output within the VPP that cannot be absorbed. In Model 2, since the benchmark price is no longer used for carbon emissions, the total cost of the VPP increases significantly. Although the unit operation cost of Model 3 is the lowest of the four models, huge amounts of CO_2 are emitted because carbon transaction costs are not considered in the scheduling process which require VPP operators to pay high carbon emission costs. After Model 4 divides the carbon emission range,

TABLE 2 Operation results of these four scenarios of VPP1.

Model	Carbon emission/kg	Carbon trading cost/yuan	Operation cost/yuan	Total/yuan
1	35,139.1	1,998.9	24,834.3	26,833.2
2	32,587.2	1,465.6	26,273.9	27,739.5
3	36,832.6	4,106.3	24,468.1	28,574.4
4	33,932.4	2,071.6	25,324.5	27,396.1

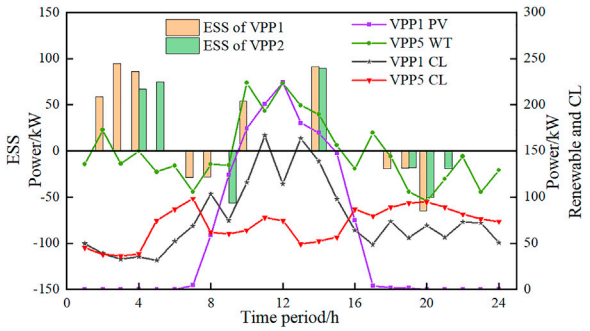


FIGURE 6
Day-ahead dispatching results.

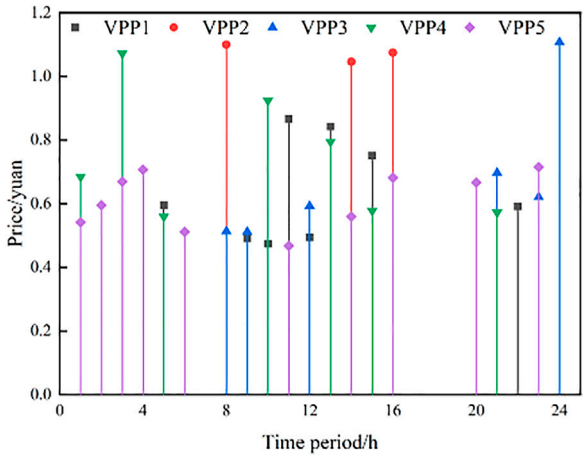


FIGURE 7
Electricity sale price of power-sale VPPs in *t* period.

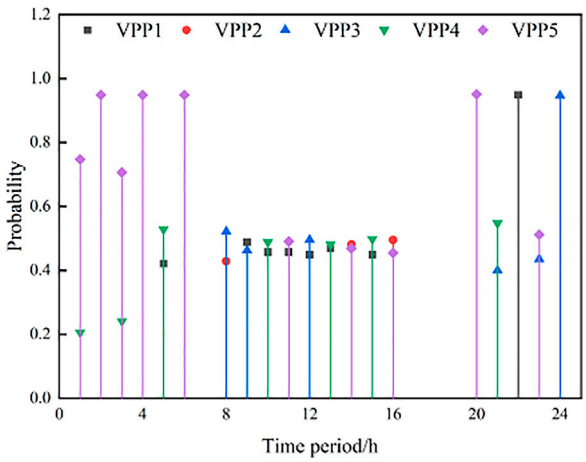


FIGURE 8
Probability of power-sale VPPs to be chosen to purchase electricity in *t* period.

the carbon emission of VPP is significantly reduced compared with the traditional model.

According to the operational comparison of models 1 and 4 in Table 2, the carbon emission cost increases after adopting the stepped carbon trading scheduling strategy, but the carbon emission decreases by 1,206.7 kg. This is because the carbon emission interval is strictly divided after adopting the stepped carbon trading model. With the gradual increase in carbon emissions, the carbon emission price faced by VPP operators will increase exponentially, greatly limiting the call of carbon sources. Although the total cost is slightly higher for VPP operators, it promotes the absorption of internal "abandoned wind and light" and generally achieves the goal of low-carbon economic operation.

We can take VPP1 (PV) and VPP5 (WT) as examples to draw the day-ahead resource scheduling curve as shown in Figure 6. From the internal supply and demand relationship of each VPP in each period, the required external interactive power can be obtained, thus introducing the trading strategy proposed in this paper.

In Figure 6, the PV output of VPP1 shows a peak trend during the time of high light intensity (7–17), and the demand response call is less then; VPP5, which is dominated by wind power output, charges during the time of low load demand and peak wind power output at night, and its energy storage system charges for three periods in 24 h per day. Thus, the VPP energy storage system, while meeting its own load demand through charging, sells excess power to other VPPs to promote the balance between the supply and demand of the system and improve the system's overall economic effectiveness.

In each dispatching time period, the five VPPs are divided into power-sale VPP and power-purchase VPP according to the SDR of each VPP. When the multiple games reach the equilibrium state,

TABLE 3 Resource allocation of each VPP in t = 10 period.

Unit	VPP	Power/kW
Renewable energy	VPP1	169.3
	VPP2	38.8
	VPP3	340.4
	VPP4	256.0
	VPP5	191.5
Controllable loads	VPP1	59.6
	VPP2	33.1
	VPP3	17.5
	VPP4	64.7
	VPP5	109.3
Load	VPP1	91.8
	VPP2	15.5
	VPP3	215.8
	VPP4	31.1
	VPP5	146.0

each VPP will decide different power sales prices for each time period, and the probability that the power-purchase VPP selects the power-sale VPP is also different. Figures 7, 8 show the sale prices of the power sales VPPs at time t and the probability distribution of each power-sale VPP being selected by the power-purchase VPPs, respectively.

Taking time periods $t = 2$, $t = 12$, and $t = 18$ as examples, it can be seen from Figure 7 that only VPP5 is a power-sale VPP in time period 2, and VPP5 delivers power to the remaining four VPPs while satisfying its own load. The power-sale VPPs in time period 12 are VPP1 and VPP3; at $t = 18$, all five VPPs are in the state of power purchase. During this period, there is no power interaction between VPPs, and only power is purchased from the main network. It can be seen that, in 24 time periods of a day, each VPP can present different states, and its internal available power can be consumed by interacting with other VPPs when there is a surplus state.

From Figures 7, 8, it can be seen that the power-purchase VPPs prefer a power-sale VPP with low power selling prices, but it can be observed that the probability of the power-purchase VPPs choosing the power-sale VPP with the lowest price at $t = 10$, $t = 12$, $t = 14$, and $t = 16$ is not the highest.

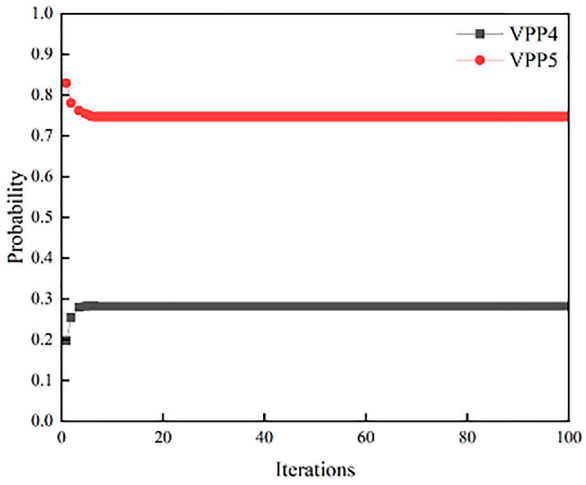


FIGURE 9
Probability convergence process of selecting power-sale VPP to purchase electricity.

Taking $t = 10$ as an example, it can be seen from Tables 3, 4 that the optimal electricity price issued by VPP1 is lower than that of VPP4, but the probability of the power-purchase VPPs choosing VPP1 is 0.4819, which is lower than the probability of choosing VPP4 by 0.5181. It can be seen that, during the dispatching process of the system, the electricity price is not the only factor to be considered by the power-purchase VPP when choosing the power sales VPP but also the influence of the controllable load call and its own load demand on the strategy.

When $t = 3$, the convergence process of the probability of each power-sale VPP being selected when the power-purchase VPP purchases electricity is shown in Figure 9. It can be seen that the probability of each power-sale VPP being selected can converge to the equilibrium value quickly; that is, the dynamic selection behavior of the power-purchase VPP can reach the evolutionary equilibrium state in a short time.

In order to verify the rationality of the proposed multiple game trading strategy for multi-VPPs considering carbon trading model, two different operation scenarios were set to analyze the multi-VPPs trading cost results.

Scenario 1: the multi-VPPs trading model with carbon trading market described in this paper, while the trading between multi-VPPs is based on the traditional method.

Scenario 2: the multi-VPPs trading model with stepped carbon trading market described in this paper and multiple game strategies among multi-VPPs are constructed.

TABLE 4 Selection strategy of power-purchase VPP in t = 10 period.

Power-purchase VPP	Power-sale VPP with the lowest price	Power-purchase VPP selection probability	
		VPP1	VPP4
VPP2 VPP3 VPP5	VPP1	0.4819	0.5181

TABLE 5 VPP scheduling cost in different scenarios.

Scenario	VPP	Carbon trading costs/yuan	VPP operating costs/yuan
1	VPP1	-2.0716×10^3	2.5324×10^4
	VPP2	-3.5010×10^3	4.6735×10^4
	VPP3	-6.6369×10^3	3.0943×10^4
	VPP4	-2.8630×10^3	2.8455×10^4
	VPP5	-3.8412×10^3	1.8950×10^4
	Total	-1.9914×10^4	1.5041×10^5
2	VPP1	-2.6873×10^3	1.8513×10^4
	VPP2	-3.8575×10^3	3.8834×10^4
	VPP3	-7.4680×10^3	1.0345×10^4
	VPP4	-3.3507×10^3	2.5275×10^4
	VPP5	-4.9848×10^3	8.7137×10^3
	Total	-2.2348×10^4	1.0168×10^5

By comparing Scenario 2 (considering multiple game strategy) with Scenario 1 (conventional game strategy) in Table 5, the total cost of carbon trading and multi-VPP system operation are reduced by 2,434 yuan and 4.873×10^4 yuan, respectively. Therefore, if there is no electric energy interaction among the VPPs, as soon as the power in the VPP cannot meet its own internal electricity demand, only purchasing power from the main network increases the operating cost of VPP system, which makes the main network face the huge pressure of electricity supply. this shows that, when the resource allocation among the subjects in the multi-VPPs is carried out through the game idea, the resource utilization rate of VPP can be improved, and the carbon transaction cost of the system can be reduced to promote the economic operation of the system.

6 Conclusion

In this paper, we analyzed a multi-VPPs unified trading strategy considering integrated carbon–electricity trading and proposed a new multiple game strategy among multi-VPPs, in which a non-cooperative game between power-sale VPPs is prioritized to determine the power sales price, and then an evolutionary game is played among power-purchase VPPs to make the choice of power purchase objects according to the power price strategy at this time. This will reduce pollution levels from VPP operations while increasing the benefits of VPP operations. After the corresponding case analysis, the following results are obtained.

1) Under the carbon trading mechanism, the multi-VPP system trading model with wind and photovoltaic is considered, which effectively promotes the consumption of renewable energy and achieves the goal of system economic operation.
2) From the perspective of multiple power-sale VPPs and multiple power-purchase VPPs, a multiple game trading strategy is adopted to improve the safe and stable operation of the system, which in turn has a positive effect on the coordinated distribution of benefits among the various entities in the system.
3) For multi-VPPs unified trading in multiple regions, a parallel distributed solution strategy is chosen in this paper, which can greatly accelerate the solution speed and avoid large amounts of computation.

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

TC conceived the idea for the manuscript and XA, YZ, XW, and LD wrote it. All authors have read and agreed to its published version.

Funding

This work is supported by the Science and Technology Project of Electric Power Research Institute of State Grid Liaoning Electric Power Supply Co., Ltd.—Research on Power Transaction and Multi regional Energy Coordination Optimization Technology of Virtual Power Plants Based on Blockchain (2022YF-78).

Conflict of interest

Authors TC, YZ and XW were employed by the company State Grid Liaoning Electric Power Co., Ltd.

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.

The reviewer editor BZ declared a shared affiliation with authors TC, WA, and LD at the time of the review.

The authors declare that this study received funding from the Science and Technology Project of Electric Power Research Institute of State Grid Liaoning Electric Power Supply Co., Ltd. The funder had the following involvement in the study: Research on Power Transaction and Multi regional Energy Coordination Optimization Technology of Virtual Power Plants Based on Blockchain (2022YF-78).

Publisher's note

References

Chen, J., Hu, Z., Chen, J., Chen, Y., Gao, M., and Lin, M. (2021). Optimal scheduling of integrated energy systems considering stepped carbon trading and flexible dual response of supply and demand. *High. Volt. Tech.* 47 (09), 3094–3106. doi:10.13336/j.1003-6520.hve.20211094

Chu, T., An, X., Zhang, W., Lu, Y., and Tian, J. (2023). Multiple virtual power plants transaction matching strategy based on alliance blockchain. *Sustainability* 15 (8), 6939. doi:10.3390/su15086939

Dou, Yangming, Ming, Chi, Liu, Zhi-Wei, Wen, Guanghui, and Sun, Qihai (2022). Distributed secondary control for voltage regulation and optimal power sharing in DC microgrids. *IEEE Trans. Control Syst. Technol.* 30, 2561–2572. doi:10.1109/TCST.2022.3156391

Feng, C., Hua, H., Yong, W., Jianfeng, H., and Mian, G. (2018). A virtual power plant bidding model considering uncertainty. *Appl. Sci. Technol.* 45 (05), 62–66.

Guotao, Z., Quan, D., Junhua, F., Guoming, Q., Ying, C., and Chao, H. (2021). Research on the low-carbon path of regional energy systems based on multi market linkage. *Electr. Power Constr.* 42 (03), 19–26.

Huang, B., Zheng, S., Wang, R., Wang, H., Xiao, J., and Wang, P. (2022). Distributed optimal control of DC microgrid considering balance of charge state. *IEEE Trans. Energy Convers.* 37 (3), 2162–2174. doi:10.1109/TEC.2022.3169462

Huang, Bonan, Liu, Lining, Zhang, Huaguang, Li, Yushuai, and Sun, Qiuye (2019). Distributed optimal economic dispatch for microgrids considering communication delays. *IEEE Trans. Syst. Man, Cybern. Syst.* 49 (8), 1634–1642. doi:10.1109/tsmc.2019.2900722

Li, Y., Gao, D. W., Gao, W., Zhang, H., and Zhou, J. (2020). Double-mode energy management for multi-energy system via distributed dynamic event-triggered Newton-raphson algorithm. *IEEE Trans. Smart Grid* 11 (6), 5339–5356. doi:10.1109/tsg.2020.3005179

Liu, Xiaoou (2022). Research on optimal dispatch method of virtual power plant considering various energy complementary and energy low carbonization. *Int. J. Electr. Power and Energy Syst.* 136, 107670. doi:10.1016/j.ijepes.2021.107670

Liu, Z., Huang, B., Hu, X., Du, P., and Sun, Q. (2023). Blockchain-based renewable energy trading using information entropy theory. *IEEE Trans. Netw. Sci. Eng.*, 1–12. doi:10.1109/TNSE.2023.3238110

Liu, Z., Liu, R., Liang, N., and Liu, X. (2021). A multi microgrid energy trading method based on game theory. *Power Grid. Technol.* 45 (02), 587–595. doi:10.13335/j.1000-3673.pst.2020.0175

Liu, Z., Xu, Y., Zhang, C., Elahi, H., and Zhou, X. (2022). A blockchain-based trustworthy collaborative power trading scheme for 5G-enabled social internet of vehicles. *Digital Commun. Netw.* 8 (6), 976–983. doi:10.1016/j.dcan.2022.10.014

Liu, Zhi-Wei, Shi, Yu-Lun, Yan, Huaicheng, Han, Bing-Xin, and Guan, Zhi-Hong. Secure consensus of multi-agent systems via impulsive control subject to deception attacks. *IEEE Trans. Circuits Syst. II Express Briefs* 170, 166–170. doi:10.1109/TCSII.2022.3196042

Shayegan-Rad, A., Badri, A., and Zanganeh, A. (2017). Day-ahead scheduling of virtual power plant in joint energy and regulation reserve markets under uncertainties. *Energy* 121 (2), 114–125. doi:10.1016/j.energy.2017.01.006

