PNAS RESEARCH ARTICLE | NEUROSCIENCE PHYSICS OPEN ACCESS

Peripheral preprocessing in *Drosophila* facilitates odor classification

Palka Puri[a], Shiuan-Tze Wu[b], Chih-Ying Su[b], and Johnatan Aljadeff[b,1]

Edited by L. Abbott, Columbia University, New York, NY; received October 1, 2023; accepted April 16, 2024

The mammalian brain implements sophisticated sensory processing algorithms along multilayered ("deep") neural networks. Strategies that insects use to meet similar computational demands, while relying on smaller nervous systems with shallow architectures, remain elusive. Using *Drosophila* as a model, we uncover the algorithmic role of odor preprocessing by a shallow network of compartmentalized olfactory receptor neurons. Each compartment operates as a ratiometric unit for specific odor-mixtures. This computation arises from a simple mechanism: electrical coupling between two differently sized neurons. We demonstrate that downstream synaptic connectivity is shaped to optimally leverage amplification of a hedonic value signal in the periphery. Furthermore, peripheral preprocessing is shown to markedly improve novel odor classification in a higher brain center. Together, our work highlights a far-reaching functional role of the sensory periphery for downstream processing. By elucidating the implementation of powerful computations by a shallow network, we provide insights into general principles of efficient sensory processing algorithms.

olfaction | *Drosophila* | shallow neural network | sensory periphery | connectome

Animals navigating in natural environments process rapidly changing stimuli (1, 2), that are high-dimensional (i.e., consist of multiple independent components). Different organisms have adopted a broad range of strategies for detecting relevant stimuli within such sensory environments to generate appropriate behavioral outputs. Major efforts in neuroscience are dedicated to understanding how the mammalian brain, utilizing elaborate and highly plastic neural architectures, solves these problems (3, 4). The compact nervous systems of insects offer an opportunity to discover alternative strategies for efficient sensory processing. In particular, the conserved compartmentalization of primary sensory neurons (5–8) suggests that hard-wired peripheral preprocessing of sensory information may enable organisms to effectively utilize shallow neural architectures for implementing powerful sensory processing algorithms.

In *Drosophila melanogaster*, the repertoire of ∼50 olfactory receptor neurons (ORNs) in the periphery is genetically determined, and their compartmentalization into ∼25 sensory hairs ("sensilla") is stereotyped (9–11). ORNs housed in the same sensillum exert nonsynaptic ("ephaptic") lateral inhibition (5). Our previous work has shown that these electrical interactions can affect odor-guided behaviors (5, 12, 13). In natural environments, fruitflies encounter high-dimensional and transient odor stimuli (2) that are first processed by this conserved array of ORNs. The olfactory system of *Drosophila* provides a unique opportunity to understand the significance of neuronal compartmentalization for processing natural stimuli due to extensive information on its neuroanatomy (12, 14–18), and rich theoretical work on how fly brains implement olfactory processing algorithms (19–21).

Olfactory signals from the periphery are transmitted to the antennal lobe (AL), from which inputs are routed along two main pathways (Fig. 1*A*). Previous work has shown that connections from the AL to lateral horn (LH) mediate innate odor-guided behaviors (16, 22). In natural environments, the concentration of many different odorant molecules can change independently, yet so far, behavioral experiments have mainly probed responses in a low dimensional stimulus space. By "low-dimensional," we mean that the number of stimulus components that are varied independently is small, irrespective of the number of glomeruli activated by the odors. This definition includes studies of behavioral and neural responses to a single odorant [e.g., ethyl acetate, 2-hexanol (22), CO_2 (23, 24)]; or studies where the overall dilution level of an odor-mixture was varied, with fixed relative odorant concentrations [e.g., apple cider vinegar (25)]. It is unclear how behavioral responses to innately meaningful, high-dimensional odors are generated and how the specific structure of AL-to-LH connections facilitates these computations.

Significance

Insects rely on compact nervous systems to process sensory information. To find principles of information processing in "shallow" neural architectures, we studied odor-mixture "preprocessing" in *Drosophila*, which is implemented by genetically predetermined electrical interactions between olfactory receptor neurons. We show that preprocessing facilitates classification of i) innately meaningful odors, by amplifying their hedonic valence signal; and ii) novel odors, by modifying the geometrical structure of neuronal responses to odor-mixture stimuli. Our work suggests that connectivity from the periphery to downstream brain-centers mediating innate and learned odor-guided behaviors is optimally aligned to take advantage of peripheral preprocessing. By elucidating how a peripheral network in the insect brain implements sophisticated computations, we give insight into general principles of efficient sensory processing algorithms.

Author contributions: P.P. and J.A. developed the model with input from S.-T.W. and C.-Y.S.; P.P. and J.A. solved, analyzed and simulated model; P.P. designed and performed data analysis with inputs from S.-T.W., C.-Y.S., and J.A.; C.-Y.S. and J.A. acquired funding and supervised project; and P.P. and J.A. wrote the paper with inputs from S.-T.W. and C.-Y.S.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: aljadeff@ucsd.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2316799121/-/DCSupplemental.

Published May 16, 2024.

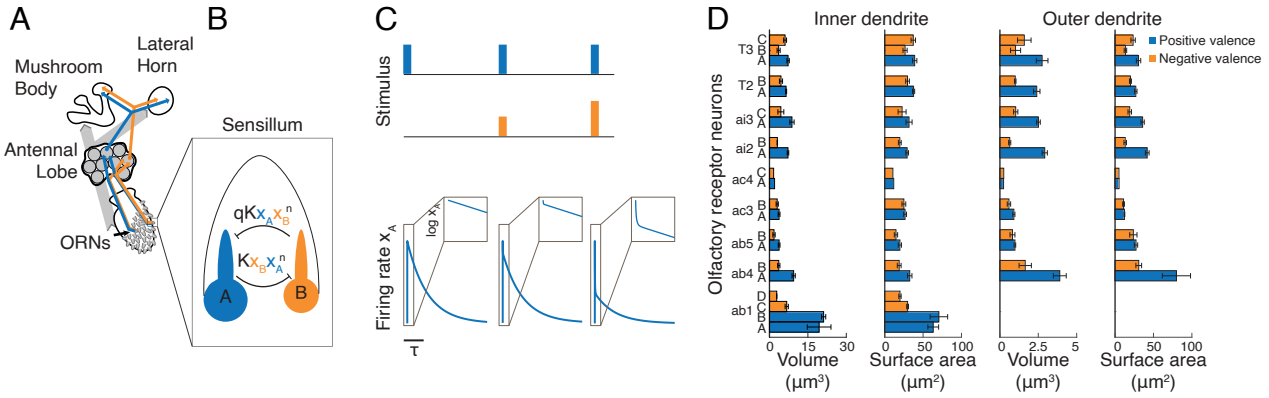

Fig. 1. Nonlinear model of peripheral ephaptic interactions. (*A*) Illustration of olfactory information flow in fruitflies. (*B*) Peripheral signal preprocessing is mediated by ephaptic interactions between cohoused ORNs, wherein the neuronal firing rates (x_A, x_B) are nonlinearly coupled. Model parameters *K, q, n* denote interaction strength, asymmetry, and nonlinearity, respectively. (*C*) Analytical solutions of the response of neuron *A* (*Bottom*) following offset of three different stimuli (*Top*). Here, the strength of the *A* odorant (blue) is constant, while the strength of the *B* odorant (orange) increases. Activating neuron *B* leads to suppression of neuron *A*'s response. *Insets*: Firing rate response on log scale illustrates a two-phase decay of the response to 0. (*D*) Valence (color) of cohoused ORNs matches the size asymmetry of their dendrites (adapted from ref. 14). Note that outer dendrite measurements for the ab1 sensillum were not performed in ref. 14.

