Using Deep Reinforcement Learning to understand Odor Plume Tracking in Walking and Flying Insects

Anonymous Author(s)
Affiliation
Address
email

Abstract

Odor plume tracking is critical for insect survival, yet how locomotion mode shapes navigation strategies remains unclear. While the "cast-and-surge" behavior of flying insects is well studied, the strategies of walking insects have received less attention. In this preliminary work, we use deep reinforcement learning to train biologically inspired artificial recurrent neural network agents to navigate to odor sources, directly comparing walking and flying modes. Walking agents, constrained by slower movement and limited turning, developed distinct strategies: fine-scale orientation relative to the plume centerline, pausing followed by localized search after plume loss, and subtle trajectory adjustments. Flying agents instead relied on broad sweeping turns and rapid plume reacquisition. Principal component analysis of recurrent activity revealed corresponding differences in neural representations: walking agents occupied compact, lower-dimensional manifolds, whereas flying agents exhibited continuous, higher-dimensional dynamics suited to flexible control. Our work illustrates how reinforcement learning can generate normative models of insect navigation and demonstrate the utility of AI as a tool for uncovering general principles linking biomechanics, behavior, and neural computation.

1 Introduction

Odor plume tracking is a fundamental behavior in many insects, enabling them to locate critical resources such as food and mates [Baker et al., 2018, Basu and Nagel, 2024]. In natural environments, odor plumes are fragmented and turbulent, requiring animals to integrate intermittent sensory cues with prior experience and motor constraints [Celani et al., 2014]. Different locomotor modes, such as walking and flying, impose distinct constraints on navigation. Walking insects move slowly with rapid turning capabilities[Katsov et al., 2017], while flying insects achieve higher speeds but face wind drift and reduced maneuverability. These biomechanical differences largely alter how sensory information translates into effective navigation strategies [Singh et al., 2023, Celani et al., 2014].

Prior work

Much work has been done to study odor navigation in insects experimentally in the field or controlled laboratory settings Basu and Nagel [2024], Baker et al. [2018]. Flying insects such as moths and fruit flies often rely on “cast-and-surge” behaviors to re-enter plumes after odor loss [Cardé and Willis, 2008, van Breugel and Dickinson, 2014, Baker et al., 2018, Pang et al., 2018]. Walking insects, by contrast, employ local reorientation, stochastic turning, and vision-odor integration [Bidaye et al., 2014, Tao et al., 2020, Demir et al., 2020, Gaudry et al., 2012, Rayshubskiy et al., 2025, Fisher, 2022, Willis et al., 2011].