Wang, R., Li, W., Sun, Q., Li, Y., Gui, Y., and Wang, P. (2022). Fully distributed dynamic edge-event-triggered current sharing control strategy for multi-bus DC microgrids with power coupling. *IEEE Trans. Industrial Inf.* 19, 5667–5678. doi:10.1109/TII.2022.3188352

Wang, R., Ma, D., Li, M. -J., Sun, Q., Zhang, H., and Wang, P. (2022). Accurate current sharing and voltage regulation in hybrid wind/solar systems: An adaptive dynamic programming approach. *IEEE Trans. Consumer Electron.* 68 (3), 261–272. doi:10.1109/tce.2022.3181105

Zhang, H., Li, Y., Gao, D. W., and Zhou, J. (2017). Distributed optimal energy management for energy internet. *IEEE Trans. Industrial Inf.* 13 (6), 3081–3097. doi:10.1109/tii.2017.2714199

Zhang, X., Liu, X., and Zhong, J. (2020). Comprehensive energy system planning considering the uncertainty of reward and punishment for tiered carbon trading and electricity heat transfer load. *Proc. CSEE* 40 (19), 6132–6142. doi:10.13334/j.0258-8013.pcsee.191302

Zhao, Haotian, Wang, Bin, Wang, Xuanyuan, Pan, Z., Sun, H., Liu, Z., et al. (2020). Active dynamic aggregation model for distributed integrated energy system as virtual power plant. *J. Mod. Power Syst. Clean Energy* 8 (5), 831–840. doi:10.35833/mpce.2020.000202

Zhi, C., Zhijian, H., Changhong, W., and Tiange, L. (2021). Multi stage planning of a comprehensive energy system in a park based on a tiered carbon trading mechanism. *Power Autom. Equip.* 41 (09), 148–155. doi:10.16081/j.epae.202109012

Zhu, Y. S., Nie, C. J., Chen, B., Li, Q. Y., Yang, Z. T., and Yang, J. L. (2022). Research on multiple game cooperative scheduling of microgrid cluster systems under mixed time scales. *Power Grid Technol.*, 1–12. doi:10.13335/j.1000-3673.pst.2022.0611

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 01 June 2023
DOI 10.3389/fenrg.2023.1201271

OPEN ACCESS

EDITED BY
Xiao-Kang Liu,
Huazhong University of Science and Technology, China

REVIEWED BY
Xiaodong Yang,
Hefei University of Technology, China
Xiaochao Hou,
Tsinghua University, China
Shichang Cui,
Huazhong University of Science and Technology, China

*CORRESPONDENCE
Zhichun Yang,
cschust@126.com

RECEIVED 06 April 2023
ACCEPTED 15 May 2023
PUBLISHED 01 June 2023

CITATION
Yang Z, Yang F and Chen J (2023), Optimal power distributed control of the DC microgrid in meshed configuration. *Front. Energy Res.* 11:1201271. doi: 10.3389/fenrg.2023.1201271

Optimal power distributed control of the DC microgrid in meshed configuration

Zhichun Yang[1]*, Fan Yang[1] and Jiawen Chen[2]

[1]State Grid Hubei Electric Power Co., Ltd., Electric Power Science Research Institute, Wuhan, China, [2]State Grid Hubei Electric Power Co., Ltd., Wuhan, China

This paper proposes a Lyapunov-based power sharing control scheme and a fixed-time-based distributed optimization algorithm to achieve optimal power sharing of sources in a DC microgrid. The Lyapunov-based controller is designed based on so-called ratio consensus protocol, where it drives the sources to a desired proportional power sharing by regulating the voltage profile of the DC microgrid. The distributed optimization optimizer is established by integrating a finite-time weighted consensus algorithm with an iterative algebraic operation, where it calculates the optimal power dispatch on the target of minimizing the generation cost. The optimizer receives the current output power of the controlled DC microgrid and sends the obtained power dispatch to the power sharing controller as the proportionality coefficients. Both the controller and optimizer are carried out in a fully distributed way. Under the framework of the Lyapunov method, stability analysis of the DC microgrid with the proposed control scheme, as well as convergence and optimality analysis of the distributed optimization algorithm, is provided. However, the influence of the time delay of the controller on the system remains to be further investigated in future work.

KEYWORDS

DC microgrid, power sharing, voltage control, power allocation, optimal DC power flow

1 Introduction

The smart grid has been attracting much attention in recent years, where it integrates the traditional power grid, renewable distributed resources, and advanced control and optimization methods on the bridge of cyber-physical techniques (Liu et al., 2021; Hou et al., 2022). With the increasing penetration of renewable energy sources (RESs) and distributed generators (DGs), traditional power systems are transforming into the form of a distributed autonomous power system, namely, the microgrid (Hatziargyriou et al., 2007). In recent years, a lot of research on the DC microgrid has been emerging since it avoids the reactive power regulation and the harmonic compensation compared with the traditional AC microgrid (Olivares et al., 2014; Papadimitriou et al., 2015; Meng et al., 2017; Wang et al., 2023).

Power sharing control is one of the important control targets of microgrids (Simpson-Porco et al., 2013; Morstyn et al., 2016a; Morstyn et al., 2016b). Under the framework of the hierarchical control (Guerrero et al., 2011; Bidram and Davoudi, 2012), the current sharing problem of parallel DC microgrids has been solved by decentralized methods (Guerrero et al., 2011; Khorsandi et al., 2014; Hamzeh et al., 2015) and distributed methods (Anand et al., 2013; Lu et al., 2014; Wang et al., 2016a), respectively. Indeed, the decentralized controller has been wildly used for the practical AC/DC microgrids (Hou et al., 2019). Recently, the distributed controller for the DC microgrid has been developed and

attracts much attention (Liu et al., 2023b; Liu et al., 2023a). The decentralized methods require transmitting the voltage of a common bus to each converter, whereas the distributed methods merely require the current or voltage information on neighbors via an information network. Similarly, distributed control schemes have also been developed in the current sharing problem of the DC microgrid with meshed topology (Nasirian et al., 2014). However, current sharing guarantees the power sharing of loads but not that of sources. In addition, the existing current sharing controllers are not applicable to accurate energy management at the source side.

When considering the generation cost of power sources, achieving optimal power sharing becomes a crucial problem that can be solved through an economic dispatch model (Ahmed et al., 2023). Distributed economic dispatch optimization algorithms have been developed, taking advantage of consensus algorithms in multi-agent systems, including the ϵ-based consensus algorithm (Yang et al., 2013), distributed bisection method (Xing et al., 2015), distributed projected gradient algorithm (Guo et al., 2016), subgradient-based consensus algorithm (Wang et al., 2016b), event-triggered consensus algorithm (Li et al., 2016), and consensus-based energy management algorithm (Zhao et al., 2016). However, most existing algorithms for economic dispatch neglect the transmission loss of power lines, despite some literature studies discussing it [e.g., Kron's loss formula models in Loia and Vaccaro (2014)]. It is noted that Kron's loss formula models transmission loss, but obtaining the loss coefficients B in practice is difficult. Optimal power sharing controllers have been designed by integrating the physical system and economic dispatch model in several studies, including Hamad et al. (2016), Li et al. (2017), Moayedi and Davoudi (2017), and Hu et al. (2018). It is noted that Li et al. (2017) formulated an optimization problem but regard the power flow as a constraint, and they then used the optimized parameters in the decentralized primary controller. However, this optimization problem was solved by a centralized heuristic algorithm, which requires global information, and may become computationally expensive once the number of sources increases. As an alternative, power sharing control schemes with distributed optimization algorithms have been widely developed by interacting with neighboring sources (Hamad et al., 2016; Moayedi and Davoudi, 2017; Hu et al., 2018).

The proposed distributed method in Moayedi and Davoudi (2017) can simultaneously optimize the power sharing of sources and regulate the voltage profile, where the generation limits of sources are also guaranteed by their incremental cost consensus protocol. The method in Hu et al. (2018) is a discrete-time control protocol using the current imbalance information, where the economical regulator generates a reference current signal for each converter to achieve the optimal power sharing of sources. Supervisory control has been designed on the basis of the sensitivity analysis, where it successfully solves the equal power sharing problem (Hamad et al., 2016). Then, the distributed equal incremental cost (DEIC) algorithm is proposed to achieve the optimal power dispatch. Optimal power sharing control has been investigated in Chang et al. (2023) for a hybrid AC/DC microgrid, but it mainly focuses on the power dispatch between the AC and DC sides while ignoring the optimal power sharing of the sources at the DC side. Optimal energy consumption has been analyzed in Xiao et al. (2022) for a practical shipboard DC microgrid, where the analysis is based on the transfer function with a linear dynamic part. However, in the aforementioned literature, the stability criteria are hard to be verified because all the poles of the transfer functions or all the eigenvalues of a big matrix should be calculated and checked to ensure them within the open left-hand plane or within the unit circle at the origin. In addition, the parameter design may fail to work if the Laplacian matrix of the communication topology or the conductance matrix of the DC microgrid is unknown.

In this paper, a distributed Lyapunov-based proportional power sharing control and a distributed initial value restoration (distributed optimization) optimization algorithm are designed to achieve the optimal power sharing of sources in a meshed DC microgrid. The Lyapunov-based controller is a consensus-like scheme based on the power information on neighbors. The proposed distributed optimization algorithm consists of a finite-time weighted consensus protocol and an algebraic operation on initial value restoration. In the process of optimization operation, the optimizer receives the real-time output power information and calculates the optimal power dispatch, and then sends back the optimized power dispatch to the controller as the proportionality coefficients. Additionally, a rigorous analysis of stability, convergence, and optimality is given. Compared with existing methods, the key contributions of this paper are summarized as follows:

1. The Lyapunov-based proportional power sharing controller for a DC microgrid is designed, which does not require to know the Laplacian matrix of the communication topology and the conductance matrix of the DC microgrid, as needed in existing approaches (Hamad et al., 2016; Moayedi and Davoudi, 2017; Hu et al., 2018).
2. The proposed distributed optimization algorithm is a fully distributed algorithm. Compared with Moayedi and Davoudi (2017), our optimization method avoids transforming the information topology once a generation reaches its limits, where the change of topology may lead to the redesign of parameters. Compared with Hu et al. (2018) and Hamad et al. (2016), our optimization method can work without knowing the exact number of sources.
3. Optimal sharing control has been investigated in Dou et al. (2022), however, on the DC microgrid with the single-bus configuration. Moreover, the consensus-based secondary control is designed using the power on the load rather than the output power of the distributed generation unit (DGU). In this paper, we focus on the DC microgrid with meshed configuration and the optimal power sharing of the DGU.
4. Optimal control of the DC microgrid is discussed in Huang et al. (2022) with a rigorous theoretical analysis and considering balance of the charge state. However, the paper focuses on the optimal voltage control rather than on the power sharing control of the DC microgrid.

The organization of the remaining part is as follows. The preliminaries and problem statement are given in **Section 2**. The distributed Lyapunov-based proportional power sharing control is presented in **Section 3**, where stability analysis is given. The distributed optimization algorithm and the convergence proof are

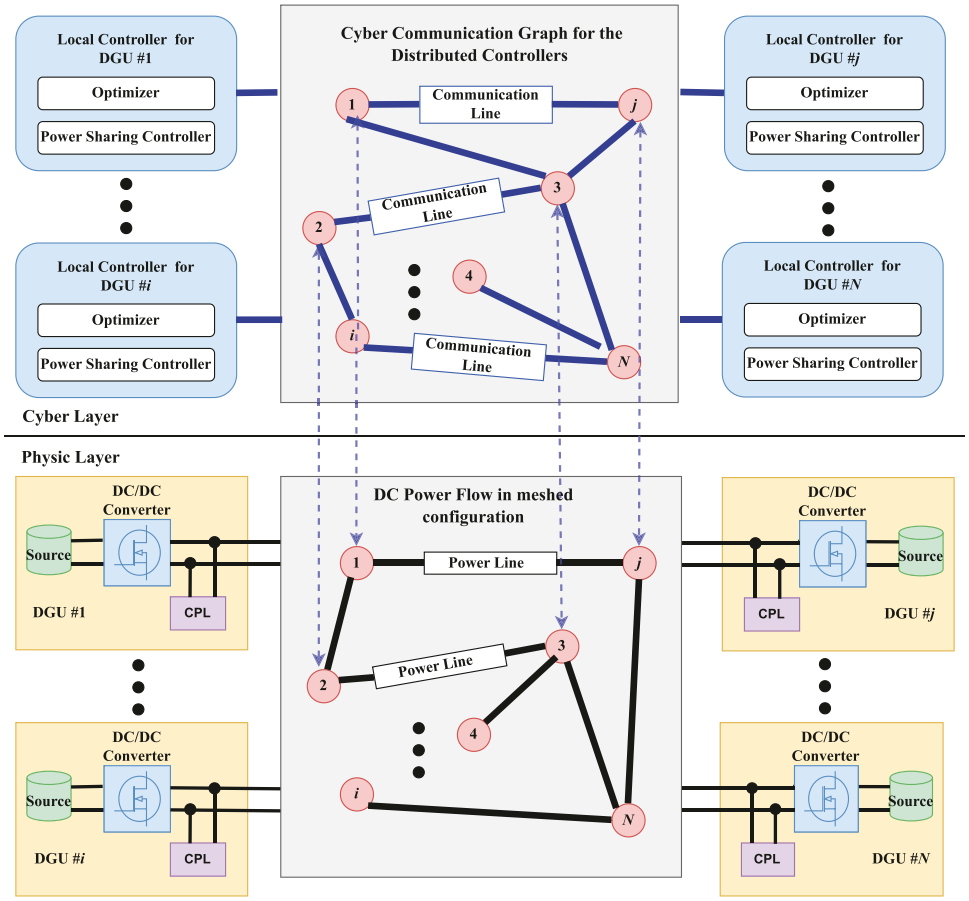


FIGURE 1
Cyber-physical framework of a DC microgrid in meshed configuration.

given in Section 4. The simulation test is given in Section 5, and the conclusion is drawn in Section 6.

2 Preliminaries

Consider a DC microgrid with N bus nodes, denoted as $\mathcal{I}_G = \{1,2,\ldots,N\}$. Define G_{ij} as the conductance of the transmission line connecting with nodes i and j, where $G_{ij} > 0$ if the i-th node and j-th node are connected via a power line, and otherwise, $G_{ij} = 0$ if they are not connected. G_{ii} is the shunt conductance of the local load. Then, the conductance matrix $Y = [Y_{ij}] \in \mathbb{R}^{N\times N}$ with $Y_{ij} = -G_{ij}$ if $i \neq j$; otherwise, $Y_{ij} = \sum_{j=1}^{N} G_{ij}$. Before describing the optimal power sharing control problem of the DC microgrid, a brief introduction of information flow is given. Denote $\mathcal{G} = (\mathcal{I}_G, \mathcal{E}, \mathcal{A})$ as the information flow of sources. The graph $\mathcal{G}$ is described with a set of nodes $\mathcal{I}_G$, a set of edges $\mathcal{E} \in \mathcal{I}_G \times \mathcal{I}_G$, and a weighted adjacency matrix $\mathcal{A} = (a_{ij})_{N\times N}$ with non-negative adjacency elements. The node i represents the ith source. Note that $a_{ij} > 0$ if and only if the ith source can obtain information from the jth source. Define $N_i = \{j | a_{ij} > 0\}$, which is the set that contains the neighbors of the node i. The Laplacian matrix $\mathcal{L} = (l_{ij})_{N\times N}$ of the graph $\mathcal{G}$ is defined as $l_{ij} = -a_{ij}$ if $i \neq j$; otherwise, $l_{ij} = \sum_{k=1,k\neq i}^{N} a_{ik}$.

The cost of each generation unit is denoted as $C_i(P_i)$, where P_i is the power generation of the i-th DG. As illustrated in Figure 1, the control and optimization framework of a DC microgrid is, in fact, a cyber-physical system which consists of a cyber system layer and a physical system layer. The physical layer is a real-time system including loads, sources, DC–DC converters, and zero-level controllers. The cyber layer is a management system that takes charge of control and optimization for the microgrid in a distributed manner through a communication network and local calculation units. In our framework, there are a power sharing controller and a generation optimizer in each source. They cooperatively calculate the reference voltages for the zero-level controllers, which directly regulate the output voltages of buses.