In contrast, AL to mushroom body (MB) connections are largely random (15, 26), and MB responses are sparse (27). Architectures with random projections and sparsification are thought to facilitate robust learning of arbitrary stimulus–response associations, by increasing the linear separability of high-dimensional stimuli (19–21). The hypothesis of enhanced linear separability in the MB relies crucially on the assumption of *clustered* neuronal responses in the AL (19), i.e., odors that elicit the same behavioral response correspond to similar AL population activity patterns. This assumption is inconsistent with empirical data showing, for example, that AL representations of different odors at the same concentration are often closer than that of the same odor at different concentrations (28, 29).

To address these fundamental questions of sensory processing in compact nervous systems, we focused on the role of lateral inhibition between compartmentalized ORNs (5, 12) that transforms olfactory signals *before* they arrive at the AL (30). We hypothesized that this peripheral preprocessing step enables efficient downstream integration of olfactory inputs.

Lateral Ephaptic Inhibition between ORNs Prepares Olfactory Signals for Downstream Processing

We constructed a simplified model of electrical ("ephaptic") inhibition between pairs of ORNs housed in the same sensillum ("coupled ORNs"). The pairing of coupled ORNs, i.e., their assignment to sensilla, is genetically determined (9). Moreover, their relative sizes exhibit stereotypical asymmetry (12, 14). The conservation of ORN pairing and relative sizes strongly suggests a functional role of the sensillum, since the characteristics of ephaptic inhibition between paired ORNs are influenced by their relative physical dimensions (12, 31). For tractability, here we focused on the typical case of a sensillum with two ORNs (Fig. 1*B*). Each ORN's label corresponds to its relative size (*A*, large; *B*, small). The time evolution of ORN *A*'s firing rate (denoted x_A) is modeled as

$$\tau \frac{\mathrm{d}x_A}{\mathrm{d}t} = -x_A - w_{AB} x_A \phi(x_B) + s_A(t). \qquad [1]$$

The three terms on the right respectively model a decay to baseline with timescale τ, nonlinear ephaptic inhibition by neuron *B*, and the time-dependent concentration of neuron *A*'s preferred odorant (i.e., the stimulus). The inhibition term contains two essential nonlinearities: i) $\phi(x_B) = x_B^n$ with $n > 1$ imposes a soft threshold for neuron *B*'s activation to influence its neighbor (32), since this term is suppressed for activation $x_B < 1$ and amplified for $x_B > 1$; and ii) the factor of x_A implies that the degree of inhibition experienced by neuron A depends on its own level of activation (5). The evolution of x_B is given by replacing *A* with *B* in Eq. **1**. We defined the asymmetry of interaction strengths as $q = w_{BA}/w_{AB}$. Since neuron B is smaller, it exerts weaker ephaptic inhibition (12). Therefore, $q \leq 1$, with smaller values of q implying more asymmetry in ephaptic interactions. We expect q to be a key sensillum-type specific parameter that tunes the asymmetry of ephaptic inhibition. Thus, multiple "copies" of Eq. **1**, with different values of q, represent an array of ORNs housed in different sensillum types that process high-dimensional olfactory stimuli.

Insects move through turbulent environments and encounter odorants sparsely in time (2, 33). Therefore, we modeled stimuli as an instantaneous rise in concentration of the two odorants ($s_{A,B}(t) = S_{A,B}\delta(t)$), which sets the "initial-condition" of the firing rates, $x_{A,B}(t = 0) = S_{A,B}$ [*SI Appendix*, §1.A, (34)]. We derived exact analytical solutions for the two-phase transient decay of $x_{A,B}(t)$ following stimulus "offset" (*SI Appendix*, §1.A). The initial decay of x_A depends strongly on the concentration of odorant *B* due to ephaptic interactions (Fig. 1*C*). After this initial phase, the firing rates decay exponentially, with a negligible effect of the coupling. Thus, our model's complex transients contain information about the odor stimulus, consistent with experimental data showing that transient neural responses in the insect AL contain more information about the odor identity than steady-state responses (35).

Previous studies have demonstrated that ORN activity carries information about the behavioral valence of incoming stimuli. In particular, large ORNs (denoted *A*; blue bars in Fig. 1*D*) are positive—their activation tends to trigger odor-guided behaviors such as attraction, egg-laying, and courtship. Conversely, small ORNs (denoted *B*; Fig. 1*D*, orange) are typically negative—their activation leads to aversion, inhibition of egg-laying, and suppression of courtship behaviors (13). Beyond the association of *A* and *B* ORNs with opposite behavioral valence, our previous work further demonstrated that paired ORNs antagonistically

modulate the same behavioral output (ref. 13 and references therein). Based on these findings, we hypothesized that the computational role of ephaptic inhibition between behaviorally antagonistic ORNs is to enhance the net valence signal extracted from countervailing stimuli. The amplification of this valence signal could trigger an appropriate behavior by downstream regions.

Our model predicts that odor stimuli satisfying the linear scaling relationship $S_A = q^{1/n} S_B$ induce equal and opposite inhibition between the ORNs. Thus, we assumed that these stimuli are *neutral* (*SI Appendix*, §1.B). We defined valence amplification as the normalized net valence conveyed by the firing rates of the coupled ORNs,

$$\alpha_q(t) = \frac{x_A(t) - q^{1/n} x_B(t)}{S_A - q^{1/n} S_B}. \quad [2]$$

The solution to Eq. **1** shows that ephaptic interactions transiently amplify valence, i.e., $\alpha_q(t)$ rises above 1 before decreasing to 0 at long times (Fig. 2 *A* and *B*). We note that the factor of $q^{1/n}$ in Eq. **2** is necessary; any alternate (e.g., unweighted) definition of valence results in misidentified valence for a significant region of stimulus space (*SI Appendix*, Fig. S1 *A–C* and §1.B). While our definition of neutral odorant ratios relies on equal and opposite ephaptic inhibition in the model, odorant ratios that are neutral from the animal's point of view may depend on additional factors not included in our model. For example, nonuniform natural stimulus statistics and heterogeneous ligand binding kinetics (9) together mean that animals face some stimulus classification problems more frequently than others. Additionally, animals may ascribe different costs to false negative (positive odor mistakenly identified as negative) and false positive errors (*SI Appendix*, Fig. S1*D*).

The analysis of valence amplification in the model implies that to accurately detect stimulus valence based on the ORNs' firing rates, downstream readouts of ORN activity must "reweigh" their inputs to calibrate the valence signal to neutral. More precisely, we predict that readout-weights are aligned to the angle $\tan^{-1}(q^{-1/n})$, which depends on the interaction parameters for each sensillum. Ephaptic inhibition between coupled ORNs induces curvature of firing-rate trajectories as they decay to zero. Depending on the relative strengths of the odorants, the trajectories curve toward axes representing the activation of either the *A* or *B* neuron. In contrast, firing-rate trajectories decay to zero along a straight line in the absence of ephaptic coupling (Fig. 2*C*, compare solid and dashed lines). Ephaptic interactions therefore facilitate the discrimination of similar odor mixtures, based on the transient ORN responses. Both discriminability and valence amplification are maximal for odor stimuli with weak (i.e., close to neutral) valence (Fig. 2 *B* and *D*).

As animals traverse through odor plumes, odorant concentrations follow complex temporal dynamics (2), which may not be captured by the pulse shape we have assumed so far (Figs. 1*C* and 2 *A–D*). To show that our conclusions for stimulus offset apply in more general cases, we extended our analysis to the scenario of noninstantaneous stimulus onset (*SI Appendix*, §1.D). Indeed, in response to linear ramp stimuli of duration T, where $s_{A,B}(t) = S_{A,B} t/T$, transient ORN responses also exhibit amplified odor valence (Fig. 2 *E* and *F*) and enhanced discriminability (Fig. 2 *G* and *H*).