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

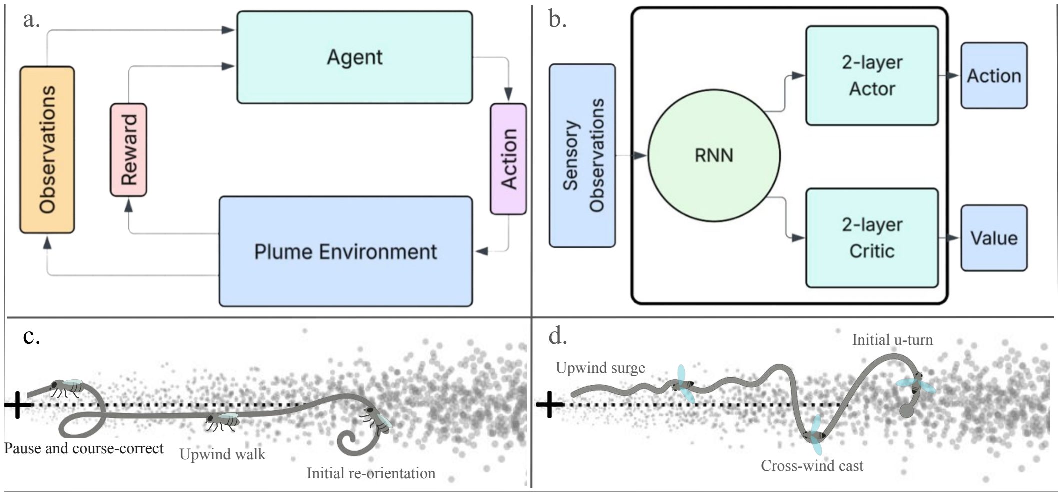

Figure 1: **Training agents that can track odor plume sources while flying or walking** (a) Reinforcement-learning loop. The agent receives egocentric wind-velocity and local odor-concentration observations from the environment and produces move and turn actions. These actions are matched to the capabilities of real flying and walking insects. The environment simulates a puff-based odor plume within a configurable wind field. (b) Network architecture used by the agent. Sensory observations feed an RNN followed by two feedforward heads. The actor head outputs actions and the critic head that predicts the value of the current belief state. (c) Walking agent: Example homing trajectory of a *walking* agent shows an initial reorientation, an upwind walk followed by a course correction on losing odor. (d) Flying agent: Example homing trajectory of a *flying* agent shows an initial course correction, similar to the walking agent. Further course corrections consist of larger cross-wind casting movements. The agent eventually tracks the edge of the plume when it finds the odor plume again.

Given the logistical challenges and costs associated with such studies, researchers have explored computational approaches that can complement such work [Verano et al., 2023, Singh et al., 2023, Stupski and van Breugel, 2024, van Breugel et al., 2008, Vouloutsi et al., 2013, Wang and Pang, 2023, Gunnarson and Dabiri, 2024]. Here we develop upon one line of such computational research that uses Deep Reinforcement Learning (DRL) with Recurrent Neural Networks (RNNs). Singh et al. [2023] showed that DRL-trained recurrent agents navigating simulated turbulent plumes spontaneously developed cast-and-surge behaviors reminiscent of flying insects, and their recurrent dynamics organized into ring-like neural manifolds encoding head direction and odor variables.

However, most computational studies have focused on a single locomotor mode, leaving open the question of how locomotion-specific constraints influence emergent strategies.

Contributions

Our preliminary work makes the following contributions:

- We introduce a comparative reinforcement learning framework for odor plume tracking that systematically contrasts walking and flying locomotor modes within a unified, physics-based simulator.
- We show that biomechanical constraints strongly shape emergent navigation strategies: flying agents develop broad surge–cast behaviors, while walking agents exhibit fine-scale adjustments, pausing, and local searches.
- Through analyses of the recurrent neural dynamics, we demonstrate that locomotor constraints also bias internal representations, with flying agents encoding task variables in continuous, distributed manifolds and walking agents relying on compact, state-dependent clusters.

2 Methods

Odor Plume Simulation Environment

We developed a physics-based plume simulation, based off OpenAI Gym [Brockman et al., 2016], following the approximate turbulent odor plume model by Farrell et al. [2002]. We generated odor landscapes through stochastic puff emissions sampled from a Poisson distribution at configurable rates (0.05–1.0). At each simulation step (25 FPS), all puffs were advected down-wind by the simulated wind field and expanded radially to simulate diffusion.

Agent Architecture: Flying vs. Walking

Agents are implemented as 64-unit recurrent actor-critic neural networks with *tanh* nonlinearities (Figure 1(b). Agents receive egocentric wind-velocity and local odor-concentration as observations. They generate continuous-valued actions controlling movement speed and turning rate, which result in movements in the simulated environment matched to the capabilities of real flying and walking insects.

Our study centers on comparing the effects of locomotor constraints implemented by constraining the maximum movement and turning speeds. Flying agents experience down-wind advection, but walking agents do not. Walking agents can move at a maximum of 0.05 m/s with π rad/s maximum turning capacity and experience no wind drift. Flying agents achieve 2.0 m/s with 6.25 π rad/s turning but must compensate for wind advection up to 0.5 m/s.