For the generation optimizer, it aims to calculate the optimal power dispatch of sources by solving an economic dispatch problem, which can be described as the following optimization formulations:

$$\min_{\widehat{P}_i,\ i\in I_G} \sum_{i=1}^{N} C_i\left(\widehat{P}_i\right), \tag{1}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \widehat{P}_i = P_{Demand}, \tag{2}$$

$$P_i^{\min} \leq \widehat{P}_i \leq P_i^{\max},\ i \in I_G \tag{3}$$

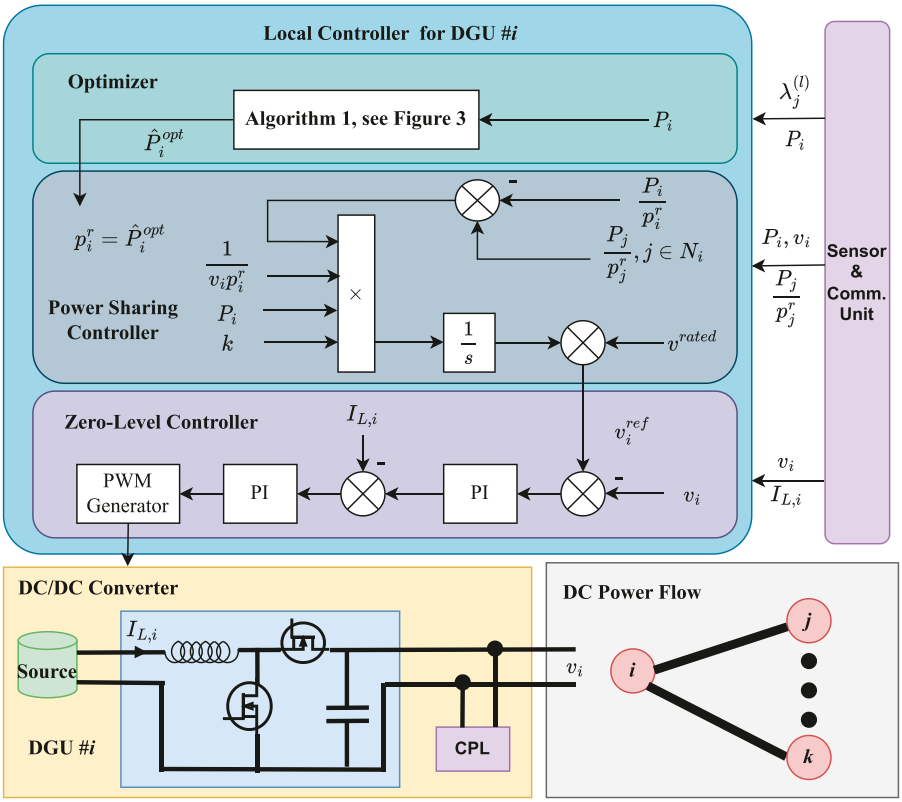


FIGURE 2
Distributed controller for each DC–DC converter.

where $C_i(\cdot)$ is the generation cost function, $\widehat{P}_i$ is the optimization variable denoted as the ith power generation, P_{Demand} is the total power demand, and $P_i^{\min}$ and $P_i^{\max}$ are the lower bound and the upper bound of the output power, respectively. The optimization result $\widehat{P}_i^{opt}$ will be further sent to the power sharing controller as a reference.

For the power sharing controller, it aims to design the reference voltage of buses V_i^{ref} for the zero-level controller such that the real-time output power could track on the optimized output power,

$$P_i = \widehat{P}_i^{opt}, \ \ for \ \ i \in \mathcal{I}_G, \tag{4}$$

where P_i is the real-time output power and $\widehat{P}_i^{opt}$ is the optimal output power dispatch generated by the generation optimizer.

Under the framework in Figure 1, the objective of optimal power sharing control is to minimize the total generation cost of the microgrid by regulating the output voltage of buses while meeting the demand and power generator constraints. To solve the optimal power sharing control of the DC microgrid, a proportional power sharing controller and a distributed optimization algorithm are designed.

3 Proportional power sharing scheme

In this section, a proportional power sharing scheme is presented to achieve the desired proportional power dispatch of sources, i.e.,

$$\frac{P_i}{p_i^r} = \frac{P_j}{p_j^r}, \ \ for \ \ \forall i,j \in \mathcal{I}_G, \tag{5}$$

where P_i is the output power and p_i^r is the proportionality coefficient.

3.1 Power sharing scheme

The power sharing control scheme for the ith source in the DC microgrid is designed as

$$v_i^{ref} = v^{rated} + \delta_i, \tag{6a}$$

$$\dot{\delta}_i = \frac{kP_i}{v_i p_i^r} \sum_{j \in N_i} a_{ij} \left(\frac{P_j}{p_j^r} - \frac{P_i}{p_i^r} \right), \tag{6b}$$

where k is a positive control parameter, v^{rated} is the nominal voltage of microgrids, and P_i is the real-time output power of the ith source. The power sharing of sources in the DC microgrid can be achieved using the active power information on neighbors, as illustrated in Figure 2. Because the dynamics of the converter is evolving in a fast time-scale, the output voltage of the source could rapidly track on the reference voltage v_i^{ref} by the zero-level controller, as shown in Figure 2. Under this circumstance, it could be assumed that $v_i = v_i^{ref}$. The following subsection gives the stability analysis of the DC microgrid under the proportional power sharing controller.

3.2 Stability analysis

Let $V = \mathrm{col}(v_1, v_2 \ldots, v_N)$, $\delta = \mathrm{col}(\delta_1, \delta_2 \ldots, \delta_N)$, and $P = \mathrm{col}(P_1, P_2 \ldots, P_N)$. Define the notation $\mathrm{diag}([b_1, \ldots, b_N]^T)$ as the diagonal matrix with diagonal elements $b_1, \ldots, b_N$; then, the compact form of the sources' dynamics is given by

$$V = V^{rated} + \delta,$$
$$\dot{\delta} = -kD^{-1}\widetilde{P}\widetilde{V}^{-1}LD^{-1}P,$$

where $V^{rated} = \mathbf{1}_N v^{rated}$, $D = \mathrm{diag}([p_1^r, p_2^r, \ldots, p_N^r]^T)$, $\widetilde{P} = \mathrm{diag}(P)$, and $\widetilde{V} = \mathrm{diag}(V)$. The dynamics of sources subject to the power flow equations

$$P = P_D + \widetilde{V}YV, \tag{7}$$

where $P_D = \mathrm{col}(P_1^d, P_2^d \ldots, P_N^d)$ and P_i^d is the demand power of local loads at bus i. Hence, the closed-loop system can be obtained as

$$\dot{V} = -kD^{-1}\widetilde{P}\widetilde{V}^{-1}LD^{-1}P, \tag{8a}$$

$$P = P_D + \widetilde{V}YV. \tag{8b}$$

Indeed, taking the derivation of both sides of Eq. 6a and substituting Eq. 6b yields Eq. 8a and Eq. 8b, which is the DC power flow for the meshed configuration of the DC microgrid. Note that Eq. 8a is a differential equation and Eq. 8b is an algebraic equation. Before giving the stability analysis of the closed-loop system, denote the equilibrium of (8) as $\mathrm{E} = (P^\star, V^\star)$, which satisfies the following equations:

$$0 = -kD^{-1}\widetilde{P}^\star \widetilde{V}^{\star -1}LD^{-1}P^\star,$$
$$P^\star = P_D + \widetilde{V}^\star YV^\star.$$

Theorem 1: *Suppose that the communication graph is connected. Consider the closed-loop system (8). The proposed distributed controller (6) ensures the following statements: i) the solution of (8) approaches the equilibrium E and ii) the power sharing* (5) *is guaranteed.*

Proof: Define the Lyapunov functional candidate as

$$\mathcal{W} = \frac{1}{2}V^TV.$$

Taking the derivative of the Lyapunov function along (8), one has

$$\begin{aligned}\dot{\mathcal{W}} &= V^T\dot{V}\\ &= -kV^T\widetilde{P}\widetilde{V}^{-1}D^{-1}LD^{-1}P\\ &= -kP^TD^{-1}LD^{-1}P\\ &\le 0.\end{aligned}$$

It is noted that $D^{-1}LD^{-1}$ is a symmetric matrix with non-negative eigenvalues because the communication graph is connected.

By LaSalle's invariant principle, the solution of (8) will approach the largest invariant set of

$$\mathcal{M} = \{(P, V)\,|\dot{\mathcal{W}} = 0\} \text{ as } t \to +\infty.$$

Indeed, the solution is also subject to the algebraic flow equation $P = P_D + \widetilde{V}YV$. Hence, it will approach the set $\mathcal{M}_s = \{(P, V)|\dot{\mathcal{W}} = 0 \text{ and } P = P_D + \widetilde{V}YV\}$. It is easy to find that $\mathrm{E} = \mathcal{M}_s$, which means the solution of (8) will approach the equilibrium E.

Note that $\dot{\mathcal{W}} = 0$ indicates $P^TD^{-1}LD^{-1}p = 0$, where it implies $LD^{-1}p = 0$. L is the Laplacian matrix, of which the row sum equals to zero. Therefore, the null space of the matrix LD^{-1} is $[p_1^r, p_2^r, \ldots, p_N^r]^T$, which indicates there exists a positive constant $p^\star$ such that $P = p^\star[p_1^r, p_2^r, \ldots, p_N^r]^T$. Obviously, it guarantees

$$\lim_{t\to+\infty}\left(\frac{P_i}{p_i^r} - \frac{P_j}{p_j^r}\right) = 0, \text{ for } \forall i, j \in \mathcal{I}_G.$$

This completes the proof of **Theorem 1**.

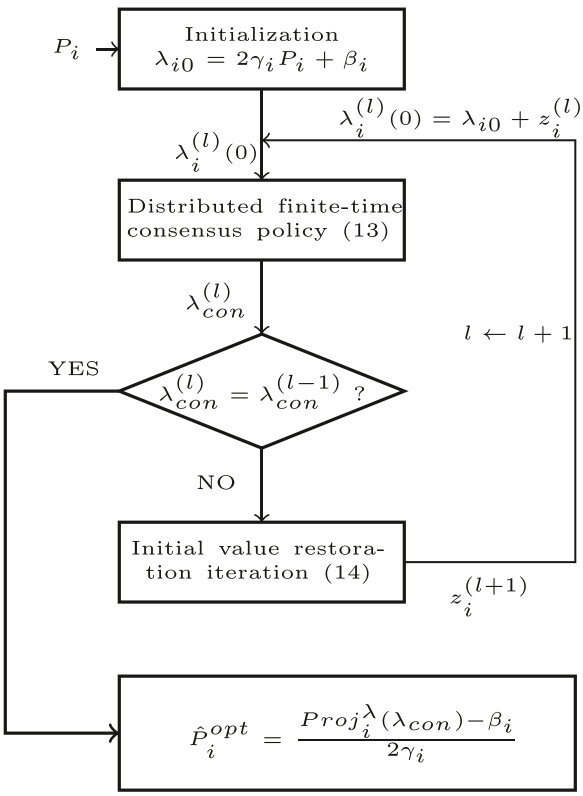


FIGURE 3
Flowchart of the distributed optimization algorithm.

Remark 1: *Set $p_i^r = \mathcal{P}_i$, where $\mathcal{P}_i$ is the desired output power of the ith source satisfying $\sum_i \mathcal{P}_i = \sum_i P_i^d + P_{loss}$ with $P_{loss} = \frac{1}{2}\sum_i\sum_j Y_{ij}(V_i^\star - V_j^\star)^2 = 1_N\widetilde{V}^\star YV^\star$, denoted as the power loss on lines. The proportional power sharing controller ensures $\frac{P_i}{\mathcal{P}_i} = \chi$, where χ is a positive constant. By the power flow Eq. 8b), one has $\sum_i P_i = \sum_i P_i^d + P_{loss}$, which indicates that $\sum_i P_i = \sum_i \mathcal{P}_i$ and $\chi = 1$. In this way, it achieves the desired power output of sources.*

Remark 2: *Based on the Lyapunov stability analysis, the parameter design of the power sharing controller is quite simple, where it only requires $k > 0$. Note that the selection of k is independent of the Laplacian matrix L and the conductance*

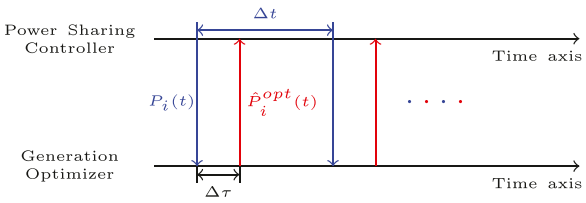


FIGURE 4
Interactive operation of the controller and the optimizer in the *i*th source.

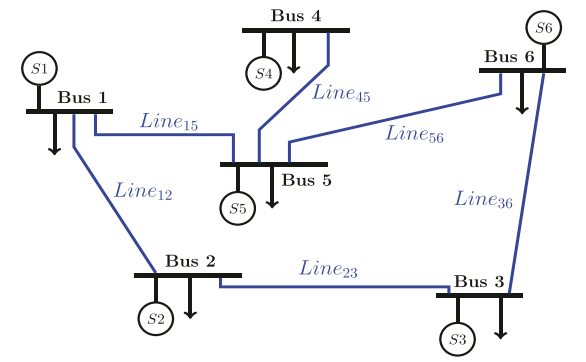


FIGURE 5
Microgrid tested with six buses.

TABLE 1 Parameters of lines and loads.

Parameter of lines and loads			
$Line_{12}$	0.15 Ω 2 mH	$Load_1$	15 kW
$Line_{23}$	0.25 Ω 2.5 mH	$Load_2$	20 kW
$Line_{36}$	0.20 Ω 3 mH	$Load_3$	15 kW
$Line_{15}$	0.15 Ω 2 mH	$Load_4$	20 kW
$Line_{45}$	0.10 Ω 1 mH	$Load_5$	15 kW
$Line_{56}$	0.20 Ω 2.5 mH	$Load_6$	20 kW

matrix Y. Moreover, the parameter k determines the speed of the convergence.

4 Optimal economic dispatch

This section aims to obtain the optimal power dispatch by solving the economic dispatch problem (Eqs 1–3).

4.1 Solution to the economic dispatch problem

Similar to the classic economic dispatch problem formulation in AC microgrids, the cost functions of dispatchable sources can be approximated as a quadratic function (Wood and Wollenberg, 2012),

$$C_i\left(\hat{P}_i\right) = \gamma_i \hat{P}_i^2 + \beta_i \hat{P}_i + \alpha_i, \tag{9}$$

where $\hat{P}_i$ is the output of the *i*th power generator, and α_i, β_i, and γ_i are the corresponded parameters.

Define the Lagrangian operator $L(\hat{P}_1, \ldots, \hat{P}_N)$ such that

$$L\left(\hat{P}_1, \ldots, \hat{P}_N\right) = \sum_{i=1}^{N} C_i\left(\hat{P}_i\right) + \lambda\left(P_D - \sum_{i=1}^{N} \hat{P}_i\right), \tag{10}$$

where λ is the Lagrange multiplier and P_D is the sum of the real-time output power transmitted from the controller. P_D includes both the load demand P_{Demand} and the transmission loss P_{loss}. For the EDP with power generation constraints, the optimal incremental cost λ^* can be obtained by verifying [Eq. (3.7) in (Wood and Wollenberg, 2012)]

$$\begin{cases} 2\gamma_i \hat{P}_i + \beta_i = \lambda^*, & \text{for } P_i^{\min} < \hat{P}_i < P_i^{\max}, \\ 2\gamma_i \hat{P}_i + \beta_i < \lambda^*, & \text{for } \hat{P}_i = P_i^{\max}, \\ 2\gamma_i \hat{P}_i + \beta_i > \lambda^*, & \text{for } \hat{P}_i = P_i^{\min}. \end{cases} \tag{11}$$

TABLE 2 Generation cost parameters.

Source	γ_i($/kW^2$)	β_i ($/kW)	α_i($)	$P_i^{\min}$ (kW)	$P_i^{\max}$ (kW)
1	0.071	2.623	68.52	5	30
2	0.091	3.143	51.81	12	20
3	0.063	2.357	38.66	5	20
4	0.087	1.715	48.47	12	45
5	0.073	2.720	53.71	8	20
6	0.067	1.934	57.50	8	45

4.2 Distributed optimization algorithm

This subsection presents a discrete-time multi-agent system that employs the local variable $\lambda_i(t)$, where $i \in \mathcal{I}_G$, to collaboratively estimate the optimal λ^* in Eq. 11. Prior to introducing the distributed optimization algorithm, it is necessary to define a projection operator that maps from $\mathbb{R} \times \Omega$ to Ω

$$\text{Proj}(x, \Omega) = \arg\min_{v \in \Omega} \|v - x\|, \tag{12}$$

where $\Omega \subseteq \mathbb{R}$ is a closed convex set, and λ_i locates within an accessible set $\Omega_i^{\lambda} = \{\lambda_i \in \mathbb{R} | \lambda_i^{\min} \leq \lambda_i \leq \lambda_i^{\max}\}$, where $\lambda_i^{\min} = 2\gamma_i P_i^{\min} + \beta_i$ and $\lambda_i^{\max} = 2\gamma_i P_i^{\max} + \beta_i$. In light of Liu et al. (2020), a two-step distributed optimization algorithm is designed based on the projection operator, taking into account the power generation constraints.