Notably, our model suggests that the interaction asymmetry, specific to each sensillum type, determines the "optimal" odorant ratio, i.e., with maximal amplification and discriminability. As q decreases, optimal stimuli are biased toward the negative valence odorant (Fig. 2*I* and *SI Appendix*, §1.C). Moreover, optimal stimuli computed for stimulus offset and onset are in good agreement, for all values of the asymmetry q (Fig. 2*I*, compare solid and dashed lines). This underscores the function of each sensillum type as a ratiometric processing unit and demonstrates how coupled ORNs could potentially extract behaviorally relevant olfactory information from odor mixtures with complex temporal dynamics of individual odorant concentrations.

We validated the model by fitting it to electrophysiological recordings of *A* and *B* ORNs in the ab2 sensillum as it was probed with binary odor mixtures which span the $S_A - S_B$ space (Fig. 2*J*). We fit the parameters q, n, K (Fig. 2*K* and *SI Appendix*, Fig. S2) by assuming that steady-state is reached within the 0.5 s stimulation, and leaving out 11% of the data for validation (*SI Appendix*, §1.E). The resulting coupling strength parameter K was strictly greater than 0 (*SI Appendix*, Fig. S2*D* and §1.E), which demonstrates that the ephaptic coupling term in the model is necessary for an accurate fit to data when paired ORNs are stimulated with a mixture of odorants. Additionally, the fact that the fitted asymmetry parameter $q = 0.019 \pm 0.007 \ll 1$ (Fig. 2*K* and *SI Appendix*, Fig. S2) indicates a strong asymmetry in interactions between ab2 ORNs, which is consistent with the large morphometric disparity measured for these neurons (14).

AL to LH Connectivity is Structured to Read Out Amplified Odor-Mixture Valence

We hypothesized that the LH, the primary brain region mediating innate olfactory behaviors (22, 36), takes advantage of ephaptic interactions at the periphery to robustly compute the valence of high-dimensional odor stimuli. To do so, LH neurons (LHNs) must asymmetrically weigh inputs from coupled ORNs to compute the amplified stimulus valence (numerator of Eq. **2**). Given that $q < 1$, LH readout-weights for negative-valence ORNs are expected to be smaller. To test this prediction, we analyzed AL-to-LH projections in the Hemibrain connectome (17). For simplicity, our analysis did not consider lateral interactions within the AL, or feedback connections (i.e., LH-to-AL, see *SI Appendix*, §2). Therefore, the details of the biological mechanisms of readout in the AL do not affect the results of our analyses of connectivity patterns, under the standard assumption made in the experimental and theoretical literature that high-order neurons monotonically integrate their input drive (15, 16, 19, 20).

We focused on uniglomerular projection neurons (uPNs), each of which constitutes a relay from a single ORN-type, thereby allowing us to directly examine ORN-specific signatures of projection asymmetry. Previous studies have identified one to three distinct *types* of uPNs per glomerulus, based on differences of their cell body locations and connectivity patterns (17, 37). Our analysis showed that different types of uPNs innervating the same glomerulus had LH connectivity that was nearly as dissimilar as that of uPNs from different glomeruli (*SI Appendix*, Fig. S3). Therefore, we analyzed distinct uPN-types separately. LHN-types are also defined based on their location and other anatomical features (17). As neurons of the same type have highly overlapping projection patterns (*SI Appendix*, Fig. S3), we considered a connection between each uPN-type and LHN-type to be a distinct information channel. This resulted in a connectivity matrix with 71 presynaptic uPN-types and 673 postsynaptic LHN-types (*SI Appendix*, §2.A). The weight of each