Deep Reinforcement Learning based Training

Agents were trained with Proximal Policy Gradient Optimization (PPO)[Schulman et al., 2017] using a randomized curriculum that varied initial positions, orientations and plume concentrations to ensure diverse training conditions. Large positive rewards were given for homing in on the source of the plume and negative rewards were given for leaving arena bounds. A small shaping reward for radial progress towards the odor source was added to encourage faster convergence. The training was carried out for 10^7 time steps with 3 to 5 random seeds per condition, using a learning rate of 3e-4 with decay (factor 0.00012) and regularization of the entropy (0.005). From these, we selected a single best-trained agent for further analysis. Agents were evaluated in single-source configurations, with episodes recording 25 Hz trajectories of position, orientation, odor concentration, and relative wind velocity. Evaluation on held-out plumes (240 episodes per model) measured success rate, trajectory efficiency, off-plume fraction, and plume exit duration.

3 Results

Behavioral Strategies Differ Across Locomotor Modes:

Both flying and walking agents successfully tracked turbulent odor plumes, but they employed distinct strategies (Figure 2). In low-concentration plumes, flying agents (Figure 2a,c) executed broad upwind surges interspersed with cross-wind casting movements, enabling rapid plume reacquisition after odor loss. Walking agents (Figure 2b,d), by contrast, made smaller, incremental adjustments: they oriented carefully upwind, paused when losing odor contact, and re-engaged the plume through localized searches. At higher plume concentrations, both agent types improved in reliability, though the characteristic differences in movement signatures persisted (Figure 2e–h).

Behavioral states and neural activity dimensionality

To better understand how locomotor constraints influence control, we quantified behavioral and sensory variables conditioned on navigational states: *Track* (actively following plume), *Recover* (reorienting after brief loss), and *Lost* (searching without odor contact). Flying agents (Figure 3a) exhibited broad distributions of head direction and displacement variables across states, likely due to continuous adjustments required by wind advection. Walking agents (Figure 3c) instead showed sharper and more conservative movement distributions,

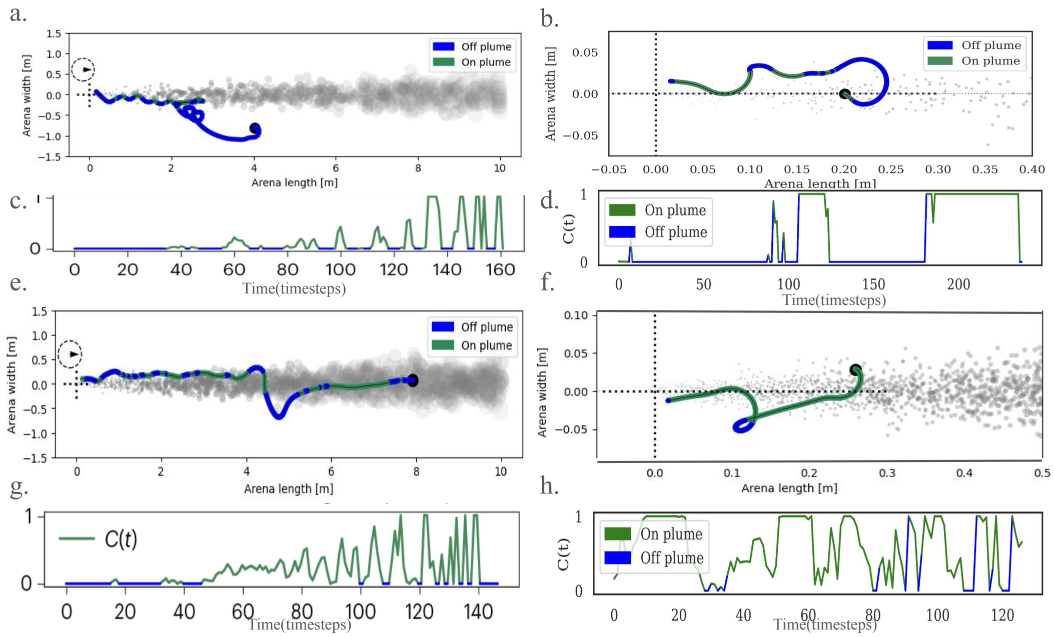

Figure 2: **Emergent plume behavior across different plume concentration levels:** (a-b) Successful odor tracking trajectories of (a) a flying agent and (b) a walking agent in a *low* odor-concentration plume. (c-d) Sensed odor concentration time-courses for (c) the flying agent and (b) the walking agent. (e-h) Successful odor tracking trajectories and odor-encounter time courses of (e,g) a flying agent and (f,h) a walking agent in a *regular* odor-concentration plume.

Dimensionality analysis further highlighted this contrast: flying networks required approximately five principal components (PCs) to explain 90% of neural variance (Figure 3b), while walking networks required only three (Figure 3d), suggesting a more compact coding scheme.

Manifold geometry reflects locomotor mode

Visual analysis of recurrent activity revealed qualitatively different representational geometries (Figure 4). For flying agents (Figure 4a–d), neural trajectories organized into smooth, ring-like manifolds. Head direction formed a continuous gradient around the ring (Figure 4a), while odor variables such as encounter history and concentration mapped onto broad continuous bands (Figure 4b–d). Walking agents instead exhibited branched, segmented manifolds (Figure 4e–h), with discrete clusters corresponding to different behavioral states.

4 Discussion

We presented a comparative study of walking and flying agents navigating turbulent odor plumes using physics-based simulation and reinforcement learning. Our results show that locomotor constraints profoundly shape both behavioral strategies and control representations: flying agents adopt broad, exploratory surge–cast cycles with rapid plume re-entry, while walking agents exhibit conservative, locally efficient adjustments. These differences extend to neural representations, with flying agents utilizing higher-dimensional, continuous manifolds similar to those found in other navigation tasks [Banino et al., 2018, Zhang et al., 2024], while walking agents employ compact, state-dependent clusters reminiscent of discrete behavioral states [Verano et al., 2023, Hennig et al., 2023].

Our study has several limitations that suggest directions for future work. We simplified sensory processing by using point odor measurements rather than bilateral antennae that enable instantaneous gradient detection [Gaudry et al., 2012, Siliciano et al., 2025], and omitted visual cues known to be crucial for plume tracking [Willis et al., 2011, van Breugel et al., 2022]. Our 2D environment neglects altitude control important for flying insects [Stupski and van Breugel, 2024] and uses static wind fields unlike the dynamic conditions insects face [Houle and van Breugel, 2023, Nag and van

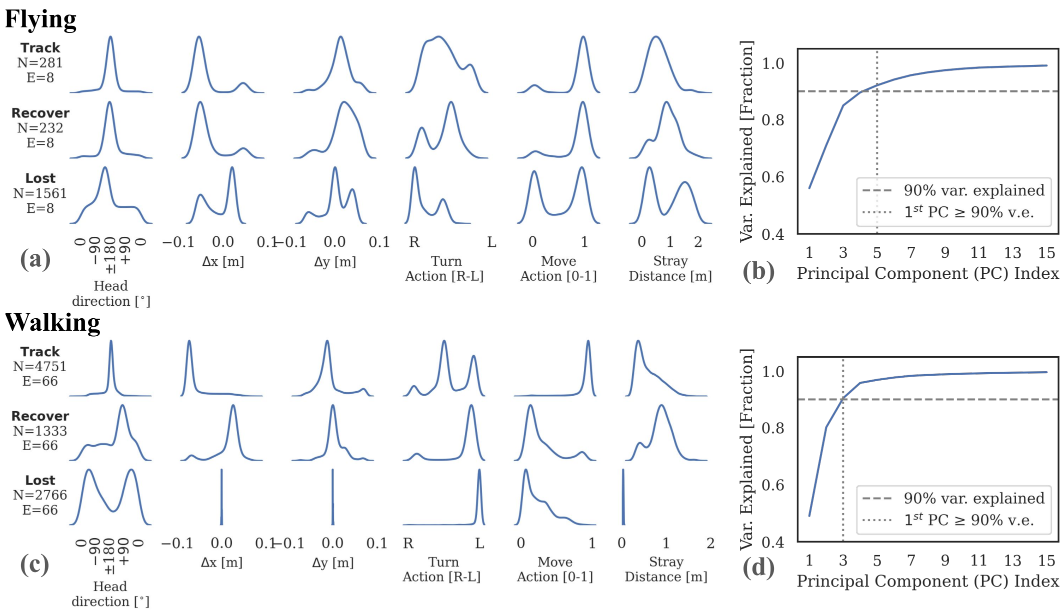

Figure 3: **Behavior and neural activity statistics across flying and walking agents:** (a & c) State-conditioned distributions of behavioral and sensory variables for (a) flying and (c) walking agents, conditioned on three navigational states: Track (actively following the plume), Recover (reorienting after brief loss), and Lost (searching without plume contact). Shown variables include head direction, translational displacements (delta x, delta y), turning action (right–left), movement action (normalized movement magnitude [0–1]), and distance strayed from the plume axis. (b & d) Fraction of variance explained by principal components (PCs) of recurrent neural activity for (b) flying and (d) walking agents. Flying-agent networks typically require more PCs ($\approx$5 PCs) than walking-agent networks ($\approx$3 PCs) to reach the 90% variance threshold, indicating richer, more distributed neural dynamics in the flying case.

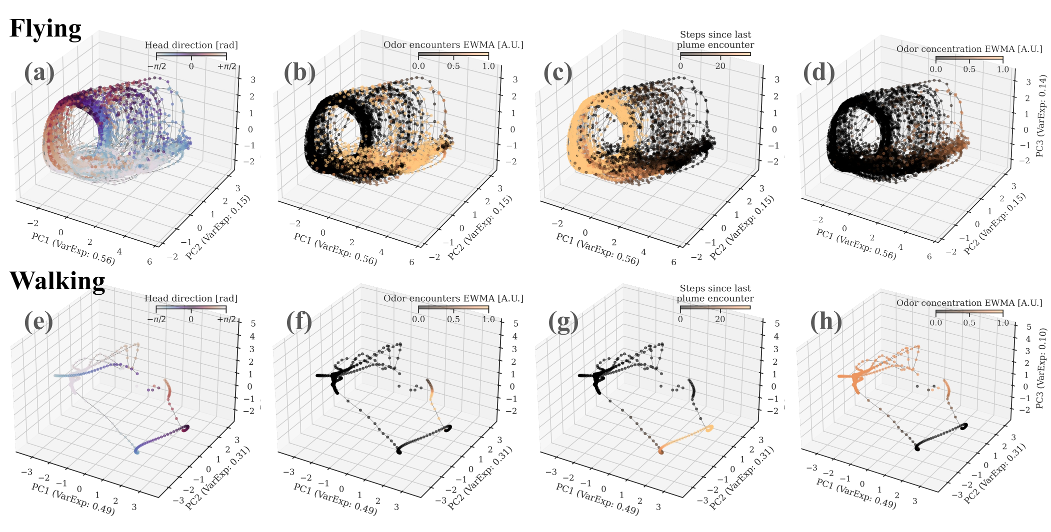

Figure 4: **Principal component analysis (PCA) of recurrent neural activity in flying and walking agents:** For each subfigure, we plot the neural activity across several plume tracking trajectories, colored by a task-relevant variable: (a, e) Head direction in radians for flying and walking agents respectively, (b, f) Exponentially weighted moving average (EWMA) of odor encounter events. (c, g) Duration (in simulation steps) since the last plume encounter (d, h) EWMA of instantaneous odor concentration.

Breugel, 2025]. The reward structure, while biologically motivated, lacks the complexity of real foraging decisions involving risk assessment and energy budgets [Reddy et al., 2022].

Future extensions could address these limitations while exploring richer behaviors. Incorporating dynamic wind shifts and multiple odor sources would test source discrimination capabilities [Celani and Panizon, 2024, Heinonen et al., 2025]. Hybrid agents capable of walk-to-flight transitions, as seen in many insects, could reveal optimal strategy switching [Rayshubskiy et al., 2025, Wang-Chen et al., 2024].

References

Keeley L Baker, Michael Dickinson, Teresa M Findley, David H Gire, Matthieu Louis, Marie P Suver, Justus V Verhagen, Katherine I Nagel, and Matthew C Smear. Algorithms for olfactory search across species. *Journal of Neuroscience*, 38(44):9383–9389, 2018.

Jayeeta Basu and Katherine Nagel. Neural circuits for goal-directed navigation across species. *Trends in Neurosciences*, 47(11):904–917, 2024.

Antonio Celani, Emmanuel Villermaux, and Massimo Vergassola. Odor landscapes in turbulent environments. *Physical Review X*, 4(4):041015, 2014.

Alexander Y Katsov, Limor Freifeld, Mark Horowitz, Seppe Kuehn, and Thomas R Clandinin. Dynamic structure of locomotor behavior in walking fruit flies. *eLife*, 6:e26410, jul 2017. ISSN 2050-084X. doi: 10.7554/eLife.26410. URL `https://doi.org/10.7554/eLife.26410`.

Satpreet H Singh, Floris van Breugel, Rajesh PN Rao, and Bingni W Brunton. Emergent behaviour and neural dynamics in artificial agents tracking odour plumes. *Nature machine intelligence*, 5(1):58–70, 2023.

Ring T Cardé and Mark A Willis. Navigational strategies used by insects to find distant, wind-borne sources of odor. *Journal of Chemical Ecology*, 34(7):854–866, 2008.

Floris van Breugel and Michael H Dickinson. Plume-tracking behavior of flying drosophila emerges from a set of distinct sensory-motor reflexes. *Current Biology*, 24(3):274–286, 2014.

Rich Pang, Floris van Breugel, Michael Dickinson, Jeffrey A Riffell, and Adrienne Fairhall. History dependence in insect flight decisions during odor tracking. *PLoS Computational Biology*, 14(2):e1005969, 2018.

Salil S Bidaye, Christian Machacek, Yang Wu, and Barry J Dickson. Neuronal control of drosophila walking direction. *Science*, 344(6179):97–101, 2014.

Liangyu Tao, Siddhi Ozarkar, and Vikas Bhandawat. Mechanisms underlying attraction to odors in walking drosophila. *PLoS computational biology*, 16(3):e1007718, 2020.

Mahmut Demir, Nirag Kadakia, Hope D Anderson, Damon A Clark, and Thierry Emonet. Walking drosophila navigate complex plumes using stochastic decisions biased by the timing of odor encounters. *Elife*, 9:e57524, 2020.

Quentin Gaudry, Katherine I Nagel, and Rachel I Wilson. Smelling on the fly: sensory cues and strategies for olfactory navigation in drosophila. *Current opinion in neurobiology*, 22(2):216–222, 2012.

Aleksandr Rayshubskiy, Stephen L Holtz, Alexander S Bates, Quinn X Vanderbeck, Laia Serratosa Capdevila, Victoria Rockwell, and Rachel Wilson. Neural circuit mechanisms for steering control in walking drosophila. *ELife*, 13:RP102230, 2025.

Yvette E Fisher. Flexible navigational computations in the drosophila central complex. *Current opinion in neurobiology*, 73:102514, 2022.

Mark A Willis, Jennifer L Avondet, and Elizabeth Zheng. The role of vision in odor-plume tracking by walking and flying insects. *Journal of Experimental Biology*, 214(24):4121–4132, 2011.

Kyrell Vann B Verano, Emanuele Panizon, and Antonio Celani. Olfactory search with finite-state controllers. *Proceedings of the National Academy of Sciences*, 120(34):e2304230120, 2023.

S David Stupski and Floris van Breugel. Wind gates olfaction-driven search states in free flight. *Current Biology*, 34(19):4397–4411, 2024.

Floris van Breugel, William Regan, and Hod Lipson. From insects to machines. *IEEE Robotics & Automation Magazine*, 15(4):68–74, 2008.

Vasiliki Vouloutsi, Lucas L LopezSerrano, Zenon Mathews, Alex Escuredo Chimeno, Andrey Ziyatdinov, Alexandre Perera i Lluna, Sergi Bermúdez i Badia, and Paul FM J Verschure. The synthetic moth: a neuromorphic approach toward artificial olfaction in robots. *Neuromorphic Olfaction*, pages 117–152, 2013.

Lingxiao Wang and Shuo Pang. Autonomous underwater vehicle based chemical plume tracing via deep reinforcement learning methods. *Journal of Marine Science and Engineering*, 11(2):366, 2023.

Peter Gunnarson and John O Dabiri. Fish-inspired tracking of underwater turbulent plumes. *Bioinspiration & Biomimetics*, 19(5):056024, 2024.

Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI gym. *arXiv preprint arXiv:1606.01540*, 2016.

Jay A Farrell, John Murlis, Xuezhu Long, Wei Li, and Ring T Cardé. Filament-based atmospheric dispersion model to achieve short time-scale structure of odor plumes. *Environmental Fluid Mechanics*, 2(1-2):143–169, 2002.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *ArXiv*, 2017. URL `https://arxiv.org/abs/1707.06347`.

Andrea Banino, Caswell Barry, Benigno Uria, Charles Blundell, Timothy Lillicrap, Piotr Mirowski, Alexander Pritzel, Martin J Chadwick, Thomas Degris, Joseph Modayil, et al. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, 2018.

Ruiyi Zhang, Xaq Pitkow, and Dora E Angelaki. Inductive biases of neural network modularity in spatial navigation. *Science Advances*, 10(29):eadk1256, 2024.

Jay A Hennig, Sandra A Romero Pinto, Takahiro Yamaguchi, Scott W Linderman, Naoshige Uchida, and Samuel J Gershman. Emergence of belief-like representations through reinforcement learning. *PLOS Computational Biology*, 19(9):e1011067, 2023.

Andrew F. Siliciano, Sun Minni, Chad Morton, Charles K. Dowell, Noelle B. Eghbali, Juliana Y. Rhee, L. F. Abbott, and Vanessa Ruta. A vector-based strategy for olfactory navigation in drosophila. *bioRxiv*, February 2025.

Floris van Breugel, Renan Jewell, and Jaleesa Houle. Active anemosensing hypothesis: how flying insects could estimate ambient wind direction through sensory integration and active movement. *Journal of The Royal Society Interface*, 19(193):20220258, 2022.

Jaleesa Houle and Floris van Breugel. Near-surface wind variability over spatiotemporal scales relevant to plume tracking insects. *Physics of Fluids*, 35(5), 2023.

Arunava Nag and Floris van Breugel. Cosmos: A data-driven probabilistic time series simulator for chemical plumes across spatial scales. *arXiv preprint arXiv:2505.22436*, 2025.

Gautam Reddy, Venkatesh N Murthy, and Massimo Vergassola. Olfactory sensing and navigation in turbulent environments. *Annual Review of Condensed Matter Physics*, 13(1):191–213, 2022.

Antonio Celani and Emanuele Panizon. Olfactory search. In *Target Search Problems*, pages 711–732. Springer, 2024.

Robin A Heinonen, Luca Biferale, Antonio Celani, and Massimo Vergassola. Optimal trajectories for bayesian olfactory search in turbulent flows: The low information limit and beyond. *Physical Review Fluids*, 10(4): 044601, 2025.

Sibo Wang-Chen, Victor Alfred Stimpfling, Thomas Ka Chung Lam, Pembe Gizem Özdil, Louise Genoud, Femke Hurtak, and Pavan Ramdya. Neuromechfly v2: simulating embodied sensorimotor control in adult drosophila. *Nature Methods*, 21(12):2353–2362, 2024.