



OPEN ACCESS

EDITED BY

Fernando Montani,
National Scientific and Technical Research
Council (CONICET), Argentina

REVIEWED BY

Ekaterina Levichkina,
The University of Melbourne, Australia
Yingjian Yang,
Shenzhen Lanmage Medical Technology Co.,
Ltd., China

*CORRESPONDENCE

Jeffrey A. Riffell
✉ jriffell@uw.edu

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

[‡]These authors have contributed equally to
this work and share last authorship

RECEIVED 27 June 2025

ACCEPTED 18 September 2025

PUBLISHED 16 October 2025

CITATION

Yang H, KC P, Chen P, Lei H, Sponberg S,
Tarokh V and Riffell JA (2025) Neuron
synchronization analyzed through
spatial-temporal attention.
Front. Comput. Neurosci. 19:1655462.
doi: 10.3389/fncom.2025.1655462

COPYRIGHT

© 2025 Yang, KC, Chen, Lei, Sponberg, Tarokh
and Riffell. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Neuron synchronization analyzed through spatial-temporal attention

Haoming Yang^{1†}, Pramod KC^{2†}, Panyu Chen³, Hong Lei⁴,
Simon Sponberg⁵, Vahid Tarokh^{2‡} and Jeffrey A. Riffell^{2*‡}

¹Department of Electrical and Computer Engineering, Duke University, Durham, NC, United States,

²Department of Biology, University of Washington, Seattle, WA, United States, ³Department of
Computer Science, Duke University, Durham, NC, United States, ⁴School of Life Sciences, Arizona
State University, Tempe, AZ, United States, ⁵Schools of Physics and Biological Sciences, Georgia
Institute of Technology, Atlanta, GA, United States

Neuronal synchronization refers to the temporal coordination of activity across populations of neurons, a process that underlies coherent information processing, supports the encoding of diverse sensory stimuli, and facilitates adaptive behavior in dynamic environments. Previous studies of synchronization have predominantly emphasized rate coding and pairwise interactions between neurons, which have provided valuable insights into emergent network phenomena but remain insufficient for capturing the full complexity of temporal dynamics in spike trains, particularly the interspike interval. To address this limitation, we performed *in vivo* neural ensemble recording in the primary olfactory center—the antennal lobe (AL) of the hawk moth *Manduca sexta*—by stimulating with floral odor blends and systematically varying the concentration of an individual odorant within one of the mixtures. We then applied machine learning methods integrating modern attention mechanisms and generative normalizing flows, enabling the extraction of semi-interpretable attention weights that characterize dynamic neuronal interactions. These learned weights not only recapitulated the established principles of neuronal synchronization but also facilitated the functional classification of two major cell types in the antennal lobe (AL) [local interneurons (LNs) and projection neurons (PNs)]. Furthermore, by experimentally manipulating the excitation/inhibition balance within the circuit, our approach revealed the relationships between synchronization strength and odorant composition, providing new insight into the principles by which olfactory networks encode and integrate complex sensory inputs.

KEYWORDS

neural synchronization, bio-inspired neural networks, generative model, attention-mechanism, antennal lobe

1 Introduction

Interconnected neural populations construct a meaningful perception of the sensory features of the complex external world (Hopfield, 1995; Reiter and Stopfer, 2013). Neural representations of the sensory world are temporally structured, and this temporal organization drives the selective behaviors by influencing which neurons are recruited, when they are activated, and how intensely they fire in the central brain (Hopfield, 1995; Laurent, 2002). The coordinated timing of neural activity of these neurons—referred to as synchronization—is thought to enhance sensory perception, sharpen neural representation, and enable complex sensory discrimination. This allows the foraging species to exhibit fast and effective decisions by facilitating effective transmission of information to the downstream targets (Riffell et al., 2009a, 2014).

Sensory signals are dynamic and multidimensional, requiring the brain to integrate the information across both time and space through distributed neuronal populations to generate meaningful representations (Laurent, 2002; Wachowiak et al., 2025). These spatial-temporal properties of the neurons necessitate the synchronization among neurons to encode identity, intensity, and valence of different sensory stimuli, including odors (Stopfer et al., 2003; Bitzenhofer et al., 2022). Odor processing involves the synchronization of the neural activity across multiple levels of the brain. In mammals, neural synchrony occurs in the olfactory bulb (OB). It may drive activity through sets of mitral and tufted cells in higher brain centers, such as the piriform cortex, entorhinal cortex, and amygdala (Laurent, 2002; Wilson and Sullivan, 2011). Similar processes may be at play in insects, where neural synchrony occurs in the primary olfactory center, the antennal lobe (AL), and activates downstream neurons in the higher brain centers, including the lateral horn and mushroom body, thought to be involved in valence and learning and memory, respectively (Riffell et al., 2009b). The synchronized activities of neurons in higher brain centers are represented not only by firing rate but also by the precise timing of neuronal activity among tens of thousands of neurons. The interaction between these neurons remains an open question (Laurent, 2002; Stopfer et al., 1997).

The insect olfactory system provides a tractable neuroanatomical model for examining the functional bases of olfactory processing. Olfactory detection starts with the activation of the olfactory sensory neurons, and axons of these neurons terminate in the distinct anatomical structures called glomeruli in the primary olfactory center, the antennal lobe (Chu et al., 2020; Kymre et al., 2021; Lei et al., 2011). Each glomerulus is innervated by several projection neurons (PNs) that relay information from the AL to the higher brain areas (Homberg et al., 1988; Kymre et al., 2021). The odor-activated PNs interact with each other via a dense network of inhibitory neurons, local interneurons (LNs) spanning throughout the AL (Hoskins et al., 1986). Several studies have shown that odorants evoke strong synchronous firing of the PNs in several species (Christensen et al., 1987; Heinbockel et al., 2004; Lei et al., 2002; Martin et al., 2013; Nagel and Wilson, 2016; Tsai et al., 2018; Wilson et al., 2004), providing odor-specific representations and encoding (Mazor and Laurent, 2005; Stopfer et al., 1997). However, studies have typically used monomolecular odorants or binary mixtures, which do not represent the complexity of the natural olfactory environment (Riffell et al., 2009a; Wachowiak et al., 2025).

Natural odors are complex mixtures made up of tens to hundreds of compounds (Bisch-Knaden et al., 2022; Riffell et al., 2009a). The proportion of odorants in the mixture is critical in odor encoding, particularly in a dynamic environment. Previous studies have shown that changing the concentration of only one compound significantly affects an insect's ability to discriminate and track the floral odor (Riffell et al., 2009b). However, the cellular and computational bases by which the olfactory system binds specific features of the complex odor mixture—including the critical odorants—are not known, and neuromorphic principles that are involved in such processes are still uncertain, given the diverse physiological and morphological properties of these neuron types (Biedlerack et al., 2006; Carlsson et al., 2007; Guerrieri et al., 2005).

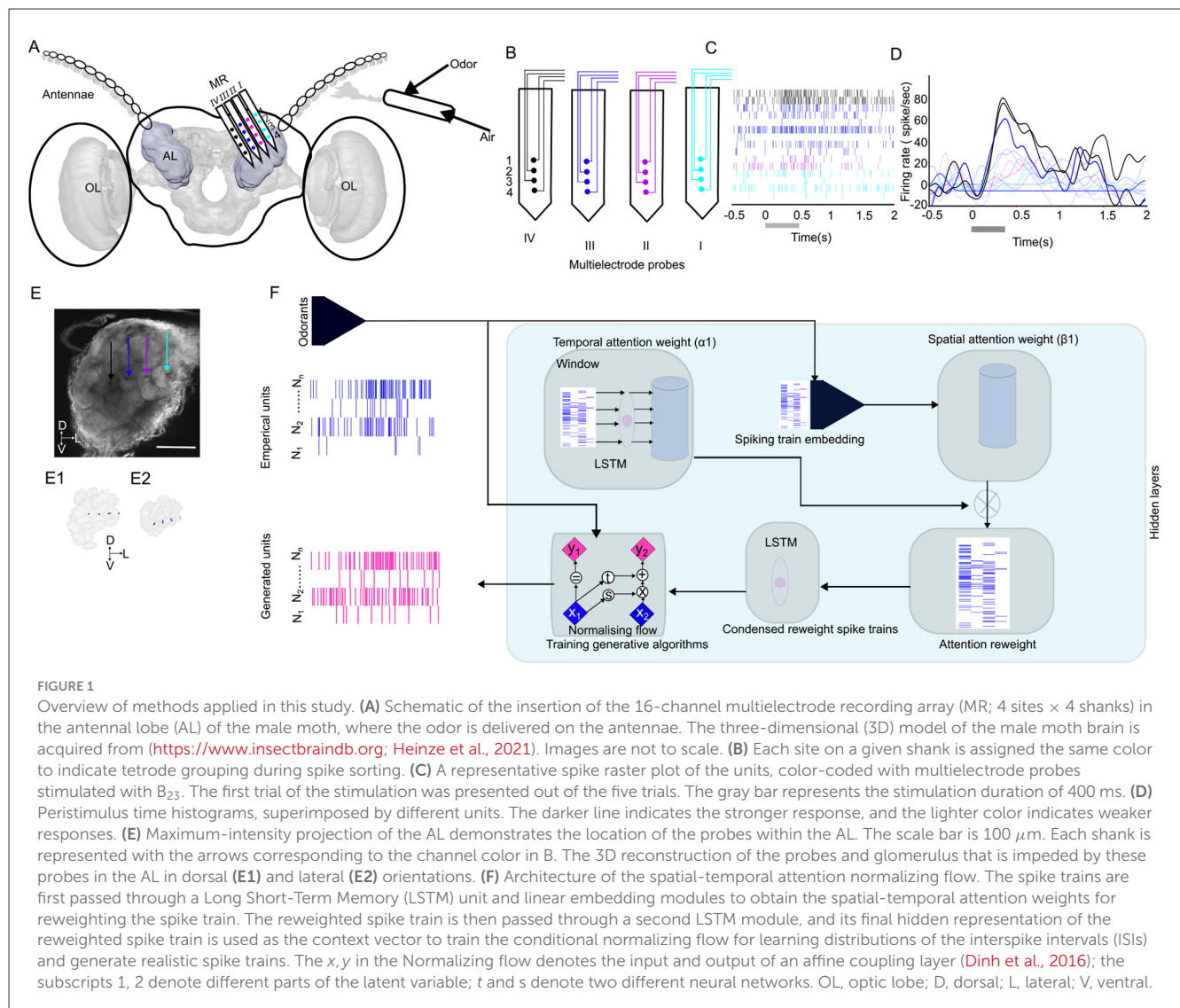
Despite advances in modeling either spatial or temporal neural dynamic patterns, current methods fall short in jointly capturing spatiotemporal dynamics due to computational and structural limitations. Previous methods for temporal [recurrent neural network (Laurent et al., 1996), Hebbian plasticity (Delahunt et al., 2018)] or spatial aspects of neural responses have been investigated. In these established methods, either a paired or a subset of the neurons (Stopfer et al., 2003) were commonly utilized in investigating the temporal dynamics, but the interaction of the neurons at the population level has been broadly overlooked. Generally, the correlation between neurons is a common strategy to investigate the sensory encoding in population codes based on the tuning similarities, stimulus effects, and the presence of higher order correlations (reviewed in Panzeri et al., 2022). In previous studies, responsive neurons were considered for investigation, whereas non-responsive units, often excluded from analysis, are frequently encountered during experiments (Fiscella et al., 2015; Ni et al., 2022; Pasupathy and Connor, 2002). However, the non-responsive units could play an important role in population coding through synergy with the responsive neurons (Haggard and Chacron, 2025). The mechanism by which individual neurons in a population selectively weigh the stimulus and influence other neurons has not been investigated.

In this study, we address how an olfactory neural population encodes complex odor information for decision making. Through *in vivo* neural ensemble recordings and computational analyses, we treat the spiking of neurons as a point process in time. Instead of assuming the arrival time of spikes follows a canonical distribution (e.g., Poisson distribution), we use the highly flexible, non-parametric, deep normalizing flow to model the probability distribution of interspike intervals (ISIs: Kobyzev et al., 2021). During the modeling process of the spike train of a specific neuron in the AL, we introduce a novel spatial-temporal attention module to learn how individual neurons synchronize with the rest of the neuron population (spatial) and are affected by population spike trains dynamically (temporal). This spatial attention weight module accounts for the higher order interactions across a population of neurons, allowing us to analyze complex population-level synchronization beyond the pairwise analyses of Ensemble Synchronization Index (Riffell et al., 2009b) and Kernelized binless methods (Martin et al., 2013). We found that our method outperforms in the classification of odors, considering both the interactions within and between LNs and PNs in the AL. In addition, our study indicates that increasing the proportion of a compound in the mixture—thereby altering the excitation/inhibition balance in the AL—could reduce the pattern of neural synchrony.