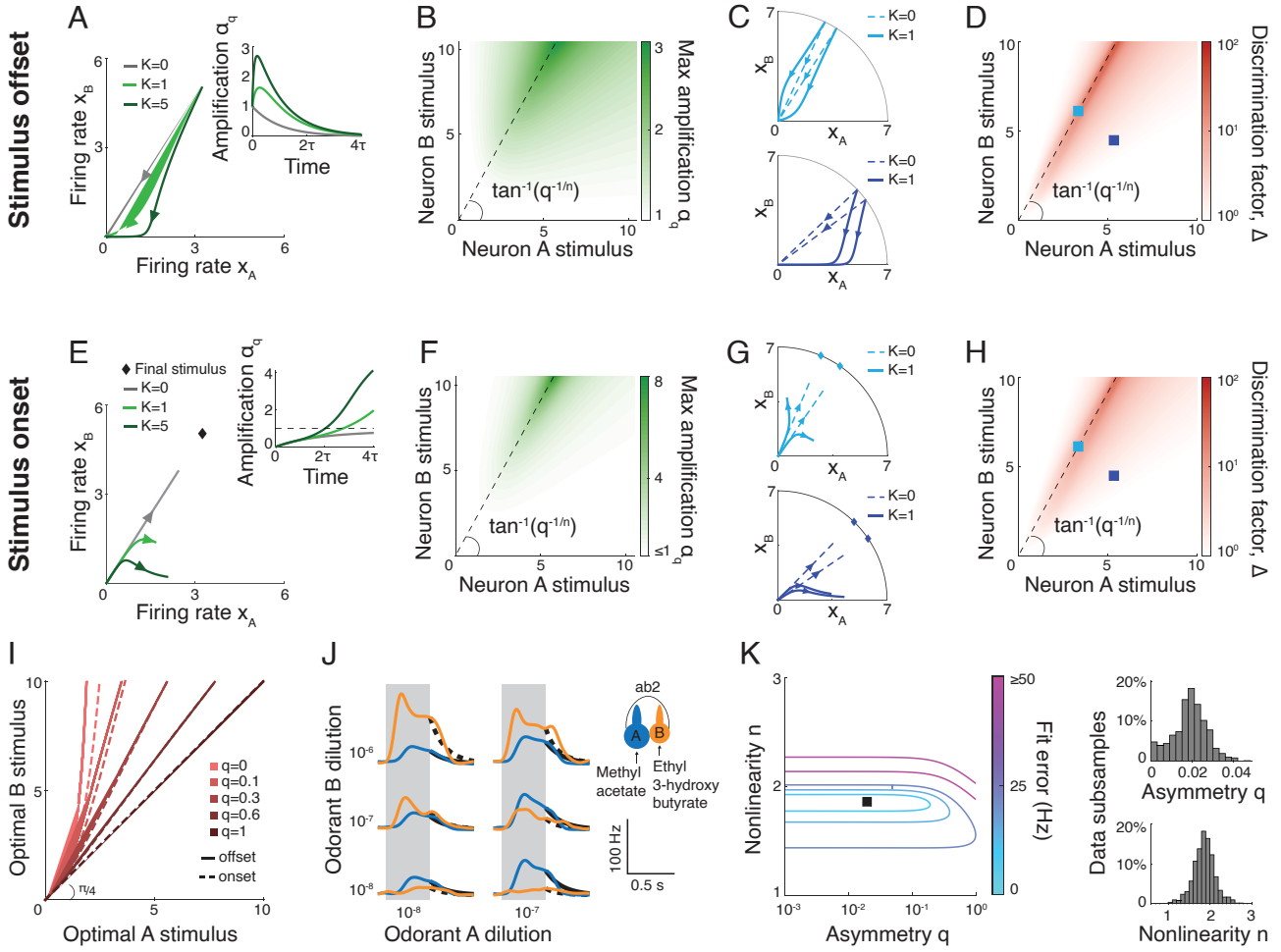

Fig. 2. Ephaptic interactions transiently amplify odor valence. (*A*) Firing-rate trajectories following stimulus offset for different values of interaction strength K. *Inset*: Corresponding valence amplification $\alpha_q(t)$. In the presence of ephaptic interactions, valence is transiently amplified $(\alpha_q(t) > 1)$. (*B*) Magnitude of valence amplification following stimulus offset [peak of $\alpha_q(t)$] as a function of the stimuli S_A, S_B. Amplification is stimulus-specific: maximal amplification is achieved for stimuli close to neutral (dashed line). (*C*) Example firing-rate trajectories in the presence of ephaptic interactions ($K = 1$, solid). *Top*: Stimuli close to neutral, i.e., "weak" valence signal (light blue). *Bottom*: strongly positive stimuli (dark blue). Firing rates follow straight lines in the absence of interactions ($K = 0$, dashed). Light blue trajectories are transiently separated, while dark blue trajectories collapse onto each other. (*D*) Discrimination factor Δ (*SI Appendix*, §1.C) is maximal for neutral combinations of S_A, S_B (dashed line). Large Δ implies improvement in stimulus discriminability based on transient firing rates. Squares indicate stimuli in (*C*). (*E–H*) Same as (*A–D*), for stimulus onset (linear ramp up, final value at $T = 4\tau$ indicated by diamond in panels *E* and *G*). (*I*) Curves of optimal stimuli (with largest Δ) for stimulus offset (solid) and onset (dashed). Optimal stimuli depend strongly on the interaction asymmetry q. (*J*) Electrophysiological recordings of ORNs in the ab2 sensillum, probed with different odorant mixtures (data from ref. 12) that were fit to the model (*SI Appendix*, §1.E). Gray shade: odor presentation epoch. Black lines (solid: ab2A, dashed: ab2B) show model predictions following stimulus offset. (*K*) Fitting error for ab2 as a function of interaction asymmetry and nonlinearity parameters (*Left*). Best-fit values, $q = 0.019 \pm 0.007$, $n = 1.9 \pm 0.26$ (mean ± SD, black square). SD obtained by subsampling trials (*Right*). For (*A–H*), $q = 0.3$, $n = 2$, $K = 1$ (in units of τ^n), unless noted otherwise.

connection was obtained by summing over the synapse counts between all individual neurons of the corresponding uPN- and LHN-types ((17), *SI Appendix*, §2.A). Henceforth, "uPN" and "LHN" refer to neuronal *types*, unless noted otherwise.

To identify salient connectivity features, it is important to compare the connectivity matrix to shuffled matrices that preserve global properties of the connections (15, 20, 26). We thus generated shuffled connectivity matrices that preserve the total number of incoming and outgoing connections per neuron in the original data (*SI Appendix*, Fig. S4). In addition to the connection frequencies, we preserved either incoming connection *weights* for each LHN (type-1 shuffles), or outgoing weights for each uPN (type-2 shuffles, see *SI Appendix*, §2.B and Fig. S5). The two types of shuffles allow us to control for overall biases in connection weights of LHNs (type-1) and uPNs (type-2).

Following the analysis in ref. 15, we used principal component analysis (PCA) on the empirical and shuffled connectivity matrices to extract nonrandom patterns of connectivity in the data that cannot be explained by the connectivity statistics pooled across all synapses in the network, or by the connectivity statistics specific to each neuron [e.g., sparse connections, (20)]. Our analysis revealed that the first few PCs of the data explain a significantly larger amount of variance than expected by chance (Fig. 3*A* and *SI Appendix*, Fig. S6). This result is in contrast to refs. 15 and 26, where PCA of AL-to-MB connections revealed no such PCs. The results of PCA in these studies suggest that projections to the MB are random. Thus, the significant PCs of projections to the LH represent structured connections from uPNs to LHNs. The leading PC (PC1) describes the most prominent pattern of AL-to-LH readouts. The PC1 coefficient of each uPN quantifies its weight in the aforementioned readout pattern.

Based on the available valence information for ORNs (13), we identified 30 uPNs that were postsynaptic to coupled,