4.2.1 bfAlgorithm 1: two-step distributed optimization algorithm

Step 1: Distributed finite-time consensus policy.

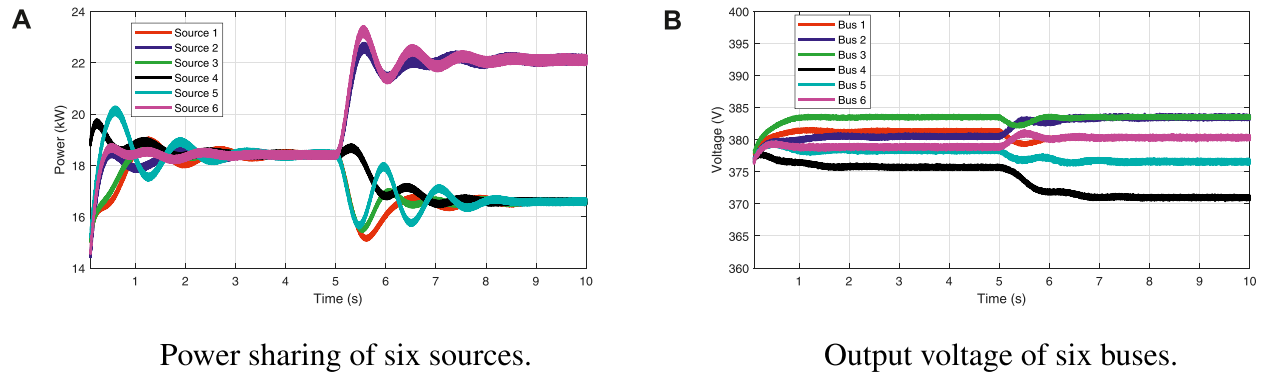


FIGURE 6
Performance of the proportional power sharing test: **(A)** output power of six sources and **(B)** voltage of six buses.

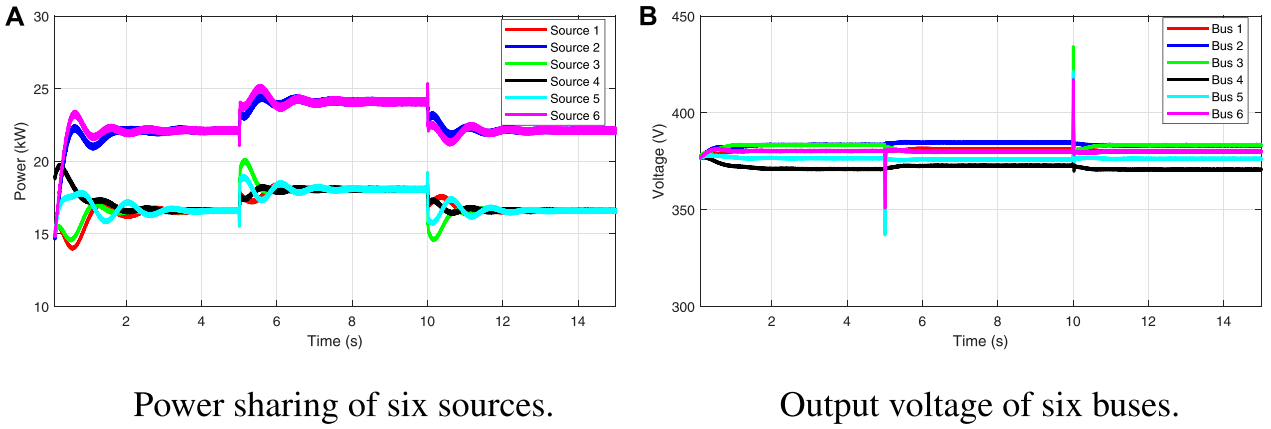


FIGURE 7
Performance of the load step test: **(A)** output power of six sources and **(B)** voltage of six buses.

$$\begin{cases} \lambda_i^{(l)}(t+1) = \lambda_i^{(l)}(t) - \mu(t)\gamma_i \sum_{j\in N_i} a_{ij}\left[\left(\lambda_i^{(l)}(t) - \lambda_j^{(l)}(t)\right)\right], \\ \lambda_i^{(l)}(0) = \lambda_{i0} + z_i^{(l)}. \end{cases} \tag{13}$$

Step 2: Initial value restoration operation.

$$z_i^{(l+1)} = \lambda_{con}^{(l)} - \mathrm{Proj}_i^{\lambda}\left(\lambda_{con}^{(l)}\right), \tag{14}$$

where $\mu(t) = \frac{1}{\lambda_{\mathcal{L},t}}$ is a time-dependent gain with $\lambda_{\mathcal{L},t}$ being the t-th eigenvalue of the Laplacian matrix $\mathcal{L}$. $\lambda_{i0} = 2\gamma_i P_i + \beta_i$, P_i is the output power, $\mathrm{Proj}_i^{\lambda}(\cdot)$ represents $\mathrm{Proj}(\cdot, \Omega_i^{\lambda})$ z_i is called as the restoration variable, and $\lambda_{con}^{(l)}$ is the consensus value to be calculated in **Step 1** at each iteration.

A major difference to Liu et al. (2020) is that we apply a fixed-time consensus algorithm via the discrete-time multi-agent system. The distributed optimization algorithm involves two steps for each source. In **Step 1**, a fixed-time discrete-time consensus algorithm is employed to drive λ_i to converge to consensus within N steps. In **Step 2**, we carry out the projection to operate and restore the initial value of λ_i according to the consensus value calculated in **Step 1**. These two steps are run alternately until $\lambda_{con}^{(l)}$ converges. The flowchart of the distributed optimization algorithm is depicted in detail in **Figure 3**.

Remark 3: *Like the optimization algorithms in Hu et al. (2018) and Hamad et al. (2016), the proposed distributed optimization algorithm utilizes the increment cost of neighbors, i.e., $\lambda_j, j \in N_i$. However, our algorithm still works without requirements on the total number of the sources N and the number of other neighbors $|N_j|, j \in N_i$.*

4.3 Interactive operation of the controller and optimizer

In the aforementioned sections, the proposed controller and optimizer separately achieve the desired proportional power sharing of sources and optimal power dispatch of sources. However, when the power sharing controller regulates the voltage profile, the transmission loss of lines changes accordingly such that the sum of real-time output power is no longer equal to the sum of optimized power, and to solve this problem, an interactive operation of the proposed controller and optimizer is presented, as shown in **Figure 4**.

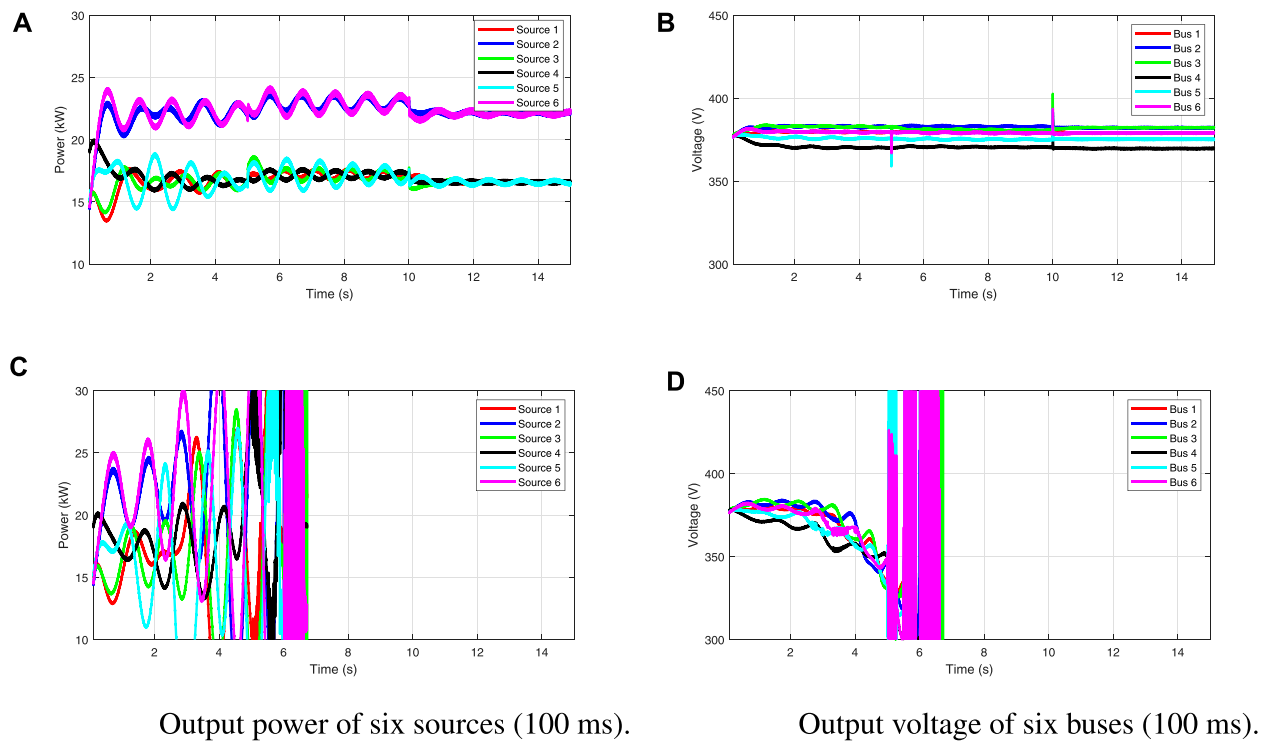


FIGURE 8
Performance of the time delay test: **(A)** output power of six sources (50 ms), **(B)** voltage of six buses (50 ms), **(C)** output power of six sources (100 ms), and **(D)** output voltage of six buses (100 ms).

In Figure 4, $\Delta\tau$ is the duration of optimization at each time and Δt is the time interval between two optimizations. During the period Δt, the power sharing controller sends the real-time output power to its generation optimizer. Subsequently, the optimizer calculates the optimal power dispatch and sends back the obtained dispatch to the controller as the proportionality coefficients. Then, the power sharing controller calculates the reference voltage for the zero-level controller, and the zero-level controller drives the DC microgrid to its steady state. Additionally, the controller will resend the current real-time output power to the generation optimizer for the next optimization. Specifically, the operation in the DC microgrid follows four steps:

1) The output power P_i of each bus at the physical system will be sent to its generation optimizer.
2) The optimal power dispatch $\widehat{P}_i^{opt}$ under the current circumstance (i.e., the current output power P_i) is calculated and sent back to the power-sharing controller.
3) The power-sharing controller works with the proportionality coefficient $p_i^r = \widehat{P}_i^{opt}$ until the DC microgrid reaches its steady state.
4) Repeat steps 1–3 until the optimal power dispatch $\widehat{P}_i^{opt}$ is convergent.

5 Simulation test

The simulation test is carried out based on MATLAB/Simulink to demonstrate the effectiveness of the proposed methods. Consider a meshed DC microgrid with six buses including six sources and six local loads, as shown in Figure 5. The rated voltage of the microgrid is selected at 380 V.

The line parameters and the load parameters are listed in Table 1. The parameters of generation cost are shown in Table 2. Each source is driven by a boost DC–DC converter with both current-loop and voltage-loop PI controllers, as shown in Figure 2, where $v_s = 220$ V, $L = 2$ mH, and $C = 470\ \mu$F. In the proposed control method, the parameter k is designed as 4,000. The constant power load is modeled via a DC/DC buck converter with a constant impedance load.

5.1 Proportional power sharing

Before $t = 5$ s, it is supposed to achieve the average power sharing, which means each source provides 16.67{%} of the total power demand. After $t = 5$ s, it is supposed to achieve a desired proportional power sharing, where six sources provide 15{%}, 15{%}, 15{%}, 20{%}, 15{%}, and 20{%} of the total power demand. In this case, p_i^r and $i \in \mathcal{I}_G$ are first set as 17,500, then p_1^r, p_2^r, p_3^r, andp_5^r will be reset as 15,750 and p_4^r and p_6^r are reset as 21,000 at $t = 5$ s. As seen in Figure 6A, the output power of six sources is converging to 17.5 kW at the first 5 s. During the last 5 s, the desired power sharing is achieved, where the output power of sources 1, 2, 3, and 5 converges to 15.75 kW and the output power of sources 4 and 6 converges to 21.0 kW as well. Moreover, voltage shifts of six buses are shown in Figure 6B. By the proposed

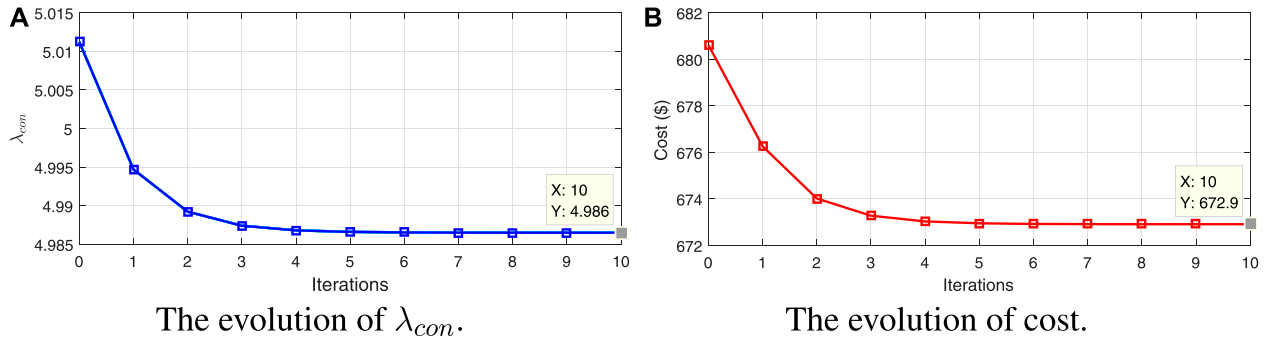


FIGURE 9
Simulation results of the distributed optimization algorithm: **(A)** evolution of λ_{con} and **(B)** evolution of the cost.

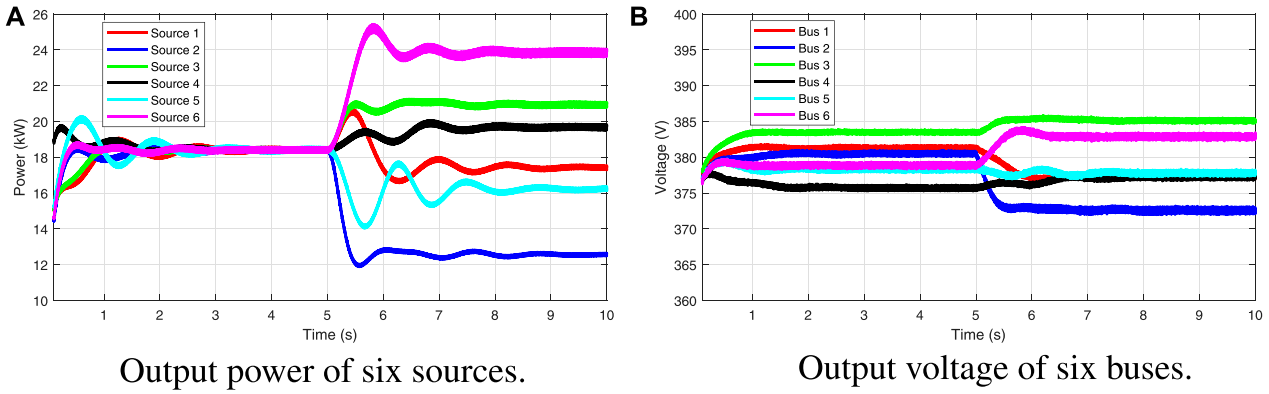


FIGURE 10
Performance of the optimal power sharing test: **(A)** output power of six sources and **(B)** voltage of six buses.

control method, the proportional power sharing of sources can be achieved.

5.2 Load step test

The controller performance of the load step test is shown in Figure 7, where two 5-kW loads are added at $t = 5$ s and removed at $t = 10$ s on buses 3 and 5, respectively. Figure 7A shows the proportional power sharing of six sources could still be maintained during the period of the test. Figure 7B shows the voltage shifts of six buses.

5.3 Power sharing with time delay

In this subsection, we have tested our power sharing algorithm with time delay, where two 2-kW loads are added at $t = 5$ s and removed at $t = 10$ s on buses 3 and 5, respectively. We have tested on different time delays, say, 50 ms and 100 ms, as shown in Figures 8A–D. It can be observed that power sharing is failed when the time delay is 100 ms, but power sharing can be achieved if the time delay is 50 ms.

5.4 Optimal power sharing

In this test, the optimizer starts to work at $t = 5$ s. The distributed optimization algorithm is employed with the setting $\mu = 3$. Within 10 iterations, it obtains the optimal power dispatch: $P_1^{opt} = 16.65$ kW, $P_2^{opt} = 12$ kW, $P_3^{opt} = 20$ kW, $P_4^{opt} = 18.80$ kW, $P_5^{opt} = 15.52$ kW, and $P_6^{opt} = 22.78$ kW. The evolution of λ_{con} and cost C are shown in Figures 9A, B . It observes that λ_{con} keeps decreasing and finally converges to 4.986. The cost of sources decreases as well and converges to a steady state within several iterations. The optimal power sharing of sources is achieved, as seen in Figure 10A. It is worth noting that the output power of source 2 reaches its lower bound, while the output power of source 3 reaches its upper bound. Moreover, voltage shifts of six buses are shown in Figure 10B.

6 Conclusion

This paper presents a solution to the optimal power sharing problem in a DC microgrid using a combination of the proportional power sharing control algorithm and the DIVR algorithm. The

proportional power sharing control algorithm, designed based on the Lyapunov method, is used to regulate the voltage profile of the microgrid. The fixed-time based distributed optimization algorithm, which uses a weighted finite-time consensus protocol and an initial value restoration algebraic operation, optimizes power sharing among the sources in the microgrid. Both algorithms are fully distributed and implemented at the cyber system layer. The algorithms work together to calculate the reference voltage for the zero-level controller to track, resulting in optimal power sharing among the sources. The effectiveness of the proposed method is demonstrated through simulation on a six-bus DC microgrid. Future work may focus on giving a theoretical bound for the time delay of the distributed optimal power-sharing controller and considering the constraint of the bus voltage during the control process.

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

ZY contributed to the conception and design of the framework. FY organized the overall paper. JC wrote sections of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the science and technology project (52153222001H) of the State Grid Hubei Electric Power Research Institute.

Acknowledgments

The authors would like to appropriate the editor and reviewers for their valuable suggestions and comments.

Conflict of interest

ZY, FY, and JC were employed by the State Grid Hubei Electric Power Co., Ltd. The authors declare that this study received funding from State Grid Hubei Electric Power Research Institute.The funder had the following involvement in the study: design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Ahmed, M. A., Abbas, G., Jumani, T. A., Rashid, N., Bhutto, A. A., and Eldin, S. M. (2023). Techno-economic optimal planning of an industrial microgrid considering integrated energy resources. *Front. Energy Res.* 11, 148. doi:10.3389/fenrg.2023.1145888

Anand, S., Fernandes, B. G., and Guerrero, J. (2013). Distributed control to ensure proportional load sharing and improve voltage regulation in low-voltage dc microgrids. *IEEE Trans. Power Electron.* 28, 1900–1913. doi:10.1109/TPEL.2012.2215055

Bidram, A., and Davoudi, A. (2012). Hierarchical structure of microgrids control system. *IEEE Trans. Smart Grid* 3, 1963–1976. doi:10.1109/TSG.2012.2197425

Chang, J.-W., Chae, S., and Lee, G.-S. (2023). Distributed optimal power sharing strategy in an islanded hybrid ac/dc microgrid to improve efficiency. *IEEE Trans. Power Deliv.* 38, 724–737. doi:10.1109/TPWRD.2022.3197434

Dou, Y., Chi, M., Liu, Z.-W., Wen, G., and Sun, Q. (2022). Distributed secondary control for voltage regulation and optimal power sharing in dc microgrids. *IEEE Trans. Control Syst. Technol.* 30, 2561–2572. doi:10.1109/TCST.2022.3156391

Guerrero, J. M., Vasquez, J. C., Matas, J., de Vicuna, L. G., and Castilla, M. (2011). Hierarchical control of droop-controlled ac and dc microgrids—A general approach toward standardization. *IEEE Trans. Industrial Electron.* 58, 158–172. doi:10.1109/TIE.2010.2066534

Guo, F., Wen, C., Mao, J., and Song, Y. D. (2016). Distributed economic dispatch for smart grids with random wind power. *IEEE Trans. Smart Grid* 7, 1572–1583. doi:10.1109/TSG.2015.2434831

Hamad, A. A., Azzouz, M. A., and El-Saadany, E. F. (2016). Multiagent supervisory control for power management in dc microgrids. *IEEE Trans. Smart Grid* 7, 1–1068. doi:10.1109/TSG.2015.2452191

Hamzeh, M., Ghazanfari, A., Mohamed, Y. A. R. I., and Karimi, Y. (2015). Modeling and design of an oscillatory current-sharing control strategy in dc microgrids. *IEEE Trans. Industrial Electron.* 62, 6647–6657. doi:10.1109/TIE.2015.2435703

Hatziargyriou, N., Asano, H., Iravani, R., and Marnay, C. (2007). Microgrids. *IEEE Power Energy Mag.* 5, 78–94. doi:10.1109/MPAE.2007.376583

Hou, X., Sun, K., Zhang, X., Sun, Y., and Lu, J. (2022). A hybrid voltage/current control scheme with low-communication burden for grid-connected series-type inverters in decentralized manner. *IEEE Trans. Power Electron.* 37, 920–931. doi:10.1109/TPEL.2021.3093080

Hou, X., Sun, Y., Zhang, X., Zhang, G., Lu, J., and Blaabjerg, F. (2019). A self-synchronized decentralized control for series-connected h-bridge rectifiers. *IEEE Trans. Power Electron.* 34, 7136–7142. doi:10.1109/TPEL.2019.2896150

Hu, J., Duan, J., Ma, H., and Chow, M.-Y. (2018). Distributed adaptive droop control for optimal power dispatch in dc microgrid. *IEEE Trans. Industrial Electron.* 65, 778–789. doi:10.1109/TIE.2017.2698425

Huang, B., Zheng, S., Wang, R., Wang, H., Xiao, J., and Wang, P. (2022). Distributed optimal control of dc microgrid considering balance of charge state. *IEEE Trans. Energy Convers.* 37, 1–2174. doi:10.1109/TEC.2022.3169462

Khorsandi, A., Ashourloo, M., and Mokhtari, H. (2014). A decentralized control method for a low-voltage dc microgrid. *IEEE Trans. Energy Convers.* 29, 793–801. doi:10.1109/TEC.2014.2329236

Li, C., de Bosio, F., Chen, F., Chaudhary, S. K., Vasquez, J. C., and Guerrero, J. M. (2017). Economic dispatch for operating cost minimization under real-time pricing in droop-controlled dc microgrid. *IEEE J. Emerg. Sel. Top. Power Electron.* 5, 587–595. doi:10.1109/JESTPE.2016.2634026

Li, C., Yu, X., Yu, W., Huang, T., and Liu, Z. W. (2016). Distributed event-triggered scheme for economic dispatch in smart grids. *IEEE Trans. Industrial Inf.* 12, 1775–1785. doi:10.1109/TII.2015.2479558

Liu, X.-K., Wang, S.-Q., Chi, M., Liu, Z.-W., and Wang, Y.-W. (2023a). Resilient secondary control and stability analysis for dc microgrids under mixed cyber attacks. *IEEE Trans. Industrial Electron.* 2023, 1–10. In press. doi:10.1109/TIE.2023.3262893

Liu, X.-K., Wen, C., Xu, Q., and Wang, Y.-W. (2021). Resilient control and analysis for dc microgrid system under DoS and impulsive FDI attacks. *IEEE Trans. Smart Grid* 12, 3742–3754. doi:10.1109/TSG.2021.3072218

Liu, X.-K., Yan, J., Xing, L., and Wen, C. (2020). Constrained consensus-based iterative algorithm for economic dispatch in power systems. *IECON 2020 46th Annu. Conf. IEEE Industrial Electron. Soc.* 2020, 3537–3542. doi:10.1109/IECON43393.2020.9255216

Liu, Z., Li, J., Su, M., Liu, X., Yuan, L., Xie, K., et al. (2023b). Evaluation of myopia-associated genes in a Han Chinese population with high myopia. *IEEE Trans. Power Syst.* 2023, 1–5. In press. doi:10.1080/13816810.2023.2191709

Loia, V., and Vaccaro, A. (2014). Decentralized economic dispatch in smart grids by self-organizing dynamic agents. *IEEE Trans. Syst. Man, Cybern. Syst.* 44, 397–408. doi:10.1109/TSMC.2013.2258909

Lu, X., Guerrero, J. M., Sun, K., and Vasquez, J. C. (2014). An improved droop control method for dc microgrids based on low bandwidth communication with dc bus voltage restoration and enhanced current sharing accuracy. *IEEE Trans. Power Electron.* 29, 1800–1812. doi:10.1109/TPEL.2013.2266419

Meng, L., Shafiee, Q., Trecate, G. F., Karimi, H., Fulwani, D., Lu, X., et al. (2017). Review on control of dc microgrids and multiple microgrid clusters. *IEEE J. Emerg. Sel. Top. Power Electron.* 5, 928–948. doi:10.1109/JESTPE.2017.2690219

Moayedi, S., and Davoudi, A. (2017). Unifying distributed dynamic optimization and control of islanded dc microgrids. *IEEE Trans. Power Electron.* 32, 2329–2346. doi:10.1109/TPEL.2016.2565517

Morstyn, T., Hredzak, B., and Agelidis, V. G. (2016a). Cooperative multi-agent control of heterogeneous storage devices distributed in a dc microgrid. *IEEE Trans. Power Syst.* 31, 2974–2986. doi:10.1109/TPWRS.2015.2469725

Morstyn, T., Hredzak, B., Demetriades, G. D., and Agelidis, V. G. (2016b). Unified distributed control for dc microgrid operating modes. *IEEE Trans. Power Syst.* 31, 802–812. doi:10.1109/TPWRS.2015.2406871

Nasirian, V., Davoudi, A., Lewis, F. L., and Guerrero, J. M. (2014). Distributed adaptive droop control for dc distribution systems. *IEEE Trans. Energy Convers.* 29, 944–956. doi:10.1109/TEC.2014.2350458

Olivares, D. E., Mehrizi-Sani, A., Etemadi, A. H., Canizares, C. A., Iravani, R., Kazerani, M., et al. (2014). Trends in microgrid control. *IEEE Trans. Smart Grid* 5, 1905–1919. doi:10.1109/TSG.2013.2295514

Papadimitriou, C., Zountouridou, E., and Hatziargyriou, N. (2015). Review of hierarchical control in dc microgrids. *Electr. Power Syst. Res.* 122, 159–167. doi:10.1016/j.epsr.2015.01.006

Simpson-Porco, J. W., Dorfler, F., and Bullo, F. (2013). Synchronization and power sharing for droop-controlled inverters in islanded microgrids. *Automatica* 49, 2603–2611. doi:10.1016/j.automatica.2013.05.018

Wang, L., Li, T., Cheng, Z., Hu, X., Li, Z., Liu, Z., et al. (2023). A unified droop control of ac microgrids under different line impedances: Revisiting droop control and virtual impedance method. *Front. Energy Res.* 11, 282–283. doi:10.1038/s41563-023-01487-3

Wang, P., Lu, X., Yang, X., Wang, W., and Xu, D. (2016a). An improved distributed secondary control method for dc microgrids with enhanced dynamic current sharing performance. *IEEE Trans. Power Electron.* 31, 6658–6673. doi:10.1109/TPEL.2015.2499310

Wang, Z., Wu, W., and Zhang, B. (2016b). A fully distributed power dispatch method for fast frequency recovery and minimal generation cost in autonomous microgrids. *IEEE Trans. Smart Grid* 7, 19–31. doi:10.1109/TSG.2015.2493638

Wood, A. J., and Wollenberg, B. F. (2012). *Power generation, operation, and control.* United States: John Wiley & Sons.

Xiao, Z.-X., Guan, Y.-Z., Fang, H.-W., Terriche, Y., and Guerrero, J. M. (2022). Dynamic and steady-state power-sharing control of high-efficiency dc shipboard microgrid supplied by diesel generators. *IEEE Syst. J.* 16, 4595–4606. doi:10.1109/JSYST.2021.3111685

Xing, H., Mou, Y., Fu, M., and Lin, Z. (2015). Distributed bisection method for economic power dispatch in smart grid. *IEEE Trans. Power Syst.* 30, 3024–3035. doi:10.1109/TPWRS.2014.2376935

Yang, S., Tan, S., and Xu, J. X. (2013). Consensus based approach for economic dispatch problem in a smart grid. *IEEE Trans. Power Syst.* 28, 4416–4426. doi:10.1109/TPWRS.2013.2271640

Zhao, C., He, J., Cheng, P., and Chen, J. (2016). Consensus-based energy management in smart grid with transmission losses and directed communication. *IEEE Trans. smart grid* 8, 2049–2061. doi:10.1109/tsg.2015.2513772

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 23 June 2023
DOI 10.3389/fenrg.2023.1203904

OPEN ACCESS

EDITED BY
Zhi-Wei Liu,
Huazhong University of Science and Technology, China

REVIEWED BY
Sahaj Saxena,
Thapar Institute of Engineering & Technology, India
Tiedong Ma,
Chongqing University, China
Ruizhuo Song,
University of Science and Technology Beijing, China

*CORRESPONDENCE
Jiru Chu,
chujr@cnpe.cc

RECEIVED 11 April 2023
ACCEPTED 08 June 2023
PUBLISHED 23 June 2023

CITATION
Qi J, Xu Z, Chu J, Zhu M and Teng Y (2023), Visual-admittance-based model predictive control for nuclear collaborative robots.
Front. Energy Res. 11:1203904.
doi: 10.3389/fenrg.2023.1203904

Visual-admittance-based model predictive control for nuclear collaborative robots

Jun Qi[1,2,3], Zhao Xu[1], Jiru Chu[1]*, Minglei Zhu[3] and Yunlong Teng[4]

[1]China Nuclear Power Engineering Co., Ltd., Beijing, China, [2]School of Automation, Chengdu University of Information Technology, Chengdu, China, [3]School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, China, [4]Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China (UESTC), Shenzhen, China

This paper presents a novel visual-admittance-based model predictive control scheme to cope with the problem of vision/force control and several constraints of a nuclear collaborative robotic visual servoing system. A visual-admittance model considering the desired image feature and force command in the image feature space is proposed. Moreover, a novel constraint scheme of the model predictive control (MPC) is proposed to cancel the overshoot in the interaction force control for most cases by taking the desired force command as the constraint of the proposed MPC. Via applying the robotic dynamic image-based visual servoing (IBVS) model, some other constraints, such as the actuator saturation, joint angle, and visual limits, can be satisfied simultaneously. The simulation results for the two-degrees-of-freedom (DOF) robot manipulator with an eye-to-hand camera are present to demonstrate the effectiveness of the proposed controller.

KEYWORDS

visual-admittance model, visual/force control, model predictive control, collaborative robot, dynamic visual servoing

1 Introduction

In recent years, collaborative robots have become more and more popular in the nuclear industry since the increasing need for safe and secure nuclear power plants (NPPs). Interaction between the robot and the environment or human is a fundamental requirement for the construction of NPPs 1, such as moving objects, deburring, grinding, polishing, and high-precision assembling (Nabat et al., 2005; Yang et al., 2008; Xu et al., 2017; Wu et al., 2017; Song et al., 2022, Figure 1). How to realize the precision position control and compliant interaction force control simultaneously in the nuclear environment is crucial for the execution of an interaction task.

To cope with this problem, numerous force control schemes have been proposed. Two main categories of those control schemes are direct and indirect force control. Direct force control adds the feedback of the force sensor directly into the closed motion control loop. For example, in the hybrid position/force control approach, the force and position controllers work separately and are connected with a diagonal section matrix (Raibert and Craig, 1981). However, these kinds of schemes lead to a trade-off between the force controller and the position controller, which implies that neither force nor position can precisely converge to their desired target. The indirect force control includes impedance control and admittance control. They all create a mass–spring–damper (MSD) model to describe the interaction force between the robot and the environment (Mason, 1981; Hogan, 1985) and are inverse

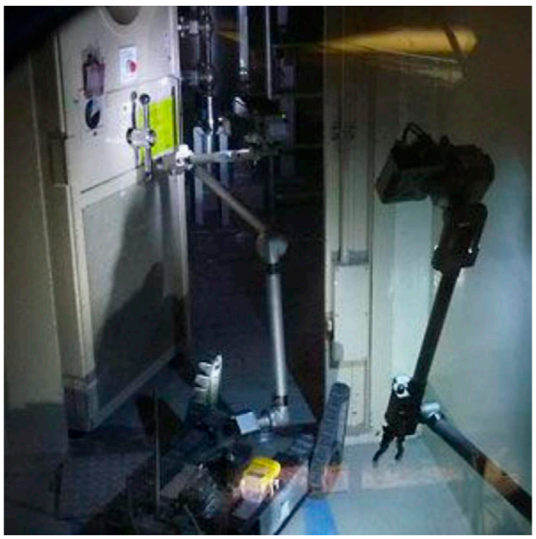


FIGURE 1
Application of collaborative robots in nuclear power plants (NPPs).

to each other. When the robot is controlled with the conventional model-based controller, precise interaction environment information and an accurate dynamic model of the robot are essential in compliant force control.

In the nuclear environment, where nuclear radiation can affect sensors and electronics, measuring accurate knowledge of the interaction environment and robot feedback is quite hard. Motivated by reducing or avoiding the need for precise preliminary knowledge of the environment, numerous researchers are trying to equip robot systems with vision and force sensors together. This method allows the robot system to get the environmental information and modify the vision/force controller online. In order to combine the visual and force information, several vision/force control schemes have been proposed to replace the motion control of the aforementioned force control with visual servoing (Bellakehal et al., 2011; Zhu et al., 2022b). For instance, in the hybrid vision/force control scheme (Zhu et al., 2022b), the force and vision are controlled separately, and a trade-off matrix is used to combine the output of two sensor controllers. Thus, these kinds of methods may result in local minima and reduce the robot control precision. In Lippiello et al. (2018), three vision-impedance controllers with the feedback of the camera and force sensor have been proposed to realize the physical interaction of a dual-arm unmanned aerial manipulator. However, the impedance control schemes can neither combine the vision and force sensing simultaneously nor avoid the trade-off between the two control loops.

To handle this problem, we propose a novel visual-admittance-based model to drive the robot with the trajectory of vision and the command of contact force. In this approach, the vision and force can be coupled in the image feature space successfully. Therefore, the convergence to a local minimum can be avoided, and the discrepancy between the modalities of vision and force sensors can be overcome.

However, due to the combination of force and vision in the image feature space, the controller is only employed to track the image feature commands generated from the aforementioned vision-admittance model without real-time perception of the contact force, which may lead to the unacceptable overshoot phenomenon and thus break the interaction environment or robot body. Therefore, we propose to consider the force command as a contact force constraint of the vision tracking controller. In practical applications, there are some other constraints on robot control systems, such as actuator saturation and visual limitation, that prevent the visual features from leaving the camera view.

Model predictive control (MPC) has been proven to be an efficient optimal control scheme aimed at addressing the disturbances and constraints of the system (Allibert et al., 2010; Deng et al., 2022). Recently, more and more researchers have concentrated on the advantages of MPC and applied this control scheme to the visual control of robots. In contrast to the aforementioned vision/force control scheme, MPC can explicitly handle the system constraints and resist disturbances. In Allibert et al. (2010), an MPC-based image-based visual servoing (IBVS) method is designed based on the conventional image Jacobian matrix and considers the constraints of image features. According to Hajiloo et al. (2016), a robust online MPC scheme based on the tensor product (TP) model of IBVS is proposed to realize the online control of MPC. By simplifying the image Jacobian matrix as a TP model, the computational speed of MPC is considerably increased. In Fusco et al. (2020b), the image feature acceleration is used to construct the model of MPC in order to get a shorter sensor trajectory and better motion in the feature space in Cartesian space. However, the aforementioned methods only consider the image feature constraints and kinematics constraints, such as the joint velocity or acceleration of the robot system, without taking into account the dynamic constraints of the robot system.

In general, in this paper, a novel visual-admittance-based MPC scheme with non-linear constraints based on the dynamic IBVS of the collaborative robot and desired force constraint is proposed to address the problem of the vision/force control with constraints. The visual-admittance model is formulated to converge the force and vision commands into image features to avoid the trade-off between two targets' (vision and force) tracking control. Considering the overshoot of the visual-admittance model, force constraints are added to the predictive model of the MPC. Moreover, the non-linear input bounds are added to the proposed MPC based on the dynamic model of the robot.

This paper is organized as follows: in Section 2, some preliminaries of the robot dynamic model and IBVS are presented. The vision-admittance-based trajectory generator is introduced to combine the force and visual information in Section 3. Section 4 presents the MPC controller with the non-linear constraints based on the dynamic model of the robot and force constraints. In Section 5, simulations based on a robot manipulator model with an eye-to-hand camera are performed to verify the effects of the proposed control scheme. Finally, the conclusions are given in Section 6.

2 Preliminaries

2.1 Dynamic model of the robot manipulator

The dynamic model of the manipulator robot with n actuators in the task space is often described as an Euler–Lagrange second-order equation (Roy et al., 2018). When the end effector of the robot comes into contact with the environment, the environment exerts an interactive force on the robot system. Considering the external force developing from the contact between the end-effector of the robot and the environment, the dynamic model of a robot can be written as

$$\mathbf{M}(q)\ddot{q} + \mathbf{C}(q,\dot{q})\dot{q} + \mathbf{G} + \mathbf{J}^T\mathbf{F}_e = \mathbf{T} + \mathbf{T}_d, \tag{1}$$

where $q, \dot{q}, \ddot{q} \in \mathcal{R}^n$ represent the position, velocity, and acceleration of the actuate joint angle, respectively. $\mathbf{M}(q(t)) \in \mathcal{R}^{n\times n}$ is the inertia matrix of the robot. $\mathbf{C}(q,\dot{q})$ denotes the Coriolis–centripetal matrix of the robot torques. $\mathbf{G} \in \mathcal{R}^{6\times 1}$ is the gravitational vector. $\mathbf{J}$ is the Jacobian matrix of the robot system. $\mathbf{F}_e \in \mathcal{R}^{6\times 1}$ represents the interactive force vector between the robot and environment. $\mathbf{T}, \mathbf{T}_d \in \mathcal{R}^n$ represent the input vector of the proposed controller and disturbance of the robotic system, respectively.

2.2 Dynamic model of image-based visual servoing

Image-based visual servoing (IBVS) is a sensor-based control scheme. It uses cameras as the main sensor to estimate the pose of the robot directly.

In this work, an eye-to-hand IBVS system is established to get the visual information, and we denote the visual feature estimated by the single fixed camera as $s \in \mathcal{R}^m$. For a point $p = (X, Y, Z)^T$ in the Cartesian space, the projection in the 2D image space is given as

$$\begin{cases} x = X/Z = (x_s - c_x)/P_x, \\ y = Y/Z = (y_s - c_y)/P_y \end{cases} \tag{2}$$

where $s = (x, y)^T$ is the projection of p in the image space and $s_i = (x_i, y_i)^T$ represents the coordinates of the image point expressed in pixel units. c_x, c_y, p_x, and p_y are the intrinsic parameters of the camera. c_x and c_y denote the coordinates of the principal point. p_x and p_y are the ratios between the focal length and pixel size.

Differentiating Eq. 2, the relationship between the time variation of the visual feature vector $\dot{s}$ and the spatial relative camera–object velocity of the robot $\dot{p}$ can be written as described in Mariottini et al. (2007)as follows:

$$\dot{s} = \mathbf{L}_s\dot{p}, \tag{3}$$

where L_s is the interaction matrix related to s (Chaumette and Hutchinson, 2006; Zhu et al., 2022a) given as

$$\begin{bmatrix} -1/Z & 0 & x/Z & xy & -(1+x^2) & y \\ 0 & -1/Z & y/Z & 1+y^2 & -xy & -x \end{bmatrix}. \tag{4}$$

The robot spatial velocity $\dot{p}$ can be transformed to the actuator velocity $\dot{q}$ with the Jacobian matrix $\mathbf{J}$. Then, Eq. 3 can be written as

$$\dot{s} = \mathbf{L}_s^c\mathbf{T}_e\mathbf{J}\dot{q}, \tag{5}$$

where ${}^c\mathbf{T}_e$ denotes the transform matrix from the kinematic screw to the camera frame. We define $\mathbf{J}_s = \mathbf{L}_s^c\mathbf{T}_e\mathbf{J}$. Equation 5 can be rearranged as follows:

$$\dot{s} = \mathbf{J}_s\dot{q}. \tag{6}$$

In order to generate the actuator torque, the relationship between the image feature and acceleration of the actuator joint has been formulated in the IBVS dynamic model (Fusco et al., 2020a; Liang et al., 2022).

By differentiating Eq. 3, the second-order interaction model can be demonstrated as

$$\ddot{s} = \mathbf{L}_s\ddot{p} + \dot{\mathbf{L}}_s\dot{p}. \tag{7}$$

When differentiating Eq. 5, we derive

$$\begin{aligned} \ddot{s} &= \dot{\mathbf{L}}_s^c\mathbf{T}_e\mathbf{J}\dot{q} + \mathbf{L}_s^c\dot{\mathbf{T}}_e\mathbf{J}\dot{q} + \mathbf{L}_s^c\mathbf{T}_e\dot{\mathbf{J}}\dot{q} + \mathbf{L}_s^c\mathbf{T}_e\mathbf{J}\ddot{q} \\ &= \mathbf{J}_s\ddot{q} + \mathbf{P}\dot{q}, \end{aligned} \tag{8}$$

where $\mathbf{P} = \dot{\mathbf{L}}_s^c\mathbf{T}_e\mathbf{J} + \mathbf{L}_s^c\dot{\mathbf{T}}_e\mathbf{J} + \mathbf{L}_s^c\mathbf{T}_e\dot{\mathbf{J}}$.

Substituting Eqs 5, 8 into Eq. 1, the dynamic model of IBVS can be written in the following form:

$$\mathbf{T} = \mathbf{M}\mathbf{J}_s^+\ddot{s} + \widehat{\mathbf{C}}\dot{s} + \mathbf{G} + \mathbf{J}^T\mathbf{F}_e - \mathbf{T}_d, \tag{9}$$

where $\widehat{\mathbf{C}} = \mathbf{C}\mathbf{J}_s^+ - \mathbf{M}\mathbf{J}_s^+\mathbf{P}\mathbf{J}_s^+$ and $^+$ represents the pseudo-inverse.

Assumption 1: During the control process, the robot and the IBVS controller do not encounter the controller singularity (Zhu et al., 2020). $\mathbf{J}$ and $\mathbf{L}_s$ are full-rank matrices.

3 Visual-admittance-based trajectory generator

In this section, a novel visual-admittance-based (VAB) trajectory generation is proposed and used to generate a reference trajectory for the robot end-effector online with feedback from the force sensor and camera in the image feature space. The structure of the robot manipulator system with an eye-to-hand visual servoing system is given in Figure 2.

The contact dynamic model between the robot end-effector and environment is often considered a second-order MSD model in the Cartesian space (Wu et al., 2022). We assume the contact model as follows:

$$\mathbf{K}_M\ddot{e}_p + \mathbf{K}_D\dot{e}_p + \mathbf{K}_S e_p = \mathbf{F}_e, \tag{10}$$

where $e_p = p_r - p_d$ represents the error between the reference trajectory and the desired trajectory of the robot end-effector in the Cartesian space. $\mathbf{K}_M, \mathbf{K}_D, \mathbf{K}_S$ denote the positive definite impedance parameters.

Due to the intrinsic technological constraints (such as the need for certain robot intrinsic parameters), the visual/force control

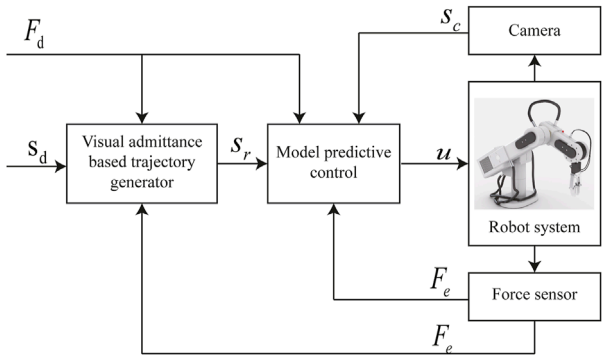


FIGURE 2
Structure of the visual-admittance-based model predictive control.

cannot be realized directly (Oliva et al., 2021). Therefore, the vision-admittance-based trajectory generator is designed to avoid those limits.

First, after some simple manipulations with Eq. 9, the relationship between image features and external force can be rewritten as follows:

$$\ddot{s} + \overline{\mathbf{C}}\dot{s} + d_s + f_{pumi} = f_u, \tag{11}$$

where

$$\begin{aligned} \overline{\mathbf{C}} &= \mathbf{J}_s\mathbf{M}^{-1}\widehat{\mathbf{C}}, \\ d_s &= \mathbf{J}_s\mathbf{M}^{-1}\left(\mathbf{G} - \mathbf{T}_d\right) \\ f_{pumi} &= \mathbf{J}_s\mathbf{M}^{-1}\mathbf{J}\mathbf{F}_e = \mathbf{J}_{sf}^T\mathbf{F}_e, \\ f_u &= \mathbf{J}_s\mathbf{M}^{-1}\mathbf{T}, \end{aligned} \tag{12}$$

where f_{pumi} represents the vector of per unit mass/inertia (p.u.m.i) virtual forces applied on the image feature space corresponding to the robot in the Cartesian space. f_u denotes the actuator torque projection in the image feature space.

Multiplying both sides of Eq. 10 by $\mathbf{J_s}$, the relationship between f_{pumi} and the position error e_p can be rearranged as follows:

$$f_{pumi} = \mathbf{J}_s\mathbf{K}_M\ddot{e}_p + \mathbf{J}_s\mathbf{K}_D\dot{e}_p + \mathbf{J}_s\mathbf{K}_S e_p. \tag{13}$$

Substituting Eqs 2, 3, 7 into Eq. 13 and denoting $p = \mathbf{P}_z s$, the impedance model corresponding to the image feature can be rearranged as

$$f_{pumi} = \widehat{\mathbf{K}}_M\ddot{e}_s + \widehat{\mathbf{K}}_D\dot{e}_s + \widehat{\mathbf{K}}_S e_s, \tag{14}$$

where

$$\begin{aligned} \widehat{\mathbf{K}}_M &= \mathbf{J}_s\mathbf{K}_M\mathbf{L}_s, \\ \widehat{\mathbf{K}}_D &= \mathbf{J}_s\left(\mathbf{K}_D - \mathbf{K}_M\mathbf{L}_s^+\dot{\mathbf{L}}_s\right)\mathbf{L}^+, \\ \widehat{\mathbf{K}}_S &= \mathbf{J}_s\mathbf{K}_S\mathbf{P}_z, \end{aligned} \tag{15}$$

where $e_s = s_r - s_d$ represents the error between the reference trajectory generated from the interaction force and the desired trajectory. When the contact force of the robot is zero, $s_r = s_d$. Otherwise, the vision-based admittance model can generate the reference trajectory to perform the force control of the system.

When the system is at equilibrium, the error of the image feature acceleration will converge to zero. To simplify the admittance model and reduce the calculation complexity, we set $\ddot{e}_s = 0$ and the simplified model can be described as

$$\dot{e}_s = -\frac{\widehat{\mathbf{K}}_S}{\widehat{\mathbf{K}}_D}e_s + \frac{f_{pumi}}{\widehat{\mathbf{K}}_D}. \tag{16}$$

By solving Eq. 16, we obtain

$$e_s(t) = s_0 e^{\mathbf{A}(t)} + e^{\mathbf{A}(t)}\int_0^t \frac{f_{pumi}}{\widehat{\mathbf{K}}_D}e^{-\mathbf{A}(\tau)}d\tau, \tag{17}$$

where $\mathbf{A}(t) = -\int_0^t \frac{\widehat{\mathbf{K}}_S(\tau)}{\widehat{\mathbf{K}}_D(\tau)}d\tau$ and s_0 is the initial state of the image feature. Then, the reference trajectory can be described as

$$\begin{aligned} s_r &= e_s(t) + s_d \\ &= s_0 e^{\mathbf{A}(t)} + e^{\mathbf{A}(t)}\int_0^t \frac{f_{pumi}}{\widehat{\mathbf{K}}_D}e^{-\mathbf{A}(\tau)}d\tau + s_d. \end{aligned} \tag{18}$$

Considering the desired force command $\mathbf{F}_d \in \mathcal{R}^6$, we propose a novel integral force control law (IFCL) to obtain the p.u.m.i virtual force f_{pumi} as follows:

$$f_{pumi}(t) = \mathbf{J}_{sf}\left(\mathbf{K}_d\dot{e}_f + \mathbf{K}_p e_f + \mathbf{K}_i\int_0^t e_f(\tau)\,d\tau\right), \tag{19}$$

where $e_f = F_d - F_e$, F_d is the predefined desired contact force, and $\mathbf{K}_d = diag(k_{d1}, \cdots k_{dn})$, $\mathbf{K}_p = diag(k_{p1}, \cdots k_{pn})$, and $\mathbf{K}_i = diag(K_{i1}, \cdots k_{in})$ are positive defined diagonal matrices.

4 Visual/force predictive controller

4.1 Model predictive control

MPC is an optimal control scheme proposed to solve the multi-variable constraint control problem. It determines the best input signal for a system by considering the finite future evolution of the system state. More precisely, it finds an optimal control sequence U^* that minimizes a cost function J_m with a set of predefined constraints over a finite predictive horizon N_p. The obtained control sequence is described as $U_t^* = \{u_t^*, \ldots, u_{t+N_c-1}^*\}$, where u_{i-1}^* is the optimal input signal in step i and $N_c \le N_p$ is called control horizon. The state predictions during the control horizon are calculated using the independent control inputs, while the remaining inputs are equal to the last elements of the control sequence. The cost function J_m is the sum of the quadratic error between the predictive state x_{t+i} and the desired state x_{t+i}^*. The optimal problem is shown as follows:

$$\mathbf{U}_t^* = \min_{\mathbf{U}_t} J_m(t), \tag{20}$$

with

$$J_m(t) = \sum_{i=1}^{N_p-1} e_{t+i}^T\mathbf{Q}_s e_{t+i} + \sum_{i=0}^{N_c-1} u_{t+i}^T\mathbf{Q}_u u_{t+i} + e_{N_p}^T\mathbf{Q}_p e_{N_p}, \tag{21}$$

which is subjected to the constraints

$$x_{t+i+1} = f\left(x_{t+i}, u_i\right), \tag{22a}$$

$$u_{t+i} \in \mathcal{U} \subseteq \mathbf{R}^{n_u}, i = 0, \ldots, N_c - 1, \tag{22b}$$

$$x_{t+i} \in \mathcal{X} \subseteq \mathbf{R}^{\mathbf{n_x}}, i = 1, \ldots, N_p, \tag{22c}$$

where $e_{t+i} = x_{t+i} - x^*_{t+i}$ and i represents time $i = iT_s$, where T_s is the control sample. $\mathbf{Q}_s$, $\mathbf{Q}_u$, and $\mathbf{Q}_p$ are the positive definite matrices that denote the relative importance of different components in cost function (Eq. 21). The constraint in Eq. 22a is the predictive model of MPC and is often chosen as the discrete-time model of a system. $\mathcal{U}$ and $\mathcal{X}$ are, respectively, the user-defined bounds of the input signal and system state. Once the optimal problem is solved, only u^*_t is fed to the system, and the obtained control sequence is used to generate the next initial control sequence of the optimal problem. Then, the optimal process is repeated.

4.2 The predictive model

In this work, we choose the joint configuration, interaction force, and image features as the state vectors of MPC. This allows the controller to explicitly take the image feature and force constraints into account and handle the redundancy during the optimization process. Then, the discrete predictive model of the IBVS robot system can be described as

$$\begin{bmatrix} q_{t+i+1} \\ \dot{q}_{t+i+1} \\ s_{t+i+1} \\ F_{t+i+1} \end{bmatrix} = \begin{bmatrix} q_{t+i} + \dot{q}_{t+}\Delta t \\ \dot{q}_{t+i} \\ s_{t+i} + \mathbf{J}_s\dot{q}_{t+i}\Delta t \\ F_{t+i} + K_{fm}\mathbf{J}\dot{q}_{t+i}\Delta t \end{bmatrix} + \begin{bmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \\ \mathbf{J}_s\Delta t \\ K_{fm}\mathbf{J}\Delta t \end{bmatrix} u_i, \tag{23}$$

where Δt is the sample time of MPC and u_i is the actuator acceleration at the step i. $\mathbf{J}_s$ and $\mathbf{J}$ are evaluated with the change in the state vector (Allibert et al., 2010). $q_t, \dot{q}_t, s_t, F_t$ are the initial states of the predictive model and can be obtained directly from the sensors of the robot system.

4.3 The joint acceleration constraints

The joint acceleration constraints aim at taking into account the physical boundaries of the actuators. Due to the interaction with the environment, it is possible that the input torque of the system exceeds the physical limit of the actuator torque, while the joint acceleration satisfies the rated acceleration constraints of the actuator. Thus, substituting the predictive state x_{t+i} into Eq. 1, the relationship between the joint acceleration and actuator torque can be rewritten as

$$\begin{aligned} u_{t+i} &= \mathbf{M}^{-1}\left(\mathbf{T}_{t+i} - \mathbf{C}\dot{q}_{t+i} - \mathbf{G} - \mathbf{J}^T\mathbf{F}_{t+i}\right) \\ &= \mathbf{M}^{-1}\mathbf{T}_{t+i} - \Pi_{t+i}. \end{aligned} \tag{24}$$

Then, the joint acceleration constraints of the MPC can be proposed as follows:

$$\hat{u}^{min}_{t+i} \le u \le \hat{u}^{max}_{t+i}, \tag{25a}$$

$$\hat{u}^{min}_{t+i} = \mathbf{M}^{-1}\mathbf{T}^{min}_{t+i} - \Pi_{t+i}, \tag{25b}$$

$$\hat{u}^{min}_{t+i} = \mathbf{M}^{-1}\mathbf{T}^{max}_{t+i} - \Pi_{t+i}, \tag{25c}$$

where $i = 1, \ldots, N_p$ and Π_{t+i} can be generated from the predictive state x_{t+i}. $\mathbf{T}^{max}_{t+i}$ and $\mathbf{T}^{min}_{t+i}$ represent the upper and lower bounds of the permissible actuator torque, respectively.

4.4 The terminal constraint (TC)

In this case, the stability of the proposed controller is ensured by a terminal constraint. TC imposes the last predicted visual features equal to the desired feature. However, it is difficult to satisfy a strict equality constraint when solving the optimizing problem. Thus, the constraints are converted into inequalities by defining a small enough threshold δ_{tc}.

$$\|s_{t+N_p} - s_d\| - \delta_{tc} \le 0. \tag{26}$$

4.5 The image feature constraints

In the context of visual serving, the image feature must remain visible. The following constraint prevents the image feature from escaping the field of the camera pixel view:

$$s^{min}_{t+i} \le s_{t+i} \le s^{max}_{t+i}, \tag{27}$$

where $i = 0, \ldots, N_c - 1$. s^{min}_{t+i} and s^{max}_{t+i}, respectively, represent the maximum and minimum bounds of the image feature in the camera pixel in sample step i.

4.6 The force constraints

Considering that the possible overshoot of force control can break the interactive environment, we set the desired force as a state constraint given as

$$\mathbf{F}_{t+i} \le \mathbf{F}^*_{t+i} + \delta_f, \quad i = 1, \ldots, N_p, \tag{28}$$

where $\mathbf{F}^*_{t+i}$ is the desired force in step i, and δ_f is the permissible error of the interaction force control.

5 Simulation

In this section, to demonstrate the control effects of the proposed visual-admittance-based model predictive controller, several simulations are illustrated in the two-link manipulator with an eye-to-hand camera. The force sensor is inserted in the end-effector of the robot, and the robot is controlled to track the predefined force and image feature trajectory.

5.1 Simulation design

In this simulation, we choose a two-link robot manipulator as the control object of the proposed controller. The structure of the

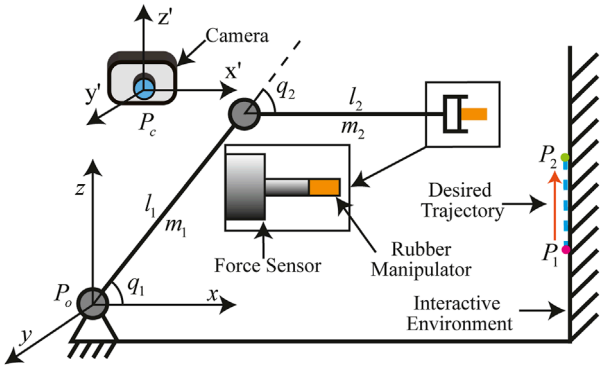


FIGURE 3
Structure of the two-link robot manipulator with eye-to-hand camera configuration.

two-link robot manipulator with eye-to-hand camera configuration is given in Figure 3. The coordinate of the origin of the robot system is $P_b = (0,0,0)$ m, and the dynamic model of this robot is given as

$$\mathbf{M}\ddot{q}(t) + \mathbf{C}\dot{q} + \mathbf{G} + \mathbf{J}^T\mathbf{F}_e = \mathbf{T}, \tag{29}$$

with

$$\mathbf{M}(q) = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22}, \end{bmatrix}$$
$$\mathbf{C}(q,\dot{q}) = \begin{bmatrix} -p_1\dot{q_2} & -p_1(\dot{q}_1+\dot{q}_2) \\ -p_1\dot{q}_1 & 0, \end{bmatrix}$$
$$\mathbf{G}(q) = \begin{bmatrix} p_2 g cos(q_1) + p_3 g cos(q_1+q_2) \\ p_3 g cos(q_1+q_2), \end{bmatrix}$$
$$\mathbf{J}(q) = \begin{bmatrix} -l_1 sin(q_1) - l_2 sin(q_s) & -l_2 sin(q_s) \\ l_1 cos(q_1) + l_2 cos(q_s) & l_2 cos(q_s), \end{bmatrix}$$

where,

$$\mathbf{M}_{11} = m_1 l_{1c}^2 + m_2(l_1^2 + l_{2c}^2 + 2l_1 l_{2c} cos(q_2)),$$
$$\mathbf{M}_{12} = m_2(l_{2c}^2 + l_1 l_{2c} cos(q_2))$$
$$\mathbf{M}_{21} = \mathbf{M}_{12},\ \mathbf{M}_{22} = m_2 l_{2c}^2,$$
$$p_1 = m_2 l_1 l_{2c} sin(q_2),\ p_2 = m_1 l_{1c} + m_2 l_1,$$
$$p_3 = m_2 l_{2c},\ q_s = q_1 + q_2,$$

where q_1 and q_2 are the joint angles of two actuators. m_1 and m_2 are the masses of links 1 and 2, respectively. l_1 and l_2 are the lengths of two links, respectively. l_{1c} and l_{2c} are the distances between the mass center of two links and the actuator joints, respectively. $p_1(0)$ and $p_2(0)$ are the initial joint angles. The aforementioned elements are listed in Table 1.

The camera set is fixed at $\mathbf{P}_c(0,0,-2)$ m, and the image plane of the camera (X'OZ') is set to be parallel to the Cartesian plane (XOZ). The camera resolution is 1280×720 pixels, and the focal length of the camera is 10 mm. The ratio of a pixel and unit length is 100 pixels/mm along the two axes of the pixel plane. The frequency of the camera observation is set at 50 Hz.

TABLE 1 Parameters and initial joint angle of the robot.

Parameter	Value	Unit	Parameter	Value	Unit
m_1	1.5	kg	m_2	1	kg
l_1	1.0	m	l_2	0.8	m
l_{c1}	0.5	m	l_{c2}	0.4	m
g	9.81	m/s^2			
$q_1(0)$	$\pi/3$	rad	$q_2(0)$	$-\pi/3$	rad
$\dot{q}_1(0)$	0	rad	$\dot{q}_2(0)$	0	rad

The constraints of actuator torque, actuator acceleration, and the image feature are, respectively, given as

$$\begin{aligned} -20\ Nm \le T \le 20\ Nm, \\ -5\ rad/s^2 \le \ddot{q} \le 5\ rad/s^2, \\ \begin{bmatrix} 0\ px \\ 0\ px \end{bmatrix} \le s \le \begin{bmatrix} 1280\ px \\ 720\ px \end{bmatrix}. \end{aligned} \tag{30}$$

In this simulation, we use Eq. 10 as the dynamic model of the force sensor. In the real interactive environment, compared with $\mathbf{k}_S$ and $\mathbf{K}_D$, the value of $\mathbf{K}_M$ is small enough to be ignored. Then, the parameters of this equation are given as $\mathbf{K}_S = 10000$ and $\mathbf{K}_D = 0.57$. The threshold of the TC is given as $\delta_{tc} = 100$.

In the visual-admittance model, the parameters of Eq. 19 are given as $\mathbf{K}_p = 5 \times 10^{-4}$, $\mathbf{K}_i = 6 \times 10^{-4}$, and $\mathbf{K}_d = 2 \times 10^{-5}$.

In order to validate the convergence capability and the control performance of the proposed controller, the position trajectory is predefined and consists of two phases: the approach phase and the interaction phase. The image feature trajectory we used in the MPC is generated by a virtual camera model from the position trajectory. During the approach phase, the robot needs to converge to P_1, (1.5, 0.4, 0) m in 1.5 s, and in the interaction phase, the robot moves from P_1 to P_2, (1.5, 0.365, 0) m in 3.5 s, and the position trajectory is given as follows:

$$p_d = \begin{bmatrix} 1.5 \\ 0.4 + 0.01^*(t-1.5) \\ 0 \end{bmatrix}. \tag{31}$$

During the interaction phase, a 10 N force command is applied along the x-axis from 2 to 4 s.

In order to verify the effects of the proposed controller, three sets of controllers are introduced in this simulation to drive the robot to track the desired trajectory:

- **Controller 1**: The MPC we proposed in this paper. The control and prediction horizon are set at $n_c = 5$ and $n_p = 6$, respectively, and the discretization time used in the controller is given as $\triangle t = 0.1$ s. The parameters of the predictive model cost function are chosen as $\mathbf{K}_{mf} =$, $10000, \mathbf{Q}_x = diag\{1000\}$, and $\mathbf{Q}_u = diag\{0.001\}$.
- **Controller 2**: The classical MPC without the torque or interaction force constraints. The parameters of this controller are similar to **Controller 1**.

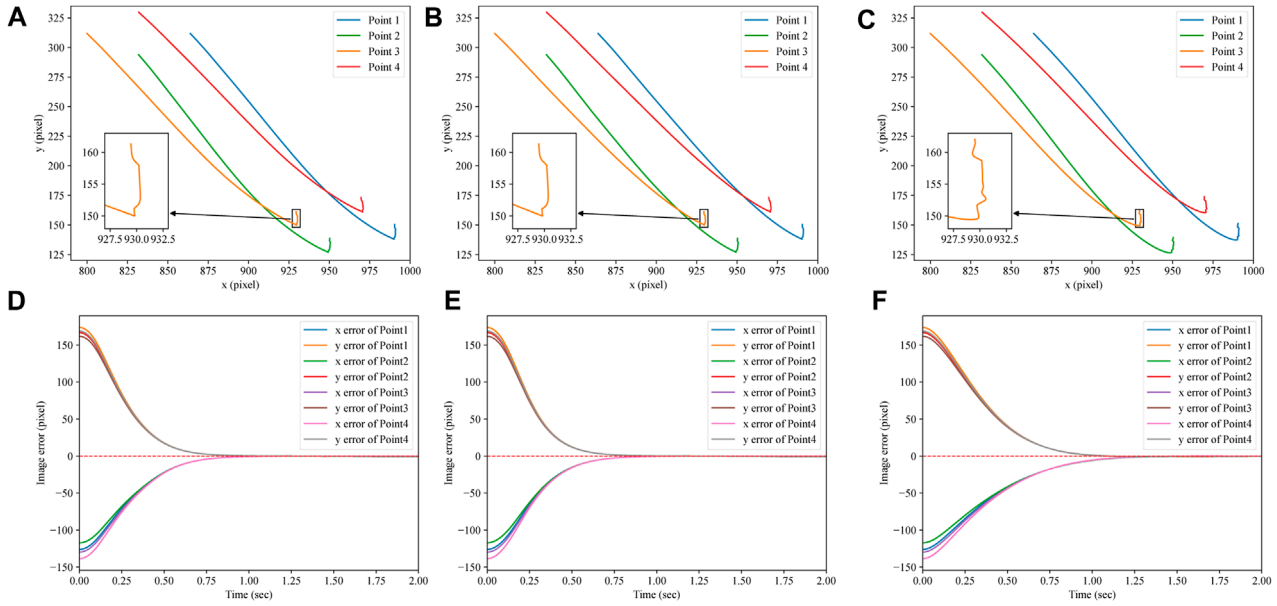


FIGURE 4
Image trajectories and errors of three controllers. **(A)** Image trajectory of the proposed MPC. **(B)** Image trajectory of the classical MPC. **(C)** Image trajectory of the PVFC. **(D)** Image errors of the proposed MPC. **(E)** Image errors of the classical MPC. **(F)** Image errors of the PVFC.

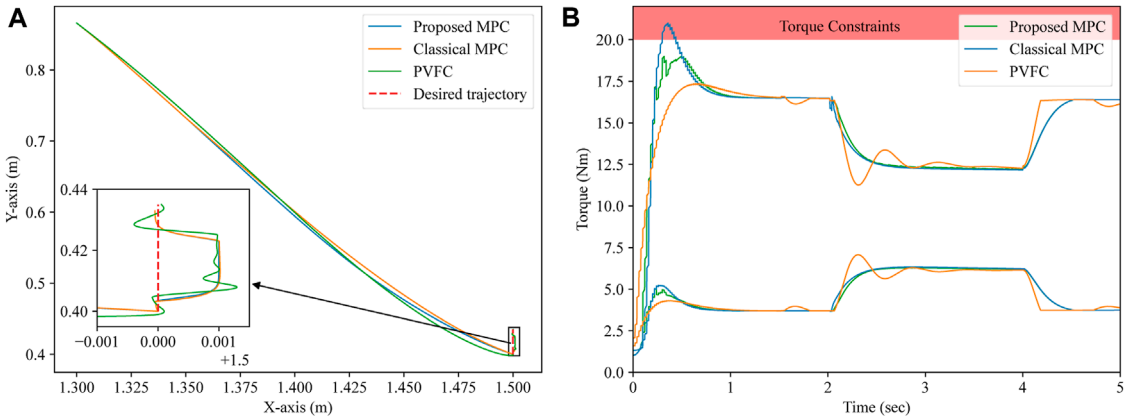


FIGURE 5
Position trajectories and input torques of three controllers. **(A)** Proposed and classical MPC. **(B)** PVFC.

- **Controller 3**: The parallel visual and force control (PVFC) scheme proposed in Zhu et al. (2022b) with input constraints. The parameters of this controller are given as $\mathbf{K}_v = 7.5$, $\mathbf{K}_p = 20$, $\mathbf{K}_i = 0.01$, and $\mathbf{K}_f = 0.01$.

5.2 Results and analysis

In this section, to illustrate the performance of the proposed controller with the torque constraints, three controllers are investigated through the simulation on a two-DOF robot manipulator with an eye-to-hand camera. We use MATLAB to conduct the simulation and sequential quadratic programming algorithm (*fmincon* function in the MATLAB optimization toolbox) to solve the optimization problem of MPC. The frequency of the MPC is equal to that of the camera which is given as 50 Hz. Similarly, the frequency of the torque generator is equal to that of the joint and force sensor given as 1,000 Hz.

The image trajectories and the image errors of the three aforementioned controllers are given in Figure 4. From Figures 4A–C, we can find that the image trajectory of the IBVS system under the proposed MPC is smoother than that under the PVFC and is similar to that under the classical MPC. When the force command is applied, the overshoot of the IBVS system under

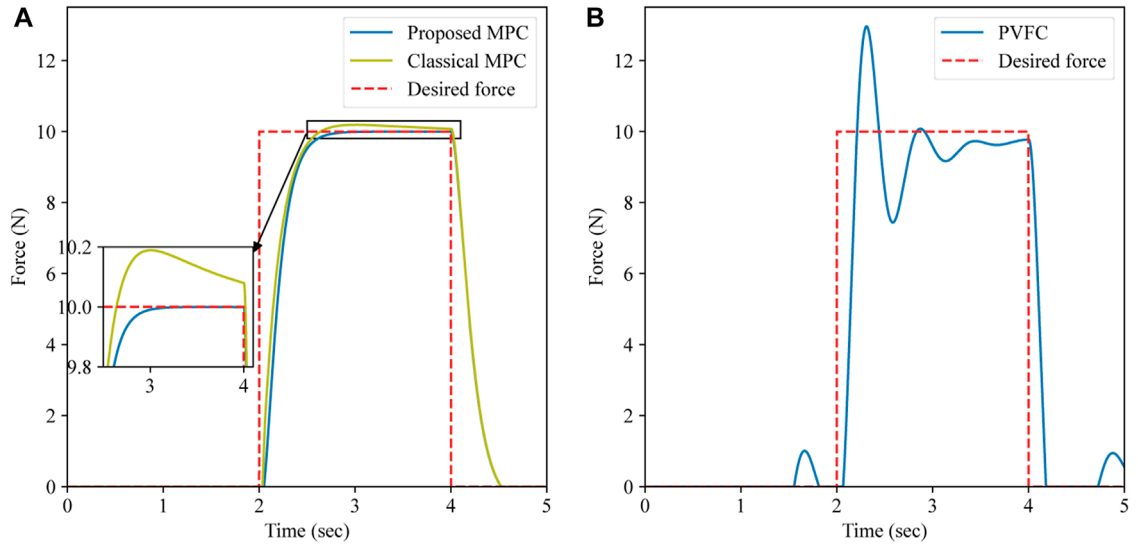


FIGURE 6
Force trajectories of the three controllers. **(A)** Proposed and classical MPC. **(B)** PVFC.

the PVFC is much more serious than that of the classical MPC and proposed MPC. As is shown in Figures 4D–F, during the approach phase, the converge time of the proposed MPC (approximately 0.8 s) is shorter than that of the PVFC (approximately 1.25 s) and is very close to the classical MPC (approximately 0.75 s).

The position and force trajectories of the three aforementioned controllers are shown in Figure 5A and Figure 6, respectively. As is shown in Figure 5A, during the approach process, the length of the position trajectory in the Cartesian space under the proposed controller is shorter than that under the PVFC and, in the interaction phase, the positioning accuracy of the IBVS robot system under the proposed MPC is better than that under the PVFC. From Figures 6A, B, we can find that the overshoot and chatters of the force control in the IBVS robot system under the PVFC are the most serious within the three aforementioned controllers, and the overshoot of the force control in the IBVS system under the classical MPC still exists. Nevertheless, due to the effectiveness of the virtual interaction model and the force constraints, the overshoot and chatters of the force control in the IBVS robot system under the proposed controller are eliminated.

The input torques of three controllers are given in Figure 5B. The red zone in Figure 5B is the predefined torque constraint. As is shown in Figure 5B, the classical MPC without the non-linear torque constraints cannot make the IBVS system obey the torque constraints strictly, and the proposed MPC can satisfy the torque constraints. Compared with the PVFC, in the interaction phase, the input torque of the IBVS system under the proposed MPC is smoother than that under the PVFC.

6 Conclusion

In this paper, a visual-admittance-based model predictive controller is developed to cope with the overshoot and trade-off of the classical position/force control scheme in the nuclear environment. A visual-admittance-based trajectory generator is presented to combine the desired image features and the force feedback into the reference image trajectory. Considering the integration element of the visual-admittance-based trajectory generator, we propose a novel predictive model constraints scheme. In this scheme, the desired force command is considered a constraint to avoid the overshoot of the force control, and the non-linear constraints based on the dynamic model of the robot are proposed to satisfy the actuator torque limits. A torque generator is used to generate the input signal of the robot system with the proposed MPC output and the real-time feedback of the robot joint sensors. An illustrative example of a two-DOF robot manipulator with an eye-to-hand camera is given to validate the effect of the proposed control scheme. Moreover, compared with the PVFC, the proposed MPC controller has better precision in force and position tracking. Compared with the classical MPC, the proposed controller can satisfy the image feature, torque, and force constraints and eliminate the overshoot of force control.

The proposed controller still needs information about the interaction environment, such as the elements of the virtual force model in the predictive model. Our ongoing research is finding a new observer to identify the elements of the interaction environment online. In addition, simplification of the optimization process in the MPC is in progress to construct a real-time MPC scheme of the IBVS robot system with non-linear constraints. In the future, experiments will be conducted to validate the proposed method.

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

JQ: methodology, software, validation, formal analysis, investigation, data curation, writing—original draft, and writing—review and editing. JC and ZX: validation, writing—review and editing, and data curation. MZ and YT: conceptualization, supervision, writing—review and editing, and funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Natural Science Foundation of Sichuan Province (2023NSFSC0872) and the Shenzhen Science and Technology Program (CYJ20210324143004012).

Conflict of interest

Authors JQ, ZX, and JC were employed by the company China Nuclear Power Engineering Co., Ltd.

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

References

Allibert, G., Courtial, E., and Chaumette, F. (2010). Predictive control for constrained image-based visual servoing. *IEEE Trans. Robot.*26, 933–939. doi:10.1109/TRO.2010.2056590

Bellakehal, S., Andreff, N., Mezouar, Y., and Tadjine, M. (2011). Vision/force control of parallel robots. *Mech. Mach. Theory.*46, 1376–1395. doi:10.1016/j.mechmachtheory.2011.05.010

Chaumette, F., and Hutchinson, S. (2006). Visual servo control. i. basic approaches. *IEEE Robot. Autom. Mag.*13, 82–90. doi:10.1109/mra.2006.250573

Deng, Y., Léchappé, V., Zhang, C., Moulay, E., Du, D., Plestan, F., et al. (2022). Designing discrete predictor-based controllers for networked control systems with time-varying delays: Application to a visual servo inverted pendulum system. *IEEE/CAA J. Autom. Sin.*9, 1763–1777. doi:10.1109/JAS.2021.1004249

Fusco, F., Kermorgant, O., and Martinet, P. (2020a). "A comparison of visual servoing from features velocity and acceleration interaction models," in 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).

Fusco, F., Kermorgant, O., and Martinet, P. (2020b). Integrating features acceleration in visual predictive control. *IEEE Robot. Autom. Lett.*5, 5197–5204. doi:10.1109/lra.2020.3004793

Hajiloo, A., Keshmiri, M., Xie, W.-F., and Wang, T.-T. (2016). Robust online model predictive control for a constrained image-based visual servoing. *IEEE Trans. Indust. Electron.*63, 1–2250. doi:10.1109/TIE.2015.2510505

Hogan, N. (1985). Impedance control: An approach to manipulation: Part III—applications. *J. Dyn. Syst. Meas. Control*107, 17–24. doi:10.1115/1.3140701

Liang, X., Wang, H., Liu, Y.-H., You, B., Liu, Z., Jing, Z., et al. (2022). Fully uncalibrated image-based visual servoing of 2dofs planar manipulators with a fixed camera. *IEEE Trans. Cybern.*52, 10895–10908. doi:10.1109/TCYB.2021.3070598

Lippiello, V., Fontanelli, G. A., and Ruggiero, F. (2018). Image-based visual-impedance control of a dual-arm aerial manipulator. *IEEE Robot. Autom. Lett.*3, 1856–1863. doi:10.1109/lra.2018.2806091

Mariottini, G. L., Oriolo, G., and Prattichizzo, D. (2007). Image-based visual servoing for nonholonomic mobile robots using epipolar geometry. *IEEE Trans. Robot.*23, 87–100. doi:10.1109/TRO.2006.886842

Mason, M. T. (1981). Compliance and force control for computer controlled manipulators. *IEEE Trans. Syst. Man. Cybern.*11, 418–432. doi:10.1109/tsmc.1981.4308708

Nabat, V., de la O Rodriguez, M., Company, O., Krut, S., and Pierrot, F. (2005). "Par4: Very high speed parallel robot for pick-and-place," in 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IEEE), 553–558.

Oliva, A. A., Giordano, P. R., and Chaumette, F. (2021). A general visual-impedance framework for effectively combining vision and force sensing in feature space. *IEEE Robot. Autom. Lett.*6, 4441–4448. doi:10.1109/LRA.2021.3068911

Raibert, M. H., and Craig, J. J. (1981). Hybrid position/force control of manipulators. *J. Dyn. Syst. Meas. Control.*102, 126–133. doi:10.1115/1.3139652

Roy, S., Roy, S. B., and Kar, I. N. (2018). Adaptive–robust control of euler–Lagrange systems with linearly parametrizable uncertainty bound. *IEEE Trans. Contr. Syst. Technol.*26, 1842–1850. doi:10.1109/TCST.2017.2739107

Song, S., Zhu, M., Dai, X., and Gong, D. (2022). "Model-free optimal tracking control of nonlinear input-affine discrete-time systems via an iterative deterministic q-learning algorithm," in IEEE Transactions on Neural Networks and Learning Systems, 1–14. doi:10.1109/TNNLS.2022.3178746

Wu, J., Gao, Y., Zhang, B., and Wang, L. (2017). Workspace and dynamic performance evaluation of the parallel manipulators in a spray-painting equipment. *Robot. Comput. Integr. Manuf.*44, 199–207. doi:10.1016/j.rcim.2016.09.002

Wu, R., Li, M., Yao, Z., Liu, W., Si, J., and Huang, H. (2022). Reinforcement learning impedance control of a robotic prosthesis to coordinate with human intact knee motion. *IEEE Robot. Autom. Lett.*7, 7014–7020. doi:10.1109/LRA.2022.3179420

Xu, P., Li, B., and Chueng, C.-F. (2017). "Dynamic analysis of a linear delta robot in hybrid polishing machine based on the principle of virtual work," in 2017 18th International Conference on Advanced Robotics (ICAR) (IEEE), 379–384.

Yang, G., Chen, I.-M., Yeo, S. H., and Lin, W. (2008). "Design and analysis of a modular hybrid parallel-serial manipulator for robotised deburring applications," in *Smart devices and machines for advanced manufacturing* (Springer), 167–188.

Zhu, M., Chriette, A., and Briot, S. (2020). "Control-based design of a DELTA robot," in ROMANSY 23 - Robot Design, Dynamics and Control, Proceedings of the 23rd CISM IFToMM Symposium.

Zhu, M., Briot, S., and Chriette, A. (2022a). Sensor-based design of a delta parallel robot. *Mechatronics*87, 102893. doi:10.1016/j.mechatronics.2022.102893

Zhu, M., Huang, C., Qiu, Z., Zheng, W., and Gong, D. (2022b). Parallel image-based visual servoing/force control of a collaborative delta robot. *Front. Neurorobot.*16, 922704. doi:10.3389/fnbot.2022.922704

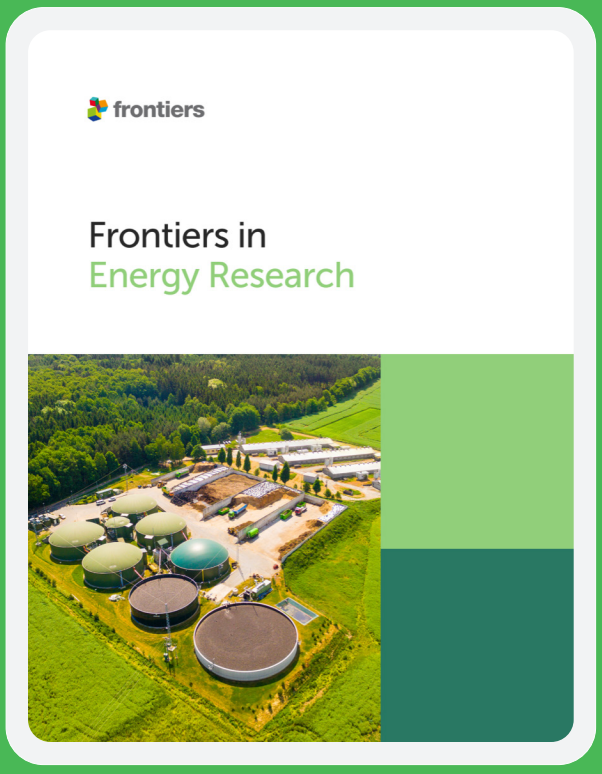
Frontiers in
Energy Research

Advances and innovation in sustainable, reliable and affordable energy

Explores sustainable and environmental developments in energy. It focuses on technological advances supporting Sustainable Development Goal 7: access to affordable, reliable, sustainable and modern energy for all.

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