2 Data curation

2.1 Insect preparation

Adult male moths (*Manduca sexta*; Lepidoptera: Sphingidae) were reared in the laboratory on an artificial diet under a long-day (17/7-h light/dark cycle) photoperiod. The moths (3 days old, post-eclosion) were secured in a 10 ml plastic pipette (ThermoFisher Scientific, USA) with dental wax (Kerr Corporation, Romulus,



MI, USA), leaving the head and antennae exposed. The cuticle on the head was carefully cut to expose the brain, and all the muscles, trachea, and neural sheath were carefully removed with fine forceps (Fine Science Tool, USA). The restrained moth was mounted to a recording platform attached to the vibration isolation table. The preparation was placed such that the ALs are orientated dorsofrontally. The brain was superfused slowly with physiological saline solution [150 mM NaCl, 3 mM CaCl₂, 3 mM KCl, 10 mM N-Tris(hydroxymethyl) methyl-2 aminoethanesulfonic acid buffer, and 25 mM sucrose, pH 6.9] throughout the experiment.

2.2 Odor stimulation

Pulses of air (100 ml/min) were pushed through a glass cartridge containing a piece of Whatman filter paper (Millipore Sigma, USA) loaded with 10 μ l of floral odorant and injected into a constant air stream (1 L/min) directed toward the moth's antennae (Figure 1A). The stimulus was pulsed through a solenoid-actuated

valve controlled by an RZ2 bioamplifier processor (Tucker-Davis Technologies, Inc., FL, USA). The outlet of the stimulus cartridge was positioned 2 cm from and orthogonal to the center of the antennal flagellum. Stimulus duration was 400 ms, and five pulses were separated by either a 5-s interval or a 10-s interval. The interstimulation duration was approximately 1 min. The tested stimuli were categorized as behavioral (B, green colored) and non-behavioral (purple and light purple colored; Figures 2C1, C2, 3A–C, 4C, D, 5C–F). Different odor stimuli under each category are annotated with the subscripted number. We classify the blend as a behavioral one if it contains all three compounds: Benzaldehyde (O₁), Benzyl alcohol (O₂), and Linalool (O₆). For our first sets of an experiment, we tested behavioral stimuli including extracts of *Datura* flowers (B₁₅), five artificial mixtures (B₁₆–B₁₉) containing the behavioral components, three dilutions of B₁₉ [10 times dilution (B₂₀), 100 times dilution (B₂₂), 10,000 times dilution (B₂₂)] and non-behavioral stimuli include mineral oil (control, no odor), five mixtures of non-behavioral components (NB₁₀–NB₁₄), and nine individual non-behavioral components (O₁–O₉). In the second experiment, to determine how modifying

the ratio of compounds in the mixture modified the encoding of floral odor, we used an odor cartridge containing the behavioral mixture (B_{23}) and a second odor cartridge containing increased concentrations of Benzaldehyde (O_1) with 10-, 100- or 1,000-fold higher concentrations (B_{24} , B_{25} , and B_{26}). The odors from the two odor cartridges were released simultaneously into the airway, allowing them to mix before reaching the flagellum. In this manner, the ratio of compounds in the behavioral mixture (B_{23}) could be dynamically altered. See [Table 1](#) for the stimuli and their compositions.

2.3 Ensemble antennal lobe recording

Odor-evoked responses were obtained from 80 units in seven male moths. Recordings were made with 16-channel silicon multielectrode recording (MR) arrays (A4X4-3 mm-50-177; NeuroNexus Technologies). These probes have four shanks (each of 15 μm in thickness) spaced 125 μm apart, each with four recording sites 50 μm apart, and have a surface area of 177 μm^2 indicating the distinct spatial activity patterns across the different regions in the antennal lobe ([Figures 1B–E](#)) ([Lei et al., 2004](#)). The MR was positioned under visual control with a stereo microscope (Narishige, Japan). As demonstrated in [Figures 1A, E](#), the four shanks were oriented in a line parallel to the antennal nerve. The MR was advanced slowly through the AL with a micromanipulator (Narishige, Japan) until the uppermost recording sites were just below the surface of the AL. Thus, the four shanks of the MR were recorded from four regions of glomerular neuropil across the AL. Ensemble activity was recorded simultaneously from the 16 channels of the MR array by using TDT amplifiers (Tucker-Davis Technologies, Inc., FL, USA). The recorded signal was digitized at 25 kHz per channel by using Synapse software (version 98, Tucker-Davis Technologies, Inc., FL, USA).

2.4 Localization of recording probes in the AL

The head was excised, and the brains were dissected in Manduca saline. The brain was washed with 0.01 M phosphate-buffered saline (PBS) (2 times: each for 20 min) and then submerged in the solution consisting of 4% PFA and 0.03% glutaraldehyde to facilitate tissue fixation. The preparation was kept overnight at 4°C and then dehydrated in a series of ethanol (50%, 70%, 90%, 96%, 100%, and 100%: each for 20 min). Finally, it was cleared in methyl salicylate (Millipore Sigma, USA). The whole-mount preparation is scanned with a laser scanning microscope (Nikon, A1R, Nikon Instruments, Inc., USA) equipped with a CFI Plan Apo 10 $\times$ Air objective, which is scanned with a 488 nm line of an argon laser. The high-resolution confocal images with 1,024 $\times$ 1,024 pixels at a distance of 2–4 μm in the z -direction were obtained. The image was imported in AMIRA version 6.5.0 (ThermoFisher Scientific, USA) and the glomerular structures were reconstructed ([Figure 1E](#)). The shank impaled in the AL was also reconstructed and visualized.

2.5 Spike sorting

The continuous waveforms are exported to an offline sorter (Offline sorter, Plexon, Version 4.7.1). The spike data were digitized at 25 kHz per channel. The filter setting (0.6–5 kHz and system gain of 1,000) were software adjustable on each channel. Spikes were sorted by tetrode configuration using a clustering algorithm based on the method of principal components (PCs) (Off-line Sorter; Plexon). We used dual thresholds, between 2 and 3 standard deviations (SD) above the baseline. The highest and lowest thresholds will eliminate the voltage deflections deriving from movement artifacts that are misinterpreted as spikes. The groups that were separated in the 3D space (PC1–PC3) after statistical verification (multivariate analysis of variance (ANOVA); $p < 0.05$) were selected for further analysis (7–19 units were isolated per ensemble; [Figures 1C, D](#)). Each spike in each cluster was time-stamped, and these data were used to create raster plots and calculate peristimulus time histograms (PSTHs). Preliminary analyses were performed with Neuroexplorer (Nex Technologies, version 5.4) using a bin width of 5 ms unless otherwise stated.

3 Model

In this section, we present Spatial-Temporal Attention Normalising Flow (STAN-Flow), a generative framework for modeling neuronal spike train dynamics through the distribution of interspike intervals (ISIs). Our approach combines a spatial-temporal attention module, which encodes both neuronal interactions and temporal spiking history, with a conditional normalizing flow that flexibly models ISI distributions without restrictive parametric assumptions. This integration enables us to capture stimulus-driven synchronization and timing patterns in a principled and interpretable manner. We then conclude the section with an introduction to the different methods and metrics we compute and compare in the results section.

3.1 Problem formulation and notation

We denote spike train $S^q \in \mathbb{R}^{N \times T}$ for a total of N neurons, T timesteps, and $q \in 1, \dots, Q$ different stimuli. For the n -th neuron, we denote the time of the i -th spike timing as t_i^{qn} , hence the previous spike's timing as t_{i-1}^{qn} . Then the interarrival time between these two spikes is $\tau_i^{qn} = t_i^{qn} - t_{i-1}^{qn}$. The goal is to model the interarrival time distribution τ_i^{qn} for arbitrary spike i .

Let $[\Delta]$ denote the window size and denote a windowed spike train as $S_{i[\Delta]}$; we assume the interspike intervals of neuron n are conditionally independent given the stimuli q , the Δ -windowed history of last spike, and time of the last spike to reduce the problem into modeling the distribution of interarrival time presented in [Equation 1](#). For ease of notation, we drop the superscript q and focus on an arbitrary stimulus; we also drop the superscript n and focus on an arbitrary neuron.

$$P(\tau_i | S_{i-1[\Delta]}, q, t_{i-1}) \quad (1)$$

Our generative modeling approach consists of two main components: (1) the spatial-temporal attention units that encode

TABLE 1 Odor stimulation and the mixture composition.

Stimuli class	Stimuli	Name	Composition
Single odorants	O ₁	Benzaldehyde	
	O ₂	Benzyl alcohol	
	O ₃	Farnesene	
	O ₄	Geraniol	
	O ₅	Caryophyllene	
	O ₆	(±)Linalool	
	O ₇	Methyl salicylate	
	O ₈	Myrcene	
	O ₉	Nerol	
Non-behavioral blend	NB ₁₀		Benzaldehyde + Benzyl alcohol
	NB ₁₁		Methyl salicylate + Caryophyllen + Farnesene
	NB ₁₂		Methyl salicylate + Caryophyllen + Farnesene+Myrcene
	NB ₁₃		Methyl salicylate + Caryophyllen + Farnesene + Myrcene + Geraniol
	NB ₁₄		Methyl salicylate + Caryophyllen + Farnesene+ Myrcene + Geraniol + Nerol
Behavioral blend	B ₁₅	Datura extract	
	B ₁₆		Benzaldehyde + benzyl alcohol + linalool
	B ₁₇		Benzaldehyde + benzyl alcohol + linalool+nerol
	B ₁₈		Benzaldehyde + benzyl alcohol + linalool+ nerol + geraniol
	B ₁₉		Benzaldehyde + benzyl alcohol + linalool + nerol + geraniol + myrcene + methyl salicylate + caryophyllene + farnesene
	B ₂₀	10× dilution of B19	
	B ₂₂	100× dilution of B19	
	B ₂₂	10,000× dilution of B19	
	B ₂₃		Benzaldehyde + benzyl alcohol + linalool + nerol+geraniol + methylsalicylate + methylbenzoate
	B ₂₄	10× increase of Benzaldehyde	
	B ₂₅	100× increase of Benzaldehyde	
	B ₂₆	1000× increase of Benzaldehyde	
Control	Ctl	Mineral oil	

the windowed spike history to latent space; and (2) a conditional generative model learned with a normalizing flow that models the target distribution presented in Equation 1. Beyond the goal of learning the conditional ISI distribution, the attention weights learned by this conditional generative system can be extracted for further analysis. We present the full spatial-temporal attention normalizing flow (STAN-Flow) architecture in Figure 1F.

3.2 Spatial-temporal attention

The spatial-temporal attention mechanism combines the vision attention mechanism in computer vision (Dosovitskiy et al., 2021) and temporal attention in natural language processing (Bahdanau et al., 2016; Vaswani et al., 2023). Synchronization can be viewed

as interactions between neurons, which can be characterized through spatial attention, where a higher spatial attention weight corresponds to a stronger interaction between neurons. The importance of particular spike timing and the general spiking rate is characterized by the temporal weights that scan through the spiking history: the higher the temporal weight, the more important a specific time is. Therefore, neurons can be synchronized in their activity even if they have different individual temporal dynamics.

The spatial-temporal attention module consists of Long Short-Term Memory (LSTM) layers and a few linear layers. The LSTM layers effectively summarize spike train time series into lower dimensional hidden states, which are then projected by the linear layers to obtain semi-interpretable attention weights. The windowed spike train $S_{i-1[\Delta]}$ is passed through the first LSTM (f_1) and a linear spatial-embedding layer parameterized as $\mathcal{W}_e$ and b_e , outputs the hidden representation h_{i-1} and d -dimensional spike train spatial embedding $E_i \in \mathbb{R}^{N \times d}$. For the temporal attention,

we further reduce h_{i-1} through another linear layer parameterized as $\mathcal{W}_\alpha$ and b_α to obtain the temporal weights, $\alpha_i^t \in \mathbb{R}^{[\Delta]}$.

$$\begin{aligned} E_i &= \mathcal{W}_e(S_{i-1[\Delta]}) + b_e = [e_i^1, \dots, e_i^N]^T \\ h_{i-1} &= f_1(S_{i-1[\Delta]}) \\ \alpha_i^t &= \text{softmax}(\text{Tanh}(\mathcal{W}_\alpha h_{i-1} + b_\alpha)) \end{aligned}$$

Additionally, Riffell et al. (2009a) shows that stimuli information is also encoded by the ensemble firing of neurons. Hence we concatenate the last hidden states of the LSTM, denoted as h^* , the spatial embedding of a particular neuron's activity $e_i^m, m \in 1 \dots N$, and the stimuli q to pass through a linear layer parameterized as $\mathcal{W}_\beta$ and b_β and obtain the spatial weights $\beta_i^m \in \mathbb{R}$. The vector that contains the spatial weights of all neurons is denoted as $\beta_i \in \mathbb{R}^N$. To isolate the higher-order ensemble patterns, we replace the traditional *softmax* of the attention mechanism with its sparsifying counterpart *sparsemax* (Martins and Astudillo, 2016) which directly projects logit values onto the simplex. Applying the *sparsemax* allows some attention weights to be reduced to zero, amplifying the effect of those synchronized neurons.

$$\begin{aligned} \beta_i^m &= \text{sparsemax}(\text{Tanh}(\mathcal{W}_\beta [h^*; e_i^m; q] + b_\beta)) \\ \beta_i &= [\beta_i^1, \dots, \beta_i^N]^T \end{aligned}$$

Usual applications of attention weights obtain a context vector through a weighted average of latent variables (Bahdanau et al., 2016; Vaswani et al., 2023). However, a weighted average of each neuron's representation dilutes the synchronization identified through the spatial attention weight as the resulting weighted representation becomes less identifiable. Hence, we reweight the windowed spike train with the mean-normalized spatial-temporal weights and feed the reweighted windowed spike train ($S'_{i-1}[\Delta]$) through a second LSTM layer (f_2) to obtain the final hidden representation h'_{i-1} . The symbol $\times$ denotes element-wise multiplication.

$$S'_{i-1[\Delta]} = \beta_i [\alpha_i^t]^T \times S_{i-1[\Delta]} \quad (\text{Reweighting Spike Train})$$

$$h'_{i-1} = f_2(S'_{i-1[\Delta]}) \quad (\text{Obtain final hidden representation})$$

The output of f_2 , h'_{i-1} , can be seen as a context vector derived from the reweighted spike train. This context vector combines the temporal dynamic spike train, the neuron interaction, and the influence of different stimuli into a continuous representation to inform the conditional generative model.

It is worth noting that the attention mechanism introduced in this section is inspired by the attention mechanism first introduced in translation tasks, rather than the more well-known self-attention mechanism for transformer-based architectures (Bahdanau et al., 2016; Vaswani et al., 2023). This modeling choice is due to several practical concerns. First, the self-attention mechanism offers limited interpretability. The transformer architecture linearly embeds data into uninterpretable lower dimensions, then applies the attention mechanism on the value matrix; this limits the interpretation of the attention matrix compared to our formulation, which directly applies the attention weights to modify the spike trains. Second, fitting self-attention mechanism on the temporal

and spatial dimensions through the transformer architecture requires two transformer modules; this will induce much more computation overhead and is less efficient. Finally, due to the size of transformer-based architecture, a large amount of data is usually required to leverage its ability to process diverse contexts; however, neural datasets generally lack this level of diversity to train a transformer without heavy regularization (Ye and Pandarinath, 2021).

3.3 Conditional normalizing flow

Once we learn the synchronization and timing information, we build a modeling module to accurately reflect the ISI distribution based on synchronization and temporal dynamics. While traditionally the modeling of spike trains follows the Poisson Process, this assumes the spike train is rate-coded and the ISI distribution follows an exponential distribution. These assumptions are not always realistic and constrain the modeling process. Instead of a model based on the Poisson assumption, we build a non-parametric deep generative model conditioned on the final hidden representation h'_{i-1} to learn the probability distribution of interspike intervals given the learned history.

We chose to apply a conditional normalizing flow that directly optimizes the negative log-likelihood of the density. A normalizing flow is usually defined by a transformation of a standard Gaussian distribution into a more complex distribution (Kobyzev et al., 2021). This transformation typically involves a sequence of invertible, tractable, and differentiable mappings, enabling the evaluation of a sample's value in the simple distribution or its likelihood.

We concatenate the stimuli q , the last hidden representation of attention-reweighted spike train h'_{i-1} , and the time of last spike t_{i-1} into a context vector denoted as x_i . We propose a normalizing flow that is conditioned on x_i ; the likelihood takes the following form, where Z is drawn from a conditional Gaussian distribution. Extending recent neural network architecture (Kobyzev et al., 2021; Dinh et al., 2016), a loss through log-likelihood can be written as:

$$x_i = [h'_{i-1}; q; t_{i-1}] \quad (2)$$

$$\begin{aligned} \log P(\tau_i | x_i) &= \log P(z_i | x_i) \det \left| \frac{\partial z}{\partial \tau} \right| = \log P(f_\theta(\tau_i, x_i)) \\ &+ \log \left(\det \left| \frac{\partial f_\theta(\tau_i, x_i)}{\partial \tau} \right| \right) \end{aligned} \quad (3)$$

We specifically applied the real-valued non-volume-preserving normalizing flow architecture (RealNVP) in our study (Dinh et al., 2016), where the f_θ is characterized through a series of neural networks that construct an upper triangular Jacobian, simplifying the determinant computation of the Jacobian to be the trace.

$$\log P(\tau_i | x_i) = \log P(f_\theta(\tau_i, x_i)) + \text{Tr} \log \left(\left| \frac{\partial f_\theta(\tau_i, x_i)}{\partial \tau} \right| \right) \quad (4)$$

This framework considers the spiking history, the interaction between neurons, and the stimulus effect, and learns the ISI distribution without assuming it follows some canonical, parametric distribution. Hyperparameters related

to the architecture and training process are listed in the [Supplementary material](#).

3.4 Identifying synchronization

A crucial part of our analysis is establishing a higher order non-linear method to analyze neuron synchronization. We propose the spatial attention weight method and compare it with two previously reported neuron synchronization methods, the Ensemble Synchronization (ESI) proposed in [Lei et al. \(2004\)](#), and the Kernelized Binless Method (KB) applied in [Martin et al. \(2013\)](#); we then discuss the synchronization analysis process of our proposed spatial-attention weights method.

3.4.1 Ensemble synchronization

The traditional analysis of ensemble patterns utilizes the cross-correlation coefficient between pairs of neurons ([Lei et al., 2004](#)). In particular, the synchronization index (SI%) is calculated as

$$\text{SI\%}_{\text{ensemble}} = \frac{[\text{CE}]_{\text{raw}}^{\delta} - [\text{CE}]_{\text{shuffle}}^{\delta}}{C_1(T) + C_2(T)} \times 100\% \quad (5)$$

where $[\text{CE}]_{\text{raw}}$ is the number of coincident events in the cross-correlogram peak centered around $t = 0$ with width δ , $[\text{CE}]_{\text{shuffle}}$ is the number of coincident events after trial shuffling (shift predictor method) to correct for coincidences attributable to chance and an increased firing rate. The corrected correlograms were calculated by averaging four trial shifts and subtracting the result from the raw correlogram. T is the total response time over which spikes were counted, and C_1 and C_2 are the number of spikes recorded from units 1 and 2 during time T ([Lei et al., 2004](#)).

We calculated the ensemble SI% for all stimuli using one trial as the raw trial and corrected it by shuffling the other four trials. We applied the parameters $\delta = 5$ and $T = 1,000$ (ms) after the onset of the stimuli as suggested in [Riffell et al. \(2009a\)](#). In the [Supplementary Figure S1](#), we explore a variety of hyperparameters for δ and T .

3.4.2 Kernelized binless method

A more recent method for analyzing the synchronization of neuron firing is through the kernelized binless method ([Martin et al., 2013](#)). While it remains a pairwise synchronization analysis, it applies an exponential function kernel to smooth out the spike train. Specifically, the exponential kernel is denoted $h(t) = \exp(-t/\phi) \times u(t)$ where $u(t)$ is the heavy side function, ϕ is a kernel parameter to aggregate spikes over time; a similarity index (see [Equation 6](#)) is then calculated between a pair of neurons' kernelized spiking.

$$\text{SI\%}_{\text{binless}} = \frac{s_1 \cdot s_2}{||s_1|| ||s_2||} \quad (6)$$

We compute $\text{SI\%}_{\text{binless}}$ with the time constant $\phi = 5(\text{ms})$ similar to previous synchronization analysis ([Martin et al., 2013](#)). Trial shuffling is also applied for the kernelized binless method. In the [Supplementary Figure S2](#), we explore a variety of hyperparameters for ϕ .

3.4.3 Spatial-attention method

There are a total of five trials for each stimulus in the dataset. Therefore, we train a conditional normalizing flow for each neuron by applying a cross-validation scheme, in which we rotate three trials to form the training set, while the other two form the validation and test sets. The validation set is used for model selection, which ensures that the STAN-Flow will not be overfit to the training data. The spatial-attention module (see Section section 3.2) is learned jointly with the conditional normalizing flow through the loss function ([Equation 4](#)). For an arbitrary spike i of an arbitrary neuron n , and arbitrary stimuli q , our modeling process would determine a set of attention weights that determines the importance of each neuron in the neuron population. During our analysis, we concatenate the spatial-attention for each stimulus, then average the spatial-attention weights over all spikes, all neurons, and all runs during evaluation to output a synchronization summary matrix $\mathcal{B}$. The specific calculation is shown in [Equation 9](#).

$$\beta^{qn} = \sum_i \beta_i^{qn} \quad \beta^{qn} \in \mathbb{R}^N \quad (7)$$

$$\mathcal{B}^n = [\beta^{1n} \dots \beta^{Qn}] \quad \mathcal{B}^n \in \mathbb{R}^{Q \times N} \quad (8)$$

$$\mathcal{B} = \frac{1}{N} \sum_n \mathcal{B}^n \quad (9)$$

Each column of $\mathcal{B}$ can be seen as the “strength” of neuron for the Q different stimuli. If a subset of neurons with elevated spatial-attention weights for a particular stimulus q , it suggests that these neurons are determined to be synchronized by the neural network. An example of this matrix is provided in [Supplementary Figure S7](#).

3.5 Classification of neurons into putative PN and LN

Although our ensemble recorded neuronal data does not allow us to identify the neuron types, we follow the classification procedure described in [Lei et al. \(2011\)](#) to classify PNs and LNs in our spike-sorted units. This classification method relies on the observation that the spontaneous spiking activity of PNs and LNs is different: PNs are more likely to have bursts of spiking activity while LNs fire regularly. It adopts the criterion in [Legendy and Salcman \(1985\)](#) to detect potential bursts in spontaneous activities (5s) in the full spike train from Poisson Surprise (Poisson S) rates, which characterizes the abrupt changes in spiking rates compared to the mean spike rate.

$$\text{Poisson S} = rT - \log \left[\sum_{j=n}^{\infty} \frac{(rT)^j}{j!} \right] \quad (10)$$

The Poisson Surprise rate for a set of spikes is computed from the time span T of the set and the mean firing rate r , which is the number of spikes O in the set divided by T . The burst is detected by first finding a pair of successive spikes whose interspike interval (ISI) is less than the mean ISI of the spike train multiplied by a designated coefficient p ($0 < p < 1$). Subsequent spikes are added to the pair of spikes to formulate a spike set with the largest possible Poisson Surprise value, and the earliest spikes are pruned from the

set if that further increases the Poisson Surprise of the spike set. Finally, the spike set is regarded as a burst if it consists of at least 3 spikes.

With all burst occurrences detected throughout the spike train, we use them to calculate nine burst-related features for a particular neuron (Lei et al., 2011). A logistic regression is finally fitted with the nine burst-related features as covariates to classify the type of neurons as PN or LN.

We train a similar logistic regression model based on the spontaneous spike train obtained from intracellular recordings and staining in Lei et al. (2011), with a validation accuracy of around 85%. Then, using this logistic regression model, we classify the neurons collected through Section 2. During our initial data analysis, we found that the distribution of the average spike rate in the data from Lei et al. (2011) differs from that in our spike-sorted data. The difference in distribution resulted in scale differences in the 9 burst-related features (Supplementary Figure S5). To resolve the difference in the features, we applied the following processing steps:

1. We tune the p parameter in the procedure for detecting potential bursts to obtain burst-related features on a similar scale. We used $p = 0.2$ while $p = 0.5$ is defaulted in Lei et al. (2011). The p parameter defines the ratio between the mean spike rate (r_{mean}) and the spike rate of potential burst segments (r_{burst}) and classifies the segment as a burst when $r_{\text{mean}}/r_{\text{burst}} < p$.
2. We remove three of the nine features for which significant scale differences cannot be resolved by tuning p . The six features we used to classify the neuron types are the within-burst maximum spiking frequency, the within-burst number of spikes, the percentage of burst spikes, the burst frequency, the mean Poisson Surprise, and the maximum Poisson Surprise.
3. We apply two different min-max scalers to the training data (Lei et al., 2011) and the testing data (described in Section 2), respectively.

Once the neurons are classified, we supply our predicted labels to human experts to assist in the annotation of true neuron types. We refer the readers to Lei et al. (2011) for the details regarding the classification method.

3.6 Peristimulus time histograms

To see how the individual neurons responded to an odor, we compute a peristimulus time histograms (PSTHs) averaged across five trials for a given odorant stimulus, binned at 20 ms, and smoothed with a Gaussian of 3 standard deviations (Figures 1D, 5B). We generate the response maps of the odor-evoked population response by Equation 11. Here, we z -scored the PSTHs for each neuron by subtracting its average baseline firing rate (measured 200 ms before a stimulus onset) and dividing by the standard deviation of the baseline. The response maps indicated changes in firing rate in units of the standard deviation of spontaneous activity. The (Figures 1D, 5B) are used for visualization purposes only.

$$\text{PSTH}_{z(t)} = (\text{PSTH} - \mu_{\text{baseline}}) / \sigma_{\text{baseline}} \quad (11)$$

3.7 Response index

The Response Index was computed in our study to investigate the response of different units to the odor stimuli and to assess the similarity of generated and real spike trains under different stimuli for each unit (Riffell et al., 2009a). The response index is calculated as follows:

$$\text{RI}_{\text{odor}} = \frac{(r_{\text{odor}} - r_{\text{control}}) - r_{\text{mean}}}{SD}$$

where the r_{odor} is the firing rate of a specific odor; r_{control} is the firing rate of control (mineral oil); r_{mean} is the mean firing rate averaged over all stimuli, and SD is the standard deviation of the firing rate across all stimuli. Firing rate in this RI computation specifically means the average firing rate over the stimulation period (0–600 ms after the onset of stimuli). The response indexes of other preparations are shown in Supplementary Figure S3.

3.8 Euclidean distance

Euclidean distance is applied in our study to understand the difference between spatial-attention vectors of different stimuli. For two vectors $v_1, v_2 \in \mathbb{R}$ for an arbitrary dimension d , with v_{1i}, v_{2i} denoting the i^{th} element of the corresponding vector, the Euclidean distance is defined as

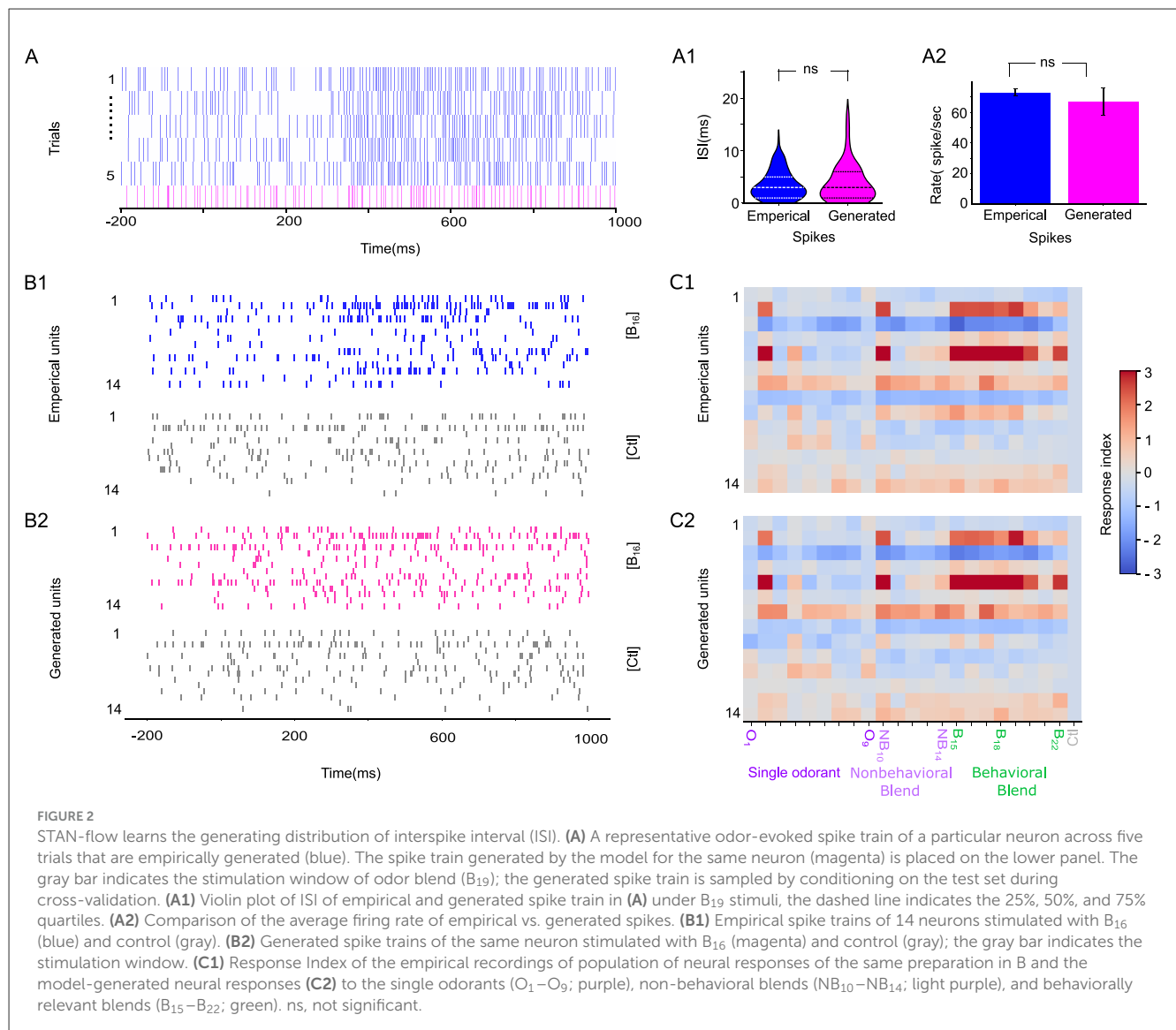
$$\text{Euclidean distance} = \sqrt{\sum_{i=1}^d (v_{1i} - v_{2i})^2} \quad (12)$$

4 Results

By leveraging well-known biological principles of the spiking and interaction between neurons in the primary olfactory center, we designed STAN-Flow to model the fundamental neuron spiking mechanism and neuron interactions in the AL (Figure 1F). We validate STAN-Flow from three different perspectives: (1) Can it generate realistic spike trains that replicate the statistical distribution of recorded data? (2) Can it distinguish between behaviorally relevant and irrelevant stimuli using spatial-attention weights that reflect neural synchronization patterns? (3) Can it infer interactions among distinct neuronal subtypes in the AL? After validating that the STAN-Flow effectively learns the spiking mechanism and network dynamics, we apply it to investigate whether synchronization is affected when component odorant concentrations diverge from those found in natural, behaviorally significant mixtures (e.g., floral blends that drive foraging behavior).

4.1 Antennal lobe network dynamics and spike train generations

One major aspect of validating the STAN-Flow architecture is examining how similar the generated spike trains are to the empirical neuron recordings. In Figure 2A, we present five trials of the spike train of a specific unit in response to a behavioral blend



odor (B₁₉) and a generated spike train given this odor. Visually, the generated spike train (magenta, **Figure 2A**), which is generated conditioned on data in the test set, realistically captures the temporal dynamics of the spiking pattern of spontaneous baseline activity, stimulation period, and post-stimulation corresponding to the empirical spike train.

We conclude that there are no differences in the empirical and generated distribution of ISI (Kolmogorov–Smirnov (KS) Test¹, $K = 0.08$, $p = 0.98$; Two One-Sided *T*-Test (TOST) for distribution mean² (Lakens, 2017) ($t = 1.29$, $p = 0.09$). We also found no significant differences in the average firing rate of a neuron for empirical and generated spike trains (two-sided *t*-test, $t = 0.65$, $df=8$, $p = 0.53$), indicating that the generated spikes are similar to the empirical ones (see **Figures 2A1, A2**).

We also examine the generated spike trains across an ensemble. We present the empirical and generated spike trains of 14 units stimulated with behavioral odor B₁₆ and control, respectively (**Figures 2B1, B2**). Surprisingly, the response index (RI; **Figures 2C1, C2, Supplementary Figure S3**) shows strong correspondence between empirical and STAN-Flow-generated spike trains across all neurons and stimuli. This close match indicates that STAN-Flow effectively captures the average spiking patterns observed in empirical recordings. The consistency of this alignment across multiple neurons and stimuli further highlights STAN-Flow's state-of-the-art capability in modeling the spiking dynamics of neurons within the antennal lobe (AL) region.

4.2 Spatial-attention weights classify stimuli

The AL relies on different cell pairs synchronizing with one another, or a specific subset of critical neurons, to encode both behavioral and non-behavioral stimuli (Riffell et al., 2009a; Lei

¹ The KS-Test will reject the null hypothesis if two distributions are statistically different.

² Also known as the test of equivalence for two independent samples; the TOST will reject the null hypothesis if the means of the two distributions are within the $\pm\Delta$ region. We choose $\Delta = 2$ ms.

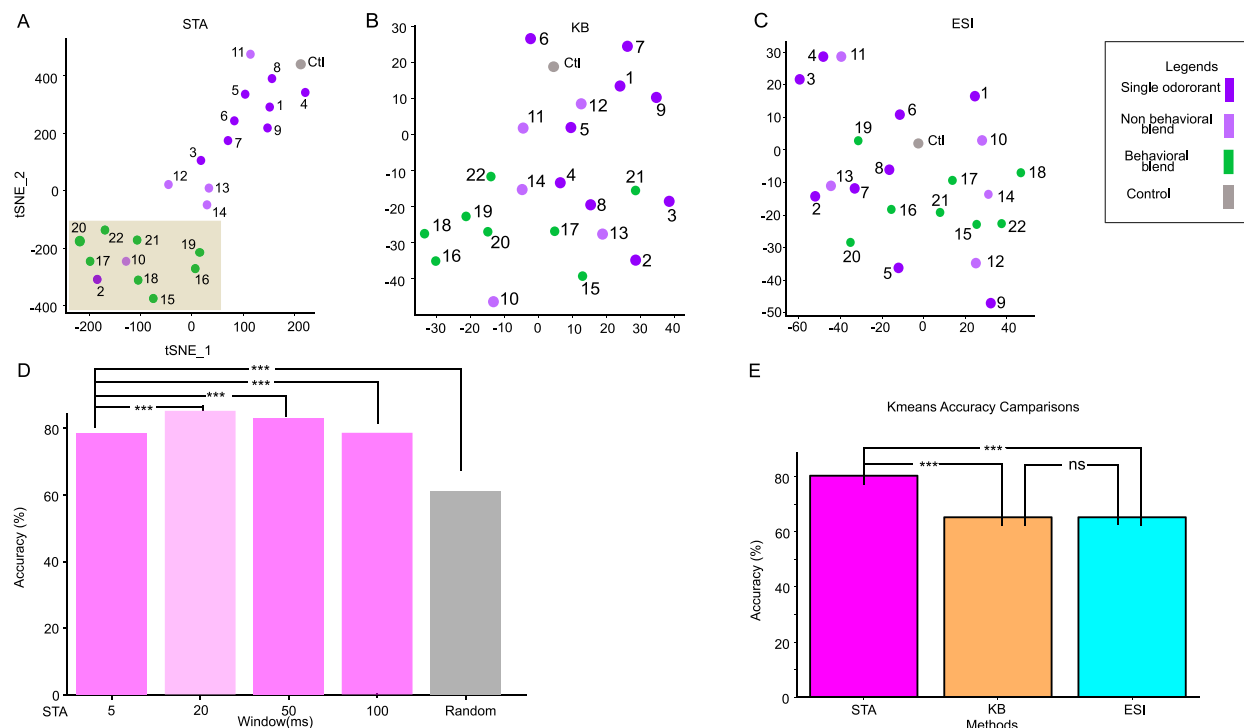


FIGURE 3

Clustering behavioral and non-behavioral stimuli through different synchronization methods. (A) spatial-attention weights separate the behavioral stimuli (green) from the non-behavioral stimuli (purple) by forming two distinctive clusters in the 2D TSNE-reduced space. The behavioral stimuli cluster is highlighted with a brown rectangle. (B, C) 2D TSNE plot of Binless synchronization (KB) and Ensemble Synchronization Index (ESI), respectively. (D) The bargraph represents the accuracy of the 2-class K-means algorithm across various window sizes. The window used in our analysis is shown against the light green background. While STA(20ms) is significantly better than other window sizes, the STA generally outperforms traditional methods. (E) The bargraph represents the 2-class K-means algorithm with the spatial-temporal attention (STA) method, repeated with 100 different initializations. K-means accuracies of the Kernel Binless (KB) and Ensemble Synchronization index (ESI) are significantly less accurate than STA. *** p -value < 0.001 and ns is non-significant.

et al., 2016). Our method captures this synchronization mechanism through the spatial attention module. The spatial-attention module assigns an attention weight to each of the units of the spike train input. Each row of the synchronization summary matrix $\mathcal{B}$ (see Equation 9) represents a stimulus, and the columns represent the relative importance of each unit. The spatial-attention method takes all neurons into account during the modeling of spike trains, thus offering the ability to characterize the synchronization of multiple neurons beyond pairwise analyses.

We first apply t-distributed Stochastic Neighbor Embedding (TSNE) to reduce the dimensionality of $\mathcal{B}$. Our result demonstrates that the spatial-attention weights distinctly separate the stimuli into two clusters: one includes all the stimuli mixtures that contain all the behavioral components (green), and the other cluster includes the individual odor molecules (purple) as well as mixtures containing non-behavioral components (light purple) (Figure 3A, Supplementary Figure S4). Two exceptions are grouped with the behavioral relevant stimuli space: the single odor (Benzylalcohol: O_2) and non-behavioral blend (NB₁₀ (mixture containing (O_1 and O_2)): Riffell et al., 2009a). These results suggest that O_1 and O_2 could be essential to the behavioral responsiveness of complex odors. To ensure reproducibility of our result, we retrained STAN-Flow with the same configuration using 100 different random initializations, then applied a 2-class K-means algorithm

(behavioral vs. non-behavioral), STA yields around 80 (± 4.5)% classification accuracy.

We next compute the pair-wise synchronization matrix for two previously known methods: the Ensemble Synchronization Index (ESI) (Riffell et al., 2009a) and the Kernelized binless method (KB) (Martin et al., 2013). We first extract the upper triangular synchronization matrix of ESI and KB, then apply TSNE to reduce its dimension. We do not observe obvious patterns of clustering in these methods (Figures 3B, C). Compared to the STA results (STAN-Flow trained with 100 different initializations), both of these methods yield around 20% less accuracy (One-sided z-test; ESI: $z = 4.67$, p -value < 0.01; KB: $z = 4.67$, p -value < 0.01), indicating that STA consistently outperforms ESI and KB (Figure 3E). As ESI and KB were previously only applied to populations of projection neurons (Riffell et al., 2009a; Martin et al., 2013), we conjectured that ESI and KB failed to separate the stimulus types in this neural population because both LNs and PNs are present. This enhanced performance in separating the behavioral and non-behavioral stimuli when different types of neurons exist in the neuron population propels us to understand how the spatial-attention models the interaction of different neuron categories in the AL.

We also compare the effect of different window sizes [Δ] in extracting neural synchronization. Similar to our comparison

before, we apply TSNE on the STA generated from models trained with different window sizes, then apply K-means on the dimension-reduced STA and compute K-means accuracy (Figure 3D, highlighted with light green). We found that while a 20 ms window performs the best, synchronization can be captured with a window size as low as 5 ms. STAN-flow trained a wide range of window sizes continue to outperform traditional methods, suggesting that STA is a robust tool to measure synchronization at any level.

4.3 Detecting the interaction between PNs and LNs

The antennal lobe consists of broadly two types of neurons, PNs and LNs. As Lei et al. (2004, 2016) and Tanaka et al. (2009) suggested, LNs modulate the PNs synchronization, and the PNs synchronize among themselves (Martin et al., 2013) to encode behavioral and non-behavioral stimuli. While STAN-Flow is trained without the knowledge of neuron types, we investigate if STAN-Flow can recover the interaction between LN and PNs through modeling the spiking process. Hence, we need to label the neurons to their corresponding neuron types. However, with the electrophysiological recordings in the AL with the current method, we cannot morphologically classify PNs and LNs. To classify the neurons, we utilize previous results that found the PNs and LNs have distinctively different spontaneous patterns of spiking activities (Figures 4A, B) (Lei et al., 2011). The PNs burst from time to time during spontaneous baseline firing, while the LNs spike regularly. Based on this observation, Lei et al. (2011) developed a simple method to classify the neurons based on their spontaneous spiking dynamics.

We refer readers of the details of this classification process to Section 3.5. The classification accuracy computed against expert label was 78% while the validation accuracy using data from Lei et al. (2011) is about 75%. Since there could be more PNs than LNs in the AL region of the moth (Homberg et al., 1988; Reisenman et al., 2011), we also compute the recall (0.73) and precision (0.73), showing that the classification method obtained a reliable result under this imbalance classification scenario.

With the neuron labels, we continue to study the essential question in this analysis to understand the interaction between the PNs and LNs. To better understand the functional roles of local neurons (LNs) and projection neurons (PNs), we trained our model on the full population of recorded neurons, then separately visualized the synchronization learned by STA of PNs (Figures 4C, C1) and LNs (Figures 4D, D1) using 2D TSNE. The PNs' embeddings showed clear clustering of behavioral vs. non-behavioral stimuli (highlighted with a gray background rectangle, while the LNs lacked any discernible pattern. When we trained the model using only the annotated PNs or LNs, the PNs alone still exhibited visible clusters corresponding to behavioral stimuli, with partial overlap from non-behavioral stimuli, suggesting inherent synchronization among PNs. One notable feature is that the individual odors are distinctly cluster (Figure 4D). Importantly, training on PNs extracted from the full model resulted in a mean classification accuracy approximately 12.5% higher than training

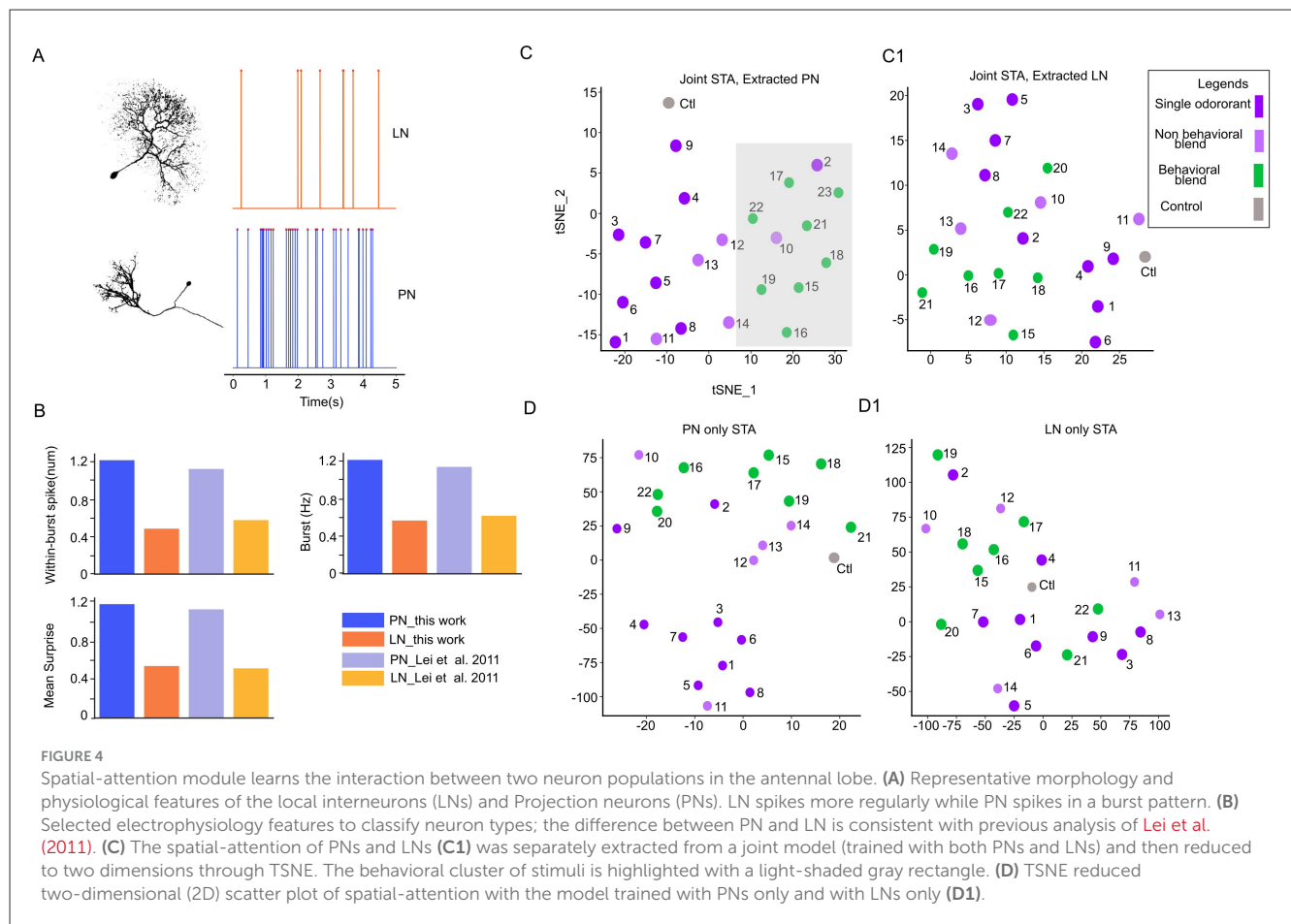
on PNs alone (one-sided z -test, $z = 20.77$, p -value < 0.001). This enhancement suggests that the optimal stimulus separability may be modulated by LNs, which could suppress responses to non-behavioral stimuli. Collectively, these results highlight the critical role of LNs in refining PN synchronization and emphasize their importance in the encoding of behavioral stimuli within the AL and the STAN-Flow model.

Our previous result has highlighted that LNs are critical for odor classification, but could there be a core neuronal unit in an ensemble that accounts for the segregation of behavioral and non-behavioral stimuli? As we showed that spatial-attention consistently clusters the behavioral and non-behavioral odors, we now test the K-means clustering accuracy by removing each unit in an ensemble. We found one LN significantly lower than the classification accuracy (86% before removal vs. 78% after removal). This neuron responds to both behavioral compounds and non-behavioral compounds (Figure 2C1), unit 14. In addition to single neuron analyses, we removed a combination of up to 3 neurons (data not shown) in an ensemble. We found a specific combination including the LN (Figure 2C1, unit 14; Supplementary Figure 3) and two other PNs (units 2 and 5, Figure 2C1; Supplementary Figure 6) will lower the clustering accuracy to 52%. These PNs responded to the majority of the behavioral compounds (broadly tuned). These results also further indicate that broadly tuned LNs could be core neurons for odor classification (Wilson et al., 2004; Olsen et al., 2010) in combination with the PNs, which are responsive to the behavioral compound (Figure 2C1).

4.4 Altering the excitatory drive attenuates synchronization

STAN-Flow preserves the fundamental spiking dynamics and functional architecture of the AL by capturing the synchronization patterns of projection neurons to the behavioral mixture, including behaviorally relevant odor diluted by several magnitudes (Figure 3A, odors (B₂₀, B₂₂, and B₂₂) in the behavioral clusters). To test the robustness of our model in capturing this synchrony, we increased the concentration of a single component, benzaldehyde (O₁), within the behavioral odor mixture B₂₃ to 10-fold (B₂₄), 100-fold (B₂₅), and 1,000-fold (B₂₆). Visually, the stimulus concentration-dependent response was prominent. The PSTHs show the sustained activity after the stimulation, forming a plateau-like PSTH to the increased O₁ ratio as compared to the original B₂₃, indicating the weak or delayed inhibitory input to this neuron (Figures 5A, B, E). These dose-dependent increases in the O₁ ratio led to a transient suppression in neuronal response patterns, which may be due to the inhibition of inhibitory neurons involved in encoding odor identity/intensity (Figures 5B, E).

We retrained STAN-Flow using neural activity. We then repeated the TSNE analysis on the summary matrix (see Figure 5C). The TSNE clustered individual components toward the top-right quadrant, grouped the behavioral mixture B₂₃ variants with elevated O₁ concentrations centrally, and positioned the control stimuli to the far left. Also, the response index of the generated heatmap (Figure 5D, Supplementary Figure S6) is similar to the empirical response index.



To evaluate how ensemble synchronization changes with increased O_1 ratio in the mixture, we calculated the dissimilarity of odor-evoked synchrony patterns using a normalized dissimilarity index (normalized by the mean, moth = 3, units = 43). This analysis compared B_{23} with $10\times$, $100\times$, $1000\times$ increase of O_1 (B_{24} – B_{26}) and also with individual odorants. As shown in [Figure 5F](#), there is an increasing trend of the dissimilarity with B_{23} , but it is not statistically significant among the group (one-way ANOVA, $F = 0.642$, $p = 0.673$). These findings could suggest that elevated O_1 concentrations in the behavioral components may shift the neural representation toward a non-behavioral classification.

Our results underscore the role of synchrony in the AL as a critical coding mechanism for odor classification. This synchrony is finely tuned to both the identity and the concentration of odor components, enabling flexible and accurate recognition of behaviorally relevant odors in dynamic natural environments.

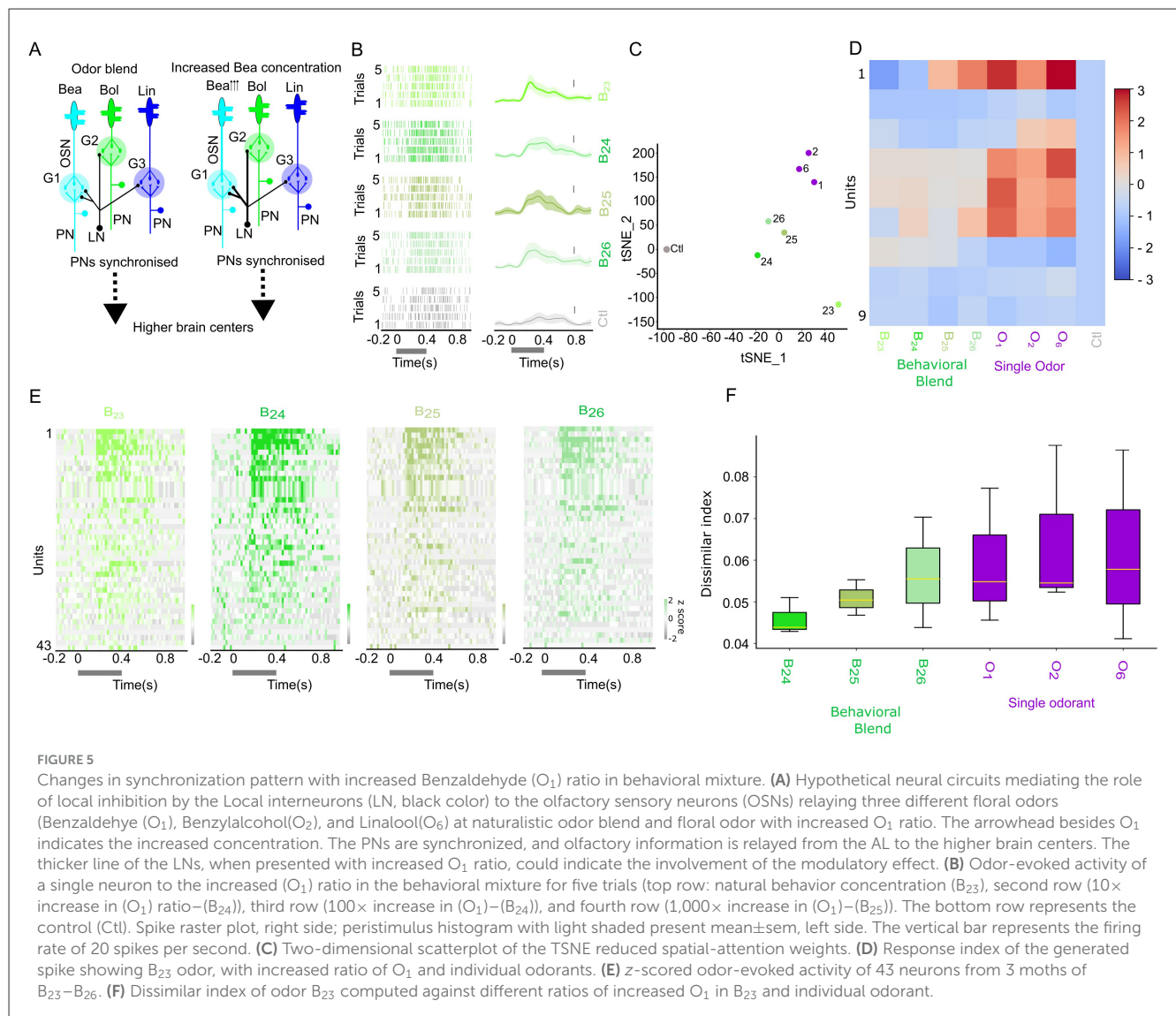
5 Discussion

Here in the current study, we elucidated the effects of network dynamics in the antennal lobe (AL) to classify behavioral and non-behavioral relevant odors. We developed the spatial-temporal attention normalizing flow (STAN-Flow), an accurate computational model representing the spiking ensemble dynamics of the AL. We adopt this model to extend the characterization

of the AL network beyond the experimental possibilities. The model effectively reproduced the AL responses' key features concerning the odor classification, through detailed neuron-level interactions. This model also agrees with the previous results that local interneurons play critical roles in the temporal encoding of odor stimuli, enabling the classification of odors into behavioral and non-behavioral stimuli. Shifting the concentration of one of the behavioral components in the odor mixture—by altering the excitation/inhibition balance in the AL—causes the neuronal representation of this stimulus to change. This computational model can be easily modified to be applied in various fields for accurately modeling and reliably interpreting complex interactions for biological and non-biological systems.

5.1 Neurophysiological computation in the AL

The AL is one of the most extensively studied neural structures in the insect brain, particularly in terms of its detailed cellular and circuit-level architectures for sensory encoding. Over the last two decades, behavioral, physiological, and modeling research have made significant strides in understanding the circuit basis of processing complex odor mixtures, their intensities, and their relationship to odor classification. Understanding the role of



the AL in odor perception has been the focus of a variety of theoretical and computational models. The dynamic and complex stimuli necessitate utilizing the computational models to extract the features of interest from the spike trains (Triplett and Goodhill, 2019). The STAN-Flow developed in this study serves this purpose: its flexibility simulates the spike train generation process and successfully discriminates and classifies complex odors. This model could be beneficial in identifying future odors, whether or not they could be relevant to insects, predicting the population response, and simulating the spike trains of the neurons.

This computational model can cluster the odors into behaviorally relevant and non-behaviorally relevant groups (Figures 3, 5). Our approach enhances clustering into behavioral and non-behavioral odor stimuli, and it is highly efficient. It could potentially facilitate further processing in the higher brain centers such as the lateral horn (LH) (Lazar et al., 2023; Strutz et al., 2014). Given the extensive knowledge of neuromorphic processing in downstream neurons from the AL and its circuits, it remains

unknown to date how this spatiotemporal information is processed within the AL and in higher brain centers.

5.2 Local interneurons necessitate the PNs' synchronous activity

Olfactory information is encoded as spatial-temporal patterns in the neural population of the AL, and through the activity of different cell types, such as LNs and PNs. A benefit of the model and resulting analyses is a dissection of the contribution of different cell types in how the complex odor stimuli are classified into behavioral and non-behavioral (Figures 3, 5). However, the glomerular processing of the behaviorally relevant odor is putatively identified in the anterior lateral regions of ordinary glomeruli in the AL investigated through measuring the calcium activity of sensory neurons in the moth (Figure 1 in Bisch-Knaden et al., 2018; Bisch-Knaden et al., 2022). It appears that the region

processing the three components of behaviorally relevant odors is colocalized in a similar position within the ordinary glomeruli in the AL. However, the PNs and LNs involved in these colocalized glomeruli are not known, even though diverse morphological PNs and LNs are reported in different insect species (Kymre et al., 2021; Matsumoto and Hildebrand, 1981; Kuebler et al., 2012; Reisenman et al., 2011; Chu et al., 2020). Future experiments with simultaneous recording from the selective recordings will answer the neural circuit involving such behavioral valence. Unlike some specific glomeruli (particularly pheromone processing glomeruli), which may not require specific connectivity between the glomeruli (or sparse connection) for the PN response, some floral odor processing glomeruli (ordinary glomeruli) and PNs arising from these glomeruli may involve the LNs that have heterogeneous branching pattern in these glomeruli and could channel to the behavioral relevance odor. Of 1,100 PNs, we were only able to identify 78% of the PNs. Even with extensive sampling of LNs and PNs through intracellular recordings and staining in the other moth species (*Helicoverpa armigera*: out of 176 reported neurons, 61% are PNs (Kymre et al., 2021)). Our method also relies on validation techniques from intracellular recordings of neuron categories in the AL of the *Manduca sexta* (Lei et al., 2011), and the number appears to be comparable. We have not subcategorized the different subclasses of LNs and PNs due to a lack of staining data. The existence of subcategories of the PNs with the distinct physiological properties during the spontaneous (Capurro et al., 2014) and after the stimulus offset and explain why we cannot achieve 100 percent accuracy. In line with other studies on response properties of PNs on the fruit flies (Wilson et al., 2004). The majority of the PN responses are dynamic under different odors; some are activated, and some are inhibited with no responses (Figure 2C, Supplementary Figure S3) due to the interaction of the olfactory sensory neurons (OSNs) and LNs. The excitatory feedforward information from OSNs activates PNs, which in turn indirectly activate LNs.

The inhibitory LN connecting the defined subsets of the glomeruli could play a crucial role in understanding perceptual constancy in the olfactory circuits, especially in understanding the synchronized activity of the PNs in response to the behaviorally relevant mixture. Previous studies have shown that pharmacological receptor antagonists, targeting the GABA receptors, abolished the synchronized activity of AL neurons and affected the olfactory behavior of the moths (Lei et al., 2011; Riffell et al., 2014). The LNs modulate the temporal patterns of PNs' spiking responses, resulting in odor-evoked activity which can enhance the synchrony of sister PNs within the same glomerulus as well as the synchrony of co-activated PNs from the other glomeruli (Lei et al., 2002; Martin et al., 2013). However, the structural connectivity information is still lacking in *Manduca sexta*, which limits our understanding of the synaptic-level connections between neurons. Electron microscopical studies in fruit flies have shown that the OSNs contribute 75% of the synaptic input to PNs, and the remaining 25% is contributed by the LNs (Tobin et al., 2017). It is possible that the LN that makes specific synaptic connections with a given glomeruli could provide the postsynaptic inhibition to the PNs that are processing the non-behavioral relevant odors and receiving information from the subset of behaviorally relevant

neurons. This can be ecologically relevant in inhibiting the input of the non-behavioral stimulus pathway, as it may be the background. In various organisms, such as moths and fruit flies, the LNs contain both pre- and postsynaptic synapses, and the density of these synapses is biased across different glomeruli. This bias could eventually affect the extent of lateral inhibition processing in the mixture (Hong and Wilson, 2015; Silbering and Galizia, 2007).

5.3 Concentration varying effects on AL network dynamics

Navigating through the complex and dynamic olfactory environment, the moth is challenged with fluctuating odor concentrations. The moth should evaluate the odor and its intensity. Behaviorally, *Manduca* has been shown that a subset of the behaviorally relevant odorants is processed in a quick (<500 ms) and reliable manner (Riffell et al., 2009b). Suppose if the ratio of one of these compounds is changed in the behavioral mixture, then they could be evaluated as a different floral compounds and the neural population responses could vary for those altered composition and could be clustered outside the neighborhood of the behavioral mixture (Figures 3A, 5C) (Riffell et al., 2009a, 2014). The behavioral compound such as Benzaldehyde (O_1) which is one of the important constituents in the behavioral mixture is clustered with the non-behavioral mixture (NB_{10}) that contains Benzaldehyde (O_1) and Benylalcohol (O_2) could play a significant role in the discrimination of odors in behavioral or non-behavioral and thus affect the olfactory navigation.

The decreasing intensity of the floral mixture, diluted upto 10,000-fold clustered with the behavioral mixture in the previous study (Riffell et al., 2009a) and this study (Figure 3) could be due to gain control of the LNs (Sachse and Galizia, 2002). Altering the ratio of (O_1) in the behavioral floral mixture (B_{23}), we noticed that different ratios are clustered outside the neighborhood of the B_{23} . The concentration of the same odor may have different or even opposite values (Sammelhack and Wang, 2009), and this odor could modify the quality of the odor (Laing et al., 2003). At the presentation of increased ratio of an odorant in the behavioral floral odor, the LNs may modulate (Supplementary Figure S5) and fine-tune the PNs for synchronization, therefore clustering the neural responses into the behavioral and non-behavioral components (Clifford and Riffell, 2013). However, upon increasing the ratio of one of the components, the GABAergic neurons may be recruited (Laurent, 2002) thus the presynaptic inhibition of sensory neurons by these neurons (disinhibition) could play a role in odor discrimination. However, how the modification to the odor panel influences the molecular mechanism or alters the plasticity within the AL remains unknown. In our experiments, we present moths to an elevated ratio of benzaldehyde within the natural concentration range. In *Datura wrightii* flowers, benzaldehyde is emitted at about 0.23 ng/h (Riffell et al., 2009a). Increasing this concentration by the 1,000-fold (230 ng/h) still falls below the natural emission rate observed in other putative hawkmoth-visited flower, such as *Petunia hybrida*: 8,000 ng/h (Boatright et al., 2004) and in blossoms and branches of Crabapple (*Malus sp*: 5,580 ng/h)

(Baghi et al., 2012). These comparisons indicate that the moths in our study were stimulated within ecologically relevant, naturally occurring ranges.

5.4 Limitations

There remain a few limitations in the STAN-Flow architectures and related analysis. One main limitation is the interpretability of the neural networks. Although through post hoc analysis, we showed that the spatial-attention weights can be interpreted to a great extent, there are limited theoretical analyses on the attention module to guarantee the interpretability of the attention weights. There are, however, ways to improve the credibility of the result: one can cross-validate with established results to check if the interpretability matches expectations, as we did with PN and LN classification. In addition, one could repeat the experiments many times (e.g., 100 different initializations) and across subjects to verify the consistency of the result. Another drawback of this current architecture is its scalability. The current training scheme of STAN-Flow for performing group-wise synchronization analysis requires neuron-specific STAN-Flow. While the number of models scales linearly with the number of interacting neurons, it remains challenging to analyze a larger population of neurons at a greater scale. Given different modeling scenarios, however, this could be resolved by a more intricate STAN-Flow that can be modified to model synchronization hierarchically when dealing with neurons from different regions; it could also be easily adapted to model a multidimensional time series instead of modeling each neuron at a time. Finally, the recent rise of diffusion generative modeling techniques can be applied to improve the normalizing flow (Song et al., 2020; Hasan et al., 2023). There also exist ways to directly connect the interaction between neurons with the generative component. For example, the interaction between neurons can be modeled as interacting particle systems characterized by a McKean–Vlasov diffusion, which can be applied as a latent process to significantly enhance the interpretability of the machine learning system (Yang et al., 2024). Despite these limitations, our machine learning approach offers a broader impact.

5.5 Significance and broader impact

Our STAN-Flow explores a new, computational avenue to study interactions within neural systems. The interaction of multiple sensory systems is common in many biological organisms, allowing them to respond swiftly and efficiently to complex and dynamic environments. For instance, in other animals and humans, multiple peripheral, central, and motor systems work in concert to produce coordinated behaviors. One prime application of the spatial-temporal attention module, a significant focus of our ongoing research, is the modeling of multisensory binding in the brain that is funneled downstream via the descending neurons to the motor program. The neural mechanism of integrating multisensory information in the brain

is not well understood, and how it drives the motor program is still under investigation. The motor programs involve the coordinated activity of individuals and groups of flight muscles that interact dynamically to produce agile movements and abrupt changes in behavior (Putney et al., 2019). Despite extensive research, the interactive mechanisms that govern this muscle dynamics remain largely unknown (Putney et al., 2023). The spatial-temporal attention module has the potential to uncover these mechanisms by providing a framework that captures the intricate timing and spatial relationships involved in motor coordination.

Another significant benefit of the STAN-Flow model is its deep generative component. Biological interactions are inherently stochastic and often do not conform to traditional statistical distributions such as the Poisson or Gaussian distributions (Lindner, 2006; Deger et al., 2012). This stochastic nature presents challenges for conventional modeling approaches that rely on these distributions. The introduction of a flexible generative model through neural networks, as seen in STAN-Flow, allows for more accurate modeling of natural phenomena with fewer assumptions about the underlying distributions. This flexibility is particularly important when exploring complex biological interactions that may be high-dimensional and highly non-linear. By leveraging the power of neural networks, STAN-Flow can capture the rich and varied nature of biological data, providing deeper insights into the interactions between multiple brain regions or systems.

STAN-Flow's success in modeling the antennal lobe region has illuminated the potential of combining attention mechanisms with deep generative neural networks to understand the complex interactive relationships between organisms and their environments. The ability of STAN-Flow to accurately model the dynamic and non-linear interactions in the AL region suggests that similar approaches could be applied to both biological and non-biological systems. This attention module could be applied to other sensory systems, such as the visual and auditory systems, for stimulus discrimination (Wilsch et al., 2020). However, whether one of the systems favors one of the modules (spatial or temporal) is elusive. This opens up new avenues for research in understanding how different neural systems interact and adapt to their environments, ultimately contributing to a more comprehensive understanding of biological complexity and adaptability. Apart from its biological context, STAN-Flow can also be applied in the digital field (Gao et al., 2023). The STAN-Flow can be naturally applied to related tasks such as analyzing videos and speech by summarizing the interaction of different graphic regions and condensing the importance of different periods of videos. By enforcing a set of constraints on the attention weights, it is also possible to extend the module to track objects through space and time. As many climatic phenomena also originate from interactions of local climate (Saupe et al., 2019), STAN-Flow also provides a generative predictive algorithm that allows explicit interaction between local climates to forecast weathering trends in the future. In general, the flexible, semi-interpretable neural network structure of STAN-Flow offers a wide range of applications that can help inform scientists with high-order interactions between groups or individuals over time.

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.

Ethics statement

The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

HY: Writing – original draft, Formal analysis, Conceptualization, Visualization, Methodology, Data curation, Investigation, Validation, Writing – review & editing. PK: Writing – review & editing, Methodology, Investigation, Writing – original draft, Conceptualization, Formal analysis, Visualization, Data curation, Validation. PC: Investigation, Writing – review & editing, Data curation, Writing – original draft, Conceptualization, Validation, Visualization, Methodology, Formal analysis. HL: Formal analysis, Visualization, Data curation, Validation, Methodology, Writing – review & editing, Conceptualization, Investigation, Writing – original draft. SS: Writing – review & editing, Supervision, Writing – original draft, Conceptualization, Investigation, Funding acquisition, Validation, Visualization. VT: Validation, Methodology, Writing – review & editing, Supervision, Formal analysis, Funding acquisition, Investigation, Visualization, Conceptualization, Writing – original draft. JR: Funding acquisition, Writing – review & editing, Conceptualization, Investigation, Visualization, Validation, Methodology, Formal analysis, Supervision, Data curation, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This project was funded by the Air Force Office of Scientific Research under grants FA9550-22-1-0315 (SS, VT, and JR), FA9550-20-1-0422 (JR), and

FA9550-21-1-0101 (JR); the National Science Foundation under IOS-2124777 (JR) and an Endowed Professorship for Excellence in Biology (JR). The funders had no role in study design, data collection and analysis, publication decision, or manuscript preparation.

Conflict of interest

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

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2025.1655462/full#supplementary-material>

References

- Baghi, R., Helmig, D., Guenther, A., Duhl, T., and Daly, R. (2012). Contribution of flowering trees to urban atmospheric biogenic volatile organic compound emissions. *Biogeosciences* 9, 3777–3785. doi:10.5194/bg-9-3777-2012
- Bahdanau, D., Cho, K., and Bengio, Y. (2016). Neural machine translation by jointly learning to align and translate. *arXiv [preprint]*. arXiv:1409.0473. doi: 10.48550/arXiv.1409.0473
- Biedlerlack, J., Castelo-Branco, M., Neuenschwander, S., Wheeler, D. W., Singer, W., and Nikolić, D. (2006). Brightness induction: rate enhancement and neuronal synchronization as complementary codes. *Neuron* 52, 1073–1083. doi: 10.1016/j.neuron.2006.11.012
- Bisch-Knaden, S., Dahake, A., Sachse, S., Knaden, M., and Hansson, B. S. (2018). Spatial representation of feeding and oviposition odors in the brain of a hawkmoth. *Cell Rep.* 22, 2482–2492. doi: 10.1016/j.celrep.2018.01.082
- Bisch-Knaden, S., Rafter, M. A., Knaden, M., and Hansson, B. S. (2022). Unique neural coding of crucial versus irrelevant plant odors in a hawkmoth. *Elife* 11:e77429. doi: 10.7554/eLife.77429.sa2
- Bitzenhofer, S. H., Westeinde, E. A., Zhang, H.-X. B., and Isaacson, J. S. (2022). Rapid odor processing by layer 2 subcircuits in lateral entorhinal cortex. *Elife* 11:e75065. doi: 10.7554/eLife.75065.sa2
- Boatright, J., Negre, F., Chen, X., Kish, C. M., Wood, B., Peel, G., et al. (2004). Understanding *in vivo* benzenoid metabolism in petunia petal tissue. *Plant Physiol.* 135, 1993?2011. doi: 10.1104/pp.104.045468
- Capurro, A., Baroni, F., Kuebler, L. S., Kárpáti, Z., Dekker, T., Lei, H., et al. (2014). Temporal features of spike trains in the moth antennal lobe revealed by a comparative time-frequency analysis. *PLoS ONE* 9:e84037. doi: 10.1371/journal.pone.0084037

- Carlsson, M. A., Chong, K. Y., Daniels, W., Hansson, B. S., and Pearce, T. C. (2007). Component information is preserved in glomerular responses to binary odor mixtures in the moth *spodoptera littoralis*. *Chem. Senses* 32, 433–443. doi: 10.1093/chemse/bjm009
- Christensen, T. A., and Waldrop, B. R. Hildebrand, J. G. (1987). Pheromonal information coding by projection neurons in the antennal lobes of the sphinx moth *manduca sexta*. *Ann. N.Y. Acad. Sci.* 510, 224–226. doi: 10.1111/j.1749-6632.1987.tb43511.x
- Chu, X., Kc, P., Ian, E., Kvello, P., Liu, Y., Wang, G., et al. (2020). Neuronal architecture of the second-order co2 pathway in the brain of a noctuid moth. *Sci. Rep.* 10:19838. doi: 10.1038/s41598-020-76918-1
- Clifford, M. R., and Riffell, J. A. (2013). Mixture and odorant processing in the olfactory systems of insects: a comparative perspective. *J. Comp. Physiol. A* 199, 911–928. doi: 10.1007/s00359-013-0818-6
- Deger, M., Helias, M., Boucsein, C., and Rotter, S. (2012). Statistical properties of superimposed stationary spike trains. *J. Comput. Neurosci.* 32, 443–463. doi: 10.1007/s10827-011-0362-8
- Delahunt, C. B., Riffell, J. A., and Kutz, J. N. (2018). Biological mechanisms for learning: a computational model of olfactory learning in the *manduca sexta* moth, with applications to neural nets. *Front. Comput. Neurosci.* 12:102. doi: 10.3389/fncom.2018.00102
- Dinh, L., Sohl-Dickstein, J., and Bengio, S. (2016). Density estimation using real NVP. *arXiv [preprint]*. arXiv:1605.08803. doi: 10.48550/arXiv.1605.08803
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2021). An image is worth 16x16 words: transformers for image recognition at scale. *arXiv [preprint]*. arXiv:2010.11929. doi: 10.48550/arXiv.2010.11929
- Fiscella, M., Franke, F., Farrow, K., M?ller, J., Roska, B., Azeredo Da Silveira, R., et al. (2015). Visual coding with a population of direction-selective neurons. *J. Neurophysiol.* 114, 2485–2499. doi: 10.1152/jn.00919.2014
- Gao, D., Zhou, L., Ji, L., Zhu, L., Yang, Y., Shou, M. Z., et al. (2023). “Mist: multi-modal iterative spatial-temporal transformer for long-form video question answering,” in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Vancouver, BC: IEEE), 14773–14783. doi: 10.1109/CVPR52729.2023.01419
- Guerrieri, F., Schubert, M., Sandoz, J.-C., and Giurfa, M. (2005). Perceptual and neural olfactory similarity in honeybees. *PLoS Biol.* 3:e0030060. doi: 10.1371/journal.pbio.0030060
- Haggard, M., and Chacron, M. J. (2025). Nonresponsive neurons improve population coding of object location. *J. Neurosci.* 45:e1068242024. doi: 10.1523/JNEUROSCI.1068-24.2024
- Hasan, A., Chen, Y., Ng, Y., Abdelghani, M., Schneider, A., Tarokh, V., et al. (2023). “Inference and sampling of point processes from diffusion excursions,” in *UAI (PMLR)*, 839–848.
- Heinbockel, T., Christensen, T., and Hildebrand, J. (2004). Representation of binary pheromone blends by glomerulus-specific olfactory projection neurons. *J. Comp. Physiol. A* 190, 1023–1037. doi: 10.1007/s00359-004-0559-7
- Heinze, S., El Jundi, B., Berg, B. G., Homberg, U., Menzel, R., Pfeiffer, K., et al. (2021). A unified platform to manage, share, and archive morphological and functional data in insect neuroscience. *eLife* 10:e65376. doi: 10.7554/eLife.65376
- Homberg, U., Montague, R., and Hildebrand, J. (1988). Anatomy of antenno-cerebral pathways in the brain of the sphinx moth *manduca sexta*. *Cell Tissue Res.* 254, 255–281. doi: 10.1007/BF00225800
- Hong, E. J., and Wilson, R. I. (2015). Simultaneous encoding of odors by channels with diverse sensitivity to inhibition. *Neuron* 85, 573–589. doi: 10.1016/j.neuron.2014.12.040
- Hopfield, J. J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature* 376, 33–36. doi: 10.1038/376033a0
- Hoskins, S. G., Homberg, U., Kingan, T. G., Christensen, T. A., and Hildebrand, J. G. (1986). Immunocytochemistry of gaba in the antennal lobes of the sphinx moth *manduca sexta*. *Cell Tissue Res.* 244, 243–252. doi: 10.1007/BF00219199
- Kobyzev, I., Prince, S. J., and Brubaker, M. A. (2021). Normalizing flows: an introduction and review of current methods. *IEEE Trans. Pattern Anal. Mach. Intell.* 43, 3964–3979. doi: 10.1109/TPAMI.2020.2992934
- Kuebler, L. S., and Schubert, M. Kárpáti, Z., Hansson, B. S., Olsson, S. B. (2012). Antennal lobe processing correlates to moth olfactory behavior. *J. Neurosci.* 32, 5772–5782. doi: 10.1523/JNEUROSCI.6225-11.2012
- Kymre, J. H., Berge, C. N., Chu, X., Ian, E., and Berg, B. G. (2021). Antennal-lobe neurons in the moth *helicoverpa armigera*: morphological features of projection neurons, local interneurons, and centrifugal neurons. *J Comp Neuro* 529, 1516–1540. doi: 10.1002/cne.25034
- Laing, D. G., Legha, P., Jinks, A. L., and Hutchinson, I. (2003). Relationship between molecular structure, concentration and odor qualities of oxygenated aliphatic molecules. *Chem. Senses* 28, 57–69. doi: 10.1093/chemse/28.1.57
- Lakens, D. (2017). Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *SPSS* 8, 355–362. doi: 10.1177/1948550617697177
- Laurent, G. (2002). Olfactory network dynamics and the coding of multidimensional signals. *Nat. Rev. Neurosci.* 3, 884–895. doi: 10.1038/nrn964
- Laurent, G., Wehr, M., and Davidowitz, H. (1996). Temporal representations of odors in an olfactory network. *J. Neurosci.* 16, 3837–3847. doi: 10.1523/JNEUROSCI.16-12-03837.1996
- Lazar, A. A., Liu, T., and Yeh, C.-H. (2023). The functional logic of odor information processing in the drosophila antennal lobe. *PLoS Comput. Biol.* 19:e1011043. doi: 10.1371/journal.pcbi.1011043
- Legendy, C., and Salzman, M. (1985). Bursts and recurrences of bursts in the spike trains of spontaneously active striate cortex neurons. *J. Neurophysiol.* 53, 926–939. doi: 10.1152/jn.1985.53.4.926
- Lei, H., Christensen, T. A., and Hildebrand, J. G. (2002). Local inhibition modulates odor-evoked synchronization of glomerulus-specific output neurons. *Nat. Neurosci.* 5, 557–565. doi: 10.1038/nn0602-859
- Lei, H., Christensen, T. A., and Hildebrand, J. G. (2004). Spatial and temporal organization of ensemble representations for different odor classes in the moth antennal lobe. *J. Neurosci.* 24, 11108–11119. doi: 10.1523/JNEUROSCI.3677-04.2004
- Lei, H., Reisenman, C. E., Wilson, C. H., Gabbur, P., and Hildebrand, J. G. (2011). Spiking patterns and their functional implications in the antennal lobe of the tobacco hornworm *manduca sexta*. *PLoS ONE* 6:e0023382. doi: 10.1371/journal.pone.0023382
- Lei, H., Yu, Y., Zhu, S., and Rangan, A. V. (2016). Intrinsic and network mechanisms constrain neural synchrony in the moth antennal lobe. *Front. Physiol.* 7:80. doi: 10.3389/fphys.2016.00080
- Lindner, B. (2006). Superposition of many independent spike trains is generally not a poisson process. *Phys. Rev. E* 73:022901. doi: 10.1103/PhysRevE.73.022901
- Martin, J. P., Lei, H., Riffell, J. A., and Hildebrand, J. G. (2013). Synchronous firing of antennal-lobe projection neurons encodes the behaviorally effective ratio of sex-pheromone components in male *manduca sexta*. *J. Comp. Physiol. A* 199, 963–979. doi: 10.1007/s00359-013-0849-z
- Martins, A., and Astudillo, R. (2016). “From softmax to sparsemax: a sparse model of attention and multi-label classification,” in *ICML (PMLR)*, 1614–1623.
- Matsumoto, S., and Hildebrand, J. (1981). Olfactory interneurons in the moth *manduca sexta*: response characteristics and morphology of central neurons in the antennal lobes. *Proc. R. Soc. London. Ser. B. Biol. Sci.* 213, 249–277. doi: 10.1098/rspb.1981.0066
- Mazor, O., and Laurent, G. (2005). Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons. *Neuron* 48, 661–673. doi: 10.1016/j.neuron.2005.09.032
- Nagel, K. I., and Wilson, R. I. (2016). Mechanisms underlying population response dynamics in inhibitory interneurons of the drosophila antennal lobe. *J. Neurosci.* 36, 4325–4338. doi: 10.1523/JNEUROSCI.3887-15.2016
- Ni, A. M., Huang, C., Doiron, B., and Cohen, M. R. (2022). A general decoding strategy explains the relationship between behavior and correlated variability. *Elife* 11:e67258. doi: 10.7554/eLife.67258
- Olsen, S. R., Bhandawat, V., and Wilson, R. I. (2010). Divisive normalization in olfactory population codes. *Neuron* 66, 287–299. doi: 10.1016/j.neuron.2010.04.009
- Panzeri, S., Moroni, M., Safaai, H., and Harvey, C. D. (2022). The structures and functions of correlations in neural population codes. *Nat. Rev. Neurosci.* 23, 551–567. doi: 10.1038/s41583-022-00606-4
- Pasupathy, A., and Connor, C. E. (2002). Population coding of shape in area v4. *Nat. Neurosci.* 5, 1332–1338. doi: 10.1038/972
- Putney, J., Conn, R., and Sponberg, S. (2019). Precise timing is ubiquitous, consistent, and coordinated across a comprehensive, spike-resolved flight motor program. *PNAS* 116, 26951–26960. doi: 10.1073/pnas.1907513116
- Putney, J., Niebur, T., Wood, L., Conn, R., and Sponberg, S. (2023). An information theoretic method to resolve millisecond-scale spike timing precision in a comprehensive motor program. *PLoS Comput. Biol.* 19:1011170. doi: 10.1371/journal.pcbi.1011170
- Reisenman, C. E., Dacks, A. M., and Hildebrand, J. G. (2011). Local interneuron diversity in the primary olfactory center of the moth *manduca sexta*. *J. Comp. Physiol. A* 197, 653–665. doi: 10.1007/s00359-011-0625-x
- Reiter, S., and Stopfer, M. (2013). “Spike timing and neural codes for odors,” in *Spike Timing: Mechanisms and Function* (Boca Raton, FL: CRC Press), 273–297.
- Riffell, J. A., Lei, H., Christensen, T. A., and Hildebrand, J. G. (2009a). Characterization and coding of behaviorally significant odor mixtures. *Curr. Biol.* 19, 335–340. doi: 10.1016/j.cub.2009.01.041
- Riffell, J. A., Lei, H., and Hildebrand, J. G. (2009b). Neural correlates of behavior in the moth *manduca sexta* in response to complex odors. *PNAS* 106, 19219–19226. doi: 10.1073/pnas.0910592106
- Riffell, J. A., Shlizerman, E., Sanders, E., Abrell, L., Medina, B., Hinterwirth, A. J., et al. (2014). Flower discrimination by pollinators in a dynamic chemical environment. *Science* 344, 1515–1518. doi: 10.1126/science.1251041

- Sachse, S., and Galizia, C. G. (2002). Role of inhibition for temporal and spatial odor representation in olfactory output neurons: a calcium imaging study. *J. Neurophysiol.* 87, 1106–1117. doi: 10.1152/jn.00325.2001
- Saupe, E. E., Myers, C. E., Townsend Peterson, A., Soberón, J., Singarayer, J., Valdes, P., et al. (2019). Spatio-temporal climate change contributes to latitudinal diversity gradients. *Nat. Ecol. Evol.* 3, 1419–1429. doi: 10.1038/s41559-019-0962-7
- Semmelhack, J. L., and Wang, J. W. (2009). Select drosophila glomeruli mediate innate olfactory attraction and aversion. *Nature* 459, 218–223. doi: 10.1038/nature07983
- Silbering, A. F., and Galizia, C. G. (2007). Processing of odor mixtures in the drosophila antennal lobe reveals both global inhibition and glomerulus-specific interactions. *J. Neurosci.* 27, 11966–11977. doi: 10.1523/JNEUROSCI.3099-07.2007
- Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., Poole, B., et al. (2020). Score-based generative modeling through stochastic differential equations. *arXiv [preprint]*. arXiv:2011.13456. doi: 10.48550/arXiv.2011.13456
- Stopfer, M., Bhagavan, S., Smith, B. H., and Laurent, G. (1997). Impaired odour discrimination on desynchronization of odour-encoding neural assemblies. *Nature* 390, 70–74. doi: 10.1038/36335
- Stopfer, M., Jayaraman, V., and Laurent, G. (2003). Intensity versus identity coding in an olfactory system. *Neuron* 39, 991–1004. doi: 10.1016/j.neuron.2003.08.011
- Strutz, A., Soelster, J., Baschwitz, A., Farhan, A., Grabe, V., Rybak, J., et al. (2014). Decoding odor quality and intensity in the drosophila brain. *Elife* 3:e04147. doi: 10.7554/eLife.04147.020
- Tanaka, N. K., Ito, K., and Stopfer, M. (2009). Odor-evoked neural oscillations in *Drosophila* are mediated by widely branching interneurons. *J. Neurosci.* 29, 8595–8603. doi: 10.1523/JNEUROSCI.1455-09.2009
- Tobin, W. F., Wilson, R. I., and Lee, W.-C. A. (2017). Wiring variations that enable and constrain neural computation in a sensory microcircuit. *Elife* 6:e24838. doi: 10.7554/eLife.24838.021
- Triplett, M. A., and Goodhill, G. J. (2019). Probabilistic encoding models for multivariate neural data. *Front. Neural Circuits* 13:1. doi: 10.3389/fncir.2019.00001
- Tsai, K.-T., Hu, C.-K., Li, K.-W., Hwang, W.-L., and Chou, Y.-H. (2018). Circuit variability interacts with excitatory-inhibitory diversity of interneurons to regulate network encoding capacity. *Sci. Rep.* 8:8027. doi: 10.1038/s41598-018-26286-8
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2023). Attention is all you need. *arXiv [preprint]*. arXiv:1706.03762. doi: 10.48550/arXiv.1706.03762
- Wachowiak, M., Dewan, A., Bozza, T., O'Connell, T. F., and Hong, E. J. (2025). Recalibrating olfactory neuroscience to the range of naturally occurring odor concentrations. *J. Neurosci.* 45:e1872242024. doi: 10.1523/JNEUROSCI.1872-24.2024
- Wilsch, A., Mercier, M. R., Obleser, J., Schroeder, C. E., and Haegens, S. (2020). Spatial attention and temporal expectation exert differential effects on visual and auditory discrimination. *J. Cogn. Neurosci.* 32, 1562–1576. doi: 10.1162/jocn_a_01567
- Wilson, D. A., and Sullivan, R. M. (2011). Cortical processing of odor objects. *Neuron* 72, 506–519. doi: 10.1016/j.neuron.2011.10.027
- Wilson, R. I., Turner, G. C., and Laurent, G. (2004). Transformation of olfactory representations in the *Drosophila* antennal lobe. *Science* 303, 366–370. doi: 10.1126/science.1090782
- Yang, H., Hasan, A., Ng, Y., and Tarokh, V. (2024). “Neural McKean-Vlasov processes: distributional dependence in diffusion processes,” in *AISTATS* (PMLR), 262–270.
- Ye, J., and Pandarinath, C. (2021). Representation learning for neural population activity with neural data transformers. *arXiv [preprint]*. arXiv:2108.01210. doi: 10.48550/arXiv.2108.01210