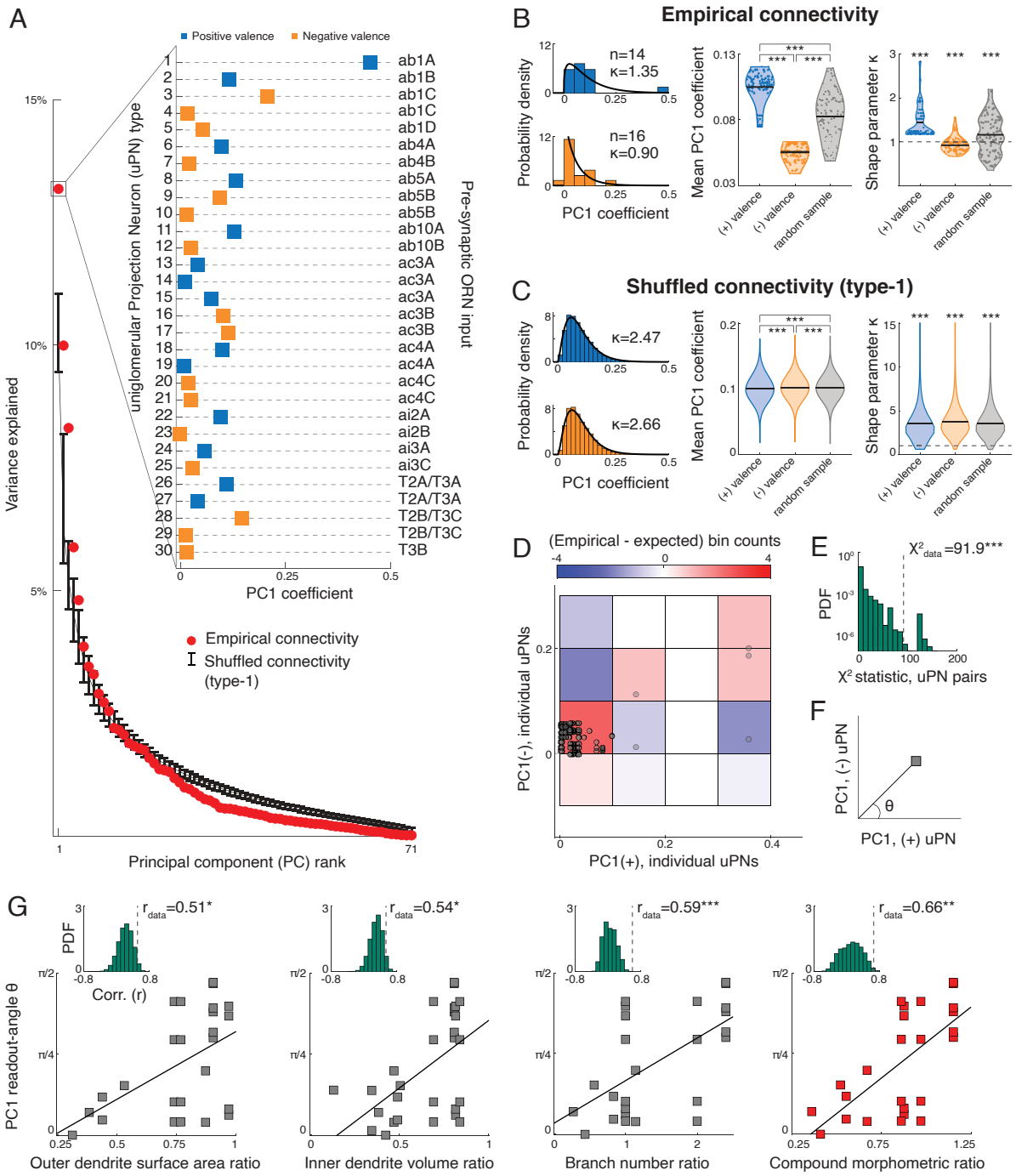

Fig. 3. AL-to-LH projections are structured to compute amplified stimulus valence. (*A*) Variance explained by principal components (PCs) of the empirical AL-to-LH connectivity matrix (red) and shuffles (black, error bars indicate 95% CI). *Inset*: PC1 coefficients of uPN-types, colored by the valence of their presynaptic ORN inputs (13). uPNs are ordered by the sensilla organization of the presynaptic ORNs (*SI Appendix*, Table S1). uPN-types 26–29 each innervate a single glomerulus which is postsynaptic to two ORN-types. (*B*) *Left*: Distributions of PC1 coefficients for positive- and negative-valence uPN-types, and the best-fit gamma distributions (black) for the empirical connectivity (*SI Appendix*, §2.C). *Right*: The average PC1 coefficient and gamma distribution shape parameter (κ) obtained from 100 subsamples of positive-valence, negative-valence, and randomly selected uPN-types (blue, orange, and gray, respectively; see *SI Appendix*, §2.C). Black lines show average over subsamples. The distribution of PC1 coefficients for negative- and positive-valence uPN-types respectively correspond to a shape parameter $\kappa < 1$ and $\kappa > 1$. *P*-values based on one-sided Student's *t* test. (*C*) Same as (*B*) for PC1 coefficients obtained from type-1 shuffled connectivity matrices (10^4 shuffles, 100 subsamples each). The difference between the mean PC1 coefficient for positive- and negative-valence uPN-types is statistically significant but is two orders of magnitude smaller than that seen in the empirical data (*SI Appendix*, §2.C). The shape parameter $\kappa > 1$ for all distributions. Overall, there is a lack of bias between positive and negative uPN-types in the shuffled connectivity. (*D*) Pearson's χ^2 independence test for PC1 coefficients of *individual* uPNs. PC1 coefficients of uPNs postsynaptic to coupled, behaviorally antagonistic ORNs (circles, $n = 124$). Colors: Difference between the empirical and expected number of data points in each bin, used to calculate the χ^2 statistic (*SI Appendix*, §2.E and Fig. S9 for other bin sizes). (*E*) Distribution of χ^2 statistic obtained under the null hypothesis of statistical independence of PC1 coefficient pairs (*SI Appendix*, §2.E). The empirical χ^2 statistic is significantly larger, indicating statistical dependencies between readout weights of uPNs postsynaptic to coupled ORNs. (*F*) Illustration of definition of PC1 readout angle θ. (*G*) PC1 readout angles of paired uPN-types versus ratios of three dendritic measurements of the corresponding ORN pairs, and a compound morphometric ratio (red, *SI Appendix*, §2.F). The quantities are positively correlated. *Insets*: Analogous correlation values for random pairs of positive- and negative-valence ORNs are significantly smaller than that for empirical ORN pairs (*SI Appendix*, §2.F). *P*-values computed using distribution percentiles. $^*P < 0.05, ^{**}P < 0.01, ^{***}P < 0.001$.

behaviorally antagonistic ORNs ($n = 14$ positive valence, $n = 16$ negative valence; Fig. 3 *A*, *Inset*; see *SI Appendix*, Table S1). The mean PC1 coefficient for negative-valence uPNs is significantly smaller than that of positive-valence uPNs (Fig. 3*B*, positive-valence uPNs: 0.105 ± 0.012, negative-valence uPNs: 0.055 ± 0.006). We found that these values were significantly different than the mean PC1 coefficient of a randomly chosen sample of 14 uPNs (0.083 ± 0.018, Fig. 3*B*), indicating that PC1 coefficients of positive- and negative-valence uPNs are distributed differently. Next, we fitted a gamma distribution to each set of PC1 coefficients separately (Fig. 3*B* and *SI Appendix*, §2.C). The distribution mode for negative-valence uPNs is 0, while for positive uPNs and random samples, it is strictly positive. This is indicated by the shape parameters of the corresponding gamma distributions (Fig. 3*B*)—positive uPNs: 1.45 ± 0.38, negative uPNs: 0.93 ± 0.18, random sample: 1.16 ± 0.44. Taken together, negative-valence uPNs are biased toward smaller PC1 coefficients. In contrast, this bias is negligible for shuffled connectivity matrices (Fig. 3*C* and *SI Appendix*, Fig. S6). This implies that the predicted asymmetry in readout-weights indeed arises from structured AL-to-LH connectivity, and cannot be explained simply by overall biases in the frequency of connections or connection weights.

Our model further predicts that in order to compute the amplified valence signal, the asymmetry in LH readout-weights of coupled ORNs must depend on the sensillum-specific interaction asymmetry (q). In other words, we expected downstream connectivity to mirror this specificity, beyond the overall differences between readouts of positive and negative ORNs (Fig. 3*B*). We identified 25 pairs of uPNs in the connectome that correspond to behaviorally antagonistic ORNs from the same sensillum (13). The connectivity of 16/25 pairs (64%) differed significantly between the data and shuffles (*SI Appendix*, Fig. S7), indicating dependencies between LHN targets of these uPN pairs. If AL-to-LH connectivity indeed reflects ORN coupling at the periphery, the readout-weights of pairs of behaviorally antagonistic, coupled ORNs would be distributed differently than that of random pairs of positive- and negative-valence ORNs. We thus compared the joint distribution of the 25 pairs of PC1 coefficients to the null distribution, obtained by independently sampling from the marginals of positive- and negative-valence uPNs (*SI Appendix*, Fig. S8*A* and §2.D). Notably, there is a region containing no data points (*SI Appendix*, Fig. S8*A*, red box), suggesting that the empirical data are not fully explained by the null distribution. We further computed readout-angles of the uPN pairs, which exhibit a marked discrepancy between the empirical and the null distributions (*SI Appendix*, Fig. S8*B*). Angles close to $\pi/4$, corresponding to symmetric readouts, are noticeably not represented in the distribution. However, we were unable to rule out the null hypothesis in *SI Appendix*, Fig. S8 based on statistical tests, likely due to the small number of uPN-types postsynaptic to ORNs with characterized behavioral valence (13).

Our model predicts *global biases* in readout-weights that apply at the level of input channels, i.e., uPN *types* postsynaptic to coupled ORNs. However, the predicted dependencies between readout-weights of coupled ORNs can be examined at the level of individual neurons. Therefore, we analyzed PC1 coefficients of *individual* uPNs to test the null hypothesis with increased statistical power (Fig. 3*D*). Indeed, we found a statistically significant dependence between the empirical PC1 coefficient pairs (Fig. 3*E*, see also *SI Appendix*, Fig. S9 *A* and *B*), and showed that the distribution of PC1 readout-angles is significantly different from the corresponding null distribution (*SI Appendix*, Fig. S9 *C* and *D*). Together these statistical tests demonstrate that the readout weights of ORNs reflect their coupling in the periphery.

Next, we tested the hypothesis that AL-to-LH connectivity reflects asymmetrical coupling between ORNs by studying the relationship between the readout-angles (Fig. 3*F*) and morphometric asymmetry of coupled ORN pairs (14). If our hypothesis is correct, a positive correlation is expected between these quantities, since asymmetry in ephaptic interactions arises from ORN size disparities (Eq. **2** and refs. 12, 14, and 31). We computed ratios of five dendritic measurements and two somatic measurements (data from ref. 14) for each pair of coupled ORNs. Three of the five dendritic ratios, but none of the somatic ratios, had a significant positive correlation with PC1 readout-angles when compared to the control: analogous correlation values obtained for random pairs of positive- and negative-valence ORNs (Fig. 3*G* and *SI Appendix*, Fig. S10). To quantify the overall morphometric asymmetry between coupled ORNs, we further defined a *compound morphometric ratio* as a linear combination of these three ratios. The coefficients of the compound ratio were optimized to maximize its correlation with PC1 readout-angles (Fig. 3*G* and *SI Appendix*, §2.F). The resulting correlation of $r = 0.66$ was highly significant relative to the control. Notably, the fact that somatic ratios were not positively correlated with the readout-angles (*SI Appendix*, Fig. S10) is consistent with the notion that ephaptic coupling arises from interactions between dendritic compartments (5, 12). We did not find similar correlations between morphometry and readout-angles for secondary PCs of the AL-to-LH connectivity matrix (*SI Appendix*, §2.G and Fig. S11), suggesting that the aligned valence readout is achieved primarily by the most prominent projection pattern.

We note that the randomization procedure used as a control in Fig. 3*G* generates random ORN pairs that have opposite valence, but do not correspond to the same sensillum. Therefore, the significance of the correlations between peripheral morphometric ratios and downstream readouts demonstrates that AL-to-LH projections reflect the compartmentalized structure of the olfactory periphery, *beyond* valence opponency. Based on our computational model of ephaptic interactions between coupled ORNs, we argue that the aligned readouts of the entire sensilla array allow for effective propagation of the transiently amplified valence signal extracted from *any* odorant mixture.

Peripheral Preprocessing Enables Odor Classification in the MB

Finally, we studied the functional impact of peripheral ephaptic inhibition on odor processing in the MB. AL-to-MB projections are largely random, and expansive: projection neurons (PNs) corresponding to ~50 glomeruli relay information to 2,000–2,500 Kenyon Cells (KCs) in the MB (15, 26, 38). KC responses are sparser than those of PNs (27). Expansive and sparse projections—which increase linear separability of high-dimensional stimuli—was proposed as the key mechanism supporting novel odor classification in the MB (19, 21). However, a crucial assumption in these studies is an initial clustering of stimuli in the input layer. This clustering implies that perceptually similar odors elicit similar AL glomerular responses (Fig. 4*A*), which is not consistent with empirical data obtained from fruitflies and other insects (28, 29). Here, we explored the possibility that expansion of odor representations in the MB can facilitate odor classification in the absence of initially clustered

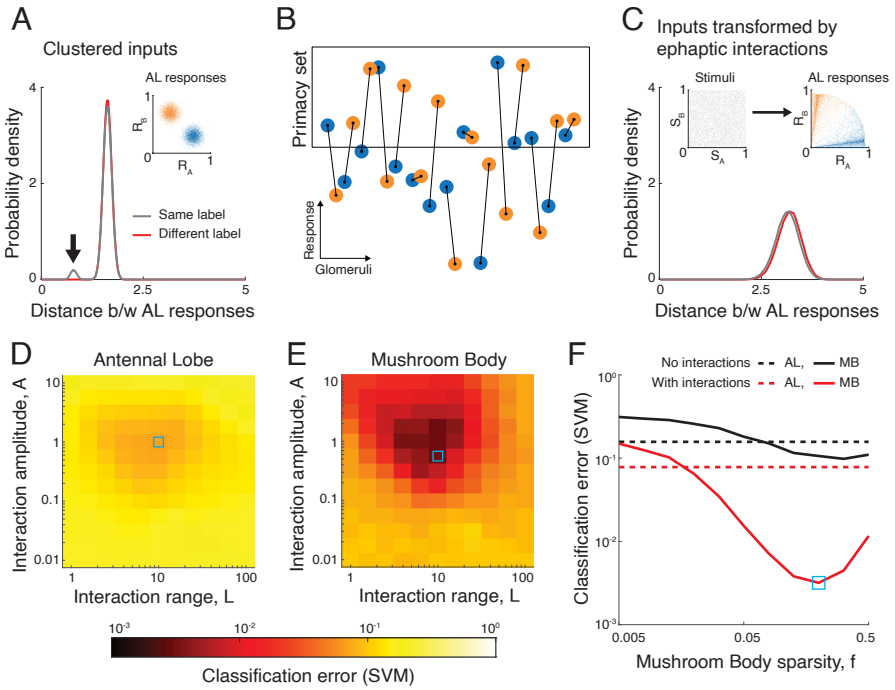

Fig. 4. Ephaptic interactions improve odor classification in the MB. (*A*) Distance between stimuli (in $N = 50$ dimensional response space) with the same (gray) or different (red) labels under the assumption of clustered AL responses. Stimuli with the same label are closer, as indicated by the peak in the gray distribution (black arrow). *Inset*: 2D illustration of clustered AL responses. (*B*) Model of primacy coding in the AL. Glomeruli with the largest responses constitute the primacy set for the stimulus, and determine the stimulus label. Black lines indicate glomeruli corresponding to ephaptically coupled ORNs. (*C*) Same as (*A*) for uniformly distributed stimuli that are transformed by ephaptic interactions and assigned a label based on the primacy set. Stimuli with the same or different labels have similar distance distributions (see also *SI Appendix*, Fig. S12). (*D*) Classification error of a support-vector machine (SVM) based on AL responses (*SI Appendix*, §3.D), as a function of the ephaptic interaction parameters. The optimal parameter combination leads to modest improvements. (*E*) Same as (*D*) for MB responses (i.e., following random expansion and sparsification, sparsity $f = 0.2$). Here, ephaptic interactions result in marked improvements (note the logarithmic color scale). (*F*) SVM classification error as a function of the sparsity of MB responses f. In the absence of ephaptic interactions (black curves), expansion and sparsification in MB does not result in a significant improvement in classification performance. See *SI Appendix* for simulation details.

stimuli. While we relax the assumption of clustering, the problem of odor classification is infeasible if the mapping from stimuli to categories is entirely random. We therefore chose a model that imposes some fine-scale structure on the mapping from AL neural responses to categories.

Given that ephaptic interactions can transiently separate ORN response trajectories (Fig. 2 *C*, *D*, *G*, and *H*), we hypothesized that preprocessing of AL responses by peripheral interactions may allow odor classification in the MB—without requiring explicit clustering. To test this hypothesis, we simulated the response of $N = 50$ ORNs selective to N uniformly distributed private odorants, each having positive or negative valence. There were $N/2$ pairs of ORNs with opposite valence that were ephaptically coupled in the simulations (Fig. 4*B* and refs. 5 and 13). For simplicity, we analyzed ORN responses at a particular "snapshot" in time. Ephaptic interactions were modeled as shifts in ORN responses relative to their initial values at $t = 0$ set by an odorant concentration "pulse." To calculate the response shift, the interactions were described by two parameters: interaction amplitude A and interaction range L (*SI Appendix*, §3.A). The range refers to the distance (in stimulus-space) from neutral stimuli for which inhibition is significant. The direction and magnitude of the shifts depend on the net stimulus valence, similar to our time-dependent model (Eq. **1**, Fig. 2 *A*–*H*, and *SI Appendix*, Fig. S12). Glomeruli were assumed to be one-to-one relays of ORN activity, as in our connectomic analysis.

We studied the classification of odor stimuli into two categories (e.g., appetitive or aversive), as in ref. 19. In contrast to previous studies (19), we *did not* assume that glomerular responses are clustered such that nearby responses are very likely to belong to the same category (Fig. 4*A*). Instead, stimuli were assigned category labels based on the principle of *primacy coding* in the AL (Fig. 4*B*). Primacy coding implies that the identity of an odor is encoded by the glomeruli that respond most strongly to it [i.e., its "primacy set," which is concentration invariant (39, 40)]. Thus, we labeled a stimulus as $+1$ if the number of positive-valence glomeruli was greater than negative-valence glomeruli in its primacy set, and -1 otherwise (Fig. 4 *C*, *Inset*). Note that this scheme for choosing labels imparts a degree of continuity on the mapping from stimuli to categories, since a change in the label of a stimulus requires changes in glomerular responses that are sufficiently large to alter the primacy set. Our analysis showed that ephaptic interactions impose a subtle structure on the distribution of glomerular responses (Fig. 4 *A* and *C*), which can be quantified using an *effective* clustering coefficient (*SI Appendix*, Figure S12). However, unlike the scenario of clustered inputs, once the uniform stimulus distribution is transformed by ephaptic coupling and labels are determined based on the primacy set, nearby glomerular responses are almost equally likely to have the same or different labels (Fig. 4*C*), consistent with refs. 28 and 29.

Next, we numerically evaluated classification performance in the AL and in the MB. We used responses computed both with and without ephaptic inhibition, for a range of A, L values. Our model allows us to systematically test the effect of ephaptic coupling, which would not be possible if we directly used empirical ORN responses to a battery of odor stimuli

(41), since those responses already include the effect of ephaptic coupling. Performance was defined as the error of a support vector machine (SVM) classifier that was trained to optimally separate neural responses with +1 and −1 labels. Training the SVM classifier mimics synaptic-plasticity-based learning in the MB (19, 21). When considering AL responses, we identified a region of A, L values where interactions lead to modest improvements (Fig. 4*D*). Remarkably, we found that random expansion of the preprocessed AL responses to the MB layer, followed by sparsification of the KC responses [(19), *SI Appendix*, §3.D] resulted in a significant improvement in classification performance (Fig. 4 *E* and *F*). In contrast, random expansion of AL responses calculated without ephaptic interactions resulted in no such improvements (Fig. 4*F*). Quantitatively similar results were obtained when we used a perceptron rather than a SVM classifier (*SI Appendix*, Fig. S13), suggesting that the improved separability is not specific to a particular definition of the learning rule or the classification error. Similar correspondence between different types of linear classifiers was reported in ref. 19. Thus, ephaptic interactions can render classification following random expansion and sparsification effective, despite the absence of a clustered input layer.

Notably, natural odors that animals encounter likely contain mixtures of innately meaningful and neutral components. We therefore explored classification performance in the MB as a function of the probability that each glomerulus' innate valence is in agreement with its contribution to the stimulus label given during associative learning (*SI Appendix*, Fig. S14). Performance is maximally enhanced by ephaptic interactions when the agreement is complete, and gradually decreases as the agreement probability decreases. In the case where glomeruli in the primacy set "vote" independently of their innate valence, ephaptic interactions still moderately enhanced classification (*SI Appendix*, Fig. S14 *A–D* and §3.E). Intuitively, this improvement arises because distances between AL representations corresponding to stimuli with opposite labels can increase if those stimuli activate ephaptically coupled ORNs. Additionally, we showed that ephaptic interactions facilitate classification in the MB even when the glomerular responses are transformed by a random rotation (*SI Appendix*, Fig. S15 and §3.F). In this scenario, the response of each "virtual" glomerulus signals the activation of a random combination of all real glomeruli. Taken together, these results imply that information preprocessing by coupled ORNs can potentially facilitate learning of *arbitrary* stimulus–label associations in the MB, though the degree of facilitation depends on the innate valence of glomeruli activated by the odor.

Discussion

Here, using a combination of mathematical modeling and analysis of experimental data, we have shown how peripheral interactions between ORNs work in coordination with the wiring diagram in the central brain to optimize complex sensory computations. Previously, anatomical wiring in fruitflies was shown to implement *specific* sensorimotor computations, for example, detection of two-dimensional visual motion within a restricted field of view (42), detection of coincident noxious stimuli (43), and evaluation of heading direction (44). By using our model of *peripheral* preprocessing and focusing on the "first pass" of feedforward processing, we made predictions for *central* connectivity necessary to perform robust, valence-based classification of complex odors. By validating these predictions, we demonstrated how the wiring diagram implements a *general purpose* computation: extracting the valence of arbitrary odors. Our study further suggests that a sparse and random expansion of sensory information (15, 19, 20) can enable classification, if AL responses are transformed by peripheral interactions. We have demonstrated that this scheme can be effective without requiring clustering, i.e., without stimuli of the same category being contained within a radius (45). Notably, ref. 19, achieved a significant boost to MB classification performance by choosing the AL-to-MB projection matrix to be partly structured and partly random. The structured component of that matrix depends on the cluster structure of the stimuli, which is absent in our model. We found that transforming initially clustered stimuli by ephaptic coupling prior to classification can negatively affect performance, since the interactions can mix clusters with different labels (*SI Appendix*, Fig. S12 *E–G*). We hypothesize that further improvements to performance can be achieved in the presence of ephaptic coupling by identifying analogous partly structured matrices that take into account the effect of the interactions (*SI Appendix*, Fig. S12 *A–D*).

Future experiments could further test our theoretical predictions, by comparing the preference of animals to odor mixtures that activate ORNs in multiple sensilla with a range of concentration ratios (analogous to a psychometric curve). We predict that the preference curve will be steep (i.e., small deviations from the neutral ratio induce more selectivity) in scenarios where compartmentalized ORNs in each of the sensilla are strongly activated. The reason is that the resulting strong ephaptic interaction will amplify the weak valence signal (Fig. 2 *A–D*). This scenario can be compared to activation of noncompartmentalized ORNs across multiple sensilla, which nevertheless keeps the overall valence of the odor mixture close to neutral. In this case, ephaptic interactions are weak, so we predict a relatively flat preference curve.

One important feature of odor processing in insects which was not explicitly included in our model is lateral inhibition in the AL. PN responses in the AL are modulated by slow and broadly tuned (but not global) inhibition (46, 47). This inhibition may have two important roles in the context of our modeling work. First, it could implement the normalization necessary to achieve primacy-coding in the AL, which we use to assign labels to stimuli in Fig. 4. Second, we note that to compute the net-valence, signals from positive- and negative-valence ORNs should have readout weights with opposite sign Eq. **2**. Broad inhibition could mediate a shift of the baseline value of the readouts. Since negative valence ORNs have smaller readout weights (Fig. 3*B*), an upward shift of the baseline would give the "correct" sign needed to compute the valence.

We believe that our study can inspire future research into other computationally inexpensive transformations of the input layer that can similarly enable classification by the random expansion and sparsification architecture (21, 48). Such architectures could support powerful sensory computations without relying on costly training or metabolically inefficient hierarchical processing. Furthermore, our findings highlight the importance of peripheral processing for central sensory computations, which may have implications in biological systems beyond olfaction. Indeed, functionally similar lateral inhibition was reported between the R7 and R8 photoreceptors (7), which have different spectral sensitivity profiles in the *Drosophila* compound eye. Future work may reveal how such color-opponent preprocessing contributes to initiation of appropriate visually guided behaviors in flies.

Materials and Methods

The sections of *Materials and Methods* correspond to sections with the same numbers in *SI Appendix*, where additional detail is provided.

1. Reduced Nonlinear Model of a Sensillum, Related to Fig. 2.

1.1. Derivation of the analytical solution to the coupled ORN firing rates. We find closed-form solutions for firing-rate trajectories following pulse stimulation ($s_{A,B}(t) = S_{A,B}\delta(t)$),

$$x_A(t) = S_A e^{-\frac{t}{\tau}} \left(\frac{S_A^n - qS_B^n}{S_A^n - qS_B^n \exp\left[-K(S_A^n - qS_B^n)(1 - e^{-\frac{nt}{\tau}})\right]} \right)^{\frac{1}{n}}$$

$$x_B(t) = S_B e^{-\frac{t}{\tau}} \left(\frac{S_A^n - qS_B^n}{S_A^n \exp\left[+K(S_A^n - qS_B^n)(1 - e^{-\frac{nt}{\tau}})\right] - qS_B^n} \right)^{\frac{1}{n}}$$

The parameters K, q, n are described in the text. The solutions were derived based on identifying an invariant of the dynamics in Eq. **1**, as described in detail in *SI Appendix*, §1.A.

1.2. Transient valence amplification. The solutions to the dynamical model (*SI Appendix*, §1.A) show that for stimulus mixtures satisfying $S_A^n = qS_B^n$, the firing rates x_A, x_B decay to zero along a straight line. In other words, for these stimuli, ORNs A and B inhibit each other equally. We interpreted stimuli satisfying this condition as *neutral* and defined the weighted difference $S_A - q^{1/n}S_B$ as the net-valence of the stimulus mixture. The time-dependent valence amplification [$\alpha_q(t)$, Eq. **2**] was calculated by directly substituting the analytical solutions for $x_{A,B}$.

1.3. Discrimination factor. We first defined the time evolution of the angle of firing rate trajectories as $\tan[\phi(t)] = x_B(t)/x_A(t)$. The initial value of ϕ is given by the relative strengths of the odorants, and at long times ($t \to \infty$) it saturates to a value that depends on the stimuli as well as ephaptic interaction parameters. To understand whether ephaptic interactions can aid in discrimination between responses to similar odor mixtures (i.e., stimuli with similar ratios S_B/S_A), we defined stimulus sensitivity $\sigma(t)$ as

$$\sigma(t) = \frac{\partial\phi(t)}{\partial\phi(0)}\Big|_S$$

evaluated at fixed stimulus strength $S = (S_A^2 + S_B^2)^{1/2}$. The discrimination factor Δ is defined as the maximum sensitivity: $\Delta = \max_t \sigma(t)$. Thus, $\Delta \geq \sigma(t = 0) = 1$. Large Δ implies that two similar odor mixtures (of the same strength) can be easily discriminated by ORN responses. *SI Appendix*, §1.C contains analytical formula for the quantities described here. One sees from the closed-form solutions that discrimination is maximal for neutral stimuli ($S_A^n \approx qS_B^n$), for which transient amplification is also maximal.

1.4. Transient dynamics in the stimulus onset scenario. We considered the response of the model to a linear ramp stimulus, $s_{A,B}(t) = S_{A,B}t/T$, which attains its final value at $t = T$. We present closed-form solutions in the absence of ephaptic coupling ($K = 0$). One sees from the solutions that for fast ramping relative to the membrane time-constant ($T \ll \tau$), the firing rates are very small $x_{A,B}(T) \ll S_{A,B}$ at the end of the onset transient. For slow ramping ($T \gg \tau$), the firing rates track the stimulus, $x_{A,B}(T) \approx S_{A,B}$. To study the effect of ephaptic interactions on the response of the system, we focused on the intermediate regime, $T \approx \tau$. We numerically solved Eq. **1** for ramp stimuli to find the firing-rate trajectories $x_{A,B}(t)$ in the presence of ephaptic inhibition (Fig. 2*E*). Valence amplification ($\alpha_q(t)$) and discrimination factor (Δ) were analogously defined and computed numerically.

1.5. Model fitting to electrophysiological data. The ab2 sensillum consists of two olfactory receptor neurons, ab2A (Or59b) and ab2B (Or85a), that respond selectively to odorants methyl acetate and ethyl 3-hydroxy butyrate respectively. In vivo single sensillum recordings were performed in response to a 0.5 s pulse of odorant mixtures at different dilution ratios [Fig. 2*J*, six dilution ratios ("conditions"), $n = 9$ trials each]. Data are taken from ref. 12. We assumed that the system is approximately at steady-state during odor presentation, and estimated stimuli strengths ($S_{A,B}$) for different odorant dilutions using ORN responses to single odorants (*SI Appendix*, Fig. S2 *A–C*) as described in *SI Appendix*, §1.E. We utilized gradient descent methods to find fit parameters (K, q, n) that minimize the fitting error for all six odor mixtures (*SI Appendix*, Figure S2 *E–G*). A distribution of fitting parameters was generated by repeatedly dropping data from six random trials (on average, one trial from each condition; repeated 1,000 times) in the implementation of the fit (Fig. 2*K* and *SI Appendix*, Fig. S2*D*).

2. Analysis of AL-to-LH Connections, Related to Fig. 3.

2.1. Analysis at the level of neuronal-types. The Hemibrain connectome (ref. 17, v1.2.1) contains 90 uPN types and 689 LHN types. uPN- and LHN-types not participating in any AL-to-LH projections were excluded from our analysis, leaving 71 uPN-types and 673 LHN-types remaining. In this section, we describe our analysis which showed that on average, uPNs of the *same* type shared 75 postsynaptic LHN-types, while uPNs of *different* types innervating *same* glomerulus shared only 39 postsynaptic LHN-types (*SI Appendix*, Fig. S3). On the other hand, uPNs innervating different glomeruli had, on average, 32 LHN targets in common, suggesting that distinct uPN-types innervating the same glomerulus are nearly as dissimilar uPN-types innervating different glomeruli. These results motivated the analysis of AL-to-LH connectivity at the level of distinct uPN-types rather than glomeruli.

2.2. Connectivity shuffles with conserved uPN-type and LHN-type statistics. We constructed shuffled versions of the 71-by-673 empirical connectivity matrix by randomizing connections between uPN- and LHN-types. We constrained the shuffles to ensure that each type of uPN and LHN participates in the same number of connections as in the original data (analogous to ref. 15). This preserves the frequency of incoming and outgoing connections in the empirical data (*SI Appendix*, Fig. S4), while removing any nonrandom, structured patterns of AL-to-LH connectivity. We implemented two types of shuffles, which differed in the constraint on connection weights that was imposed: **type-1** shuffles preserved the incoming connection weights for each LHN-type, while **type-2** shuffles preserved outgoing connection weights for each uPN-type (*SI Appendix*, Fig. S5). The only matrix that satisfies both constraints on connection weights is the empirical connectivity matrix itself. The two types of shuffles allow us to probe the significance of different properties of the empirical connectivity matrix, which is described in detail in *SI Appendix*, §2.B.

2.3. Fitting gamma distribution to PC1 coefficients. We fitted a gamma distribution to PC1 coefficients of positive- and negative-valence carrying uPN-types separately using maximum-likelihood estimation (Fig. 3*B*). In the empirical data, one negative-valence uPN-type has a small, negative PC1 coefficient (−0.0006, small relative to the distribution mean 0.057). We replaced that value with 0 to obtain a valid fit for the gamma distribution (Fig. 3*B*). In *SI Appendix*, §2.C, we describe an alternate fitting procedure based on shifting all PC1 coefficients that yielded a similar result. We further describe the subsampling procedure used in Fig. 3 *B* and *C* and *SI Appendix*, Fig. S6. Here, for both empirical and shuffled connectivity, we subsampled negative- and positive-valence uPNs by dropping 10% of PC1 coefficients (repeated 100 times). As a control, we also generated 100 samples of 14 randomly chosen uPN-types. For each subsample and random sample, we calculated the mean PC1 coefficient and fit a gamma distribution.

2.4. Null distribution for PC1 coefficient pairs. We calculated the joint distribution of PC1 coefficients corresponding to coupled ORN pairs under the null hypothesis that the coefficients are sampled independently from distinct distributions of positive- and negative-valence uPN-types. The null distribution was computed as the product of the marginals of positive and negative uPN-types. To take into account the fact that some uPN-types participate in multiple coupled pairs (*SI Appendix*, Fig. S7), the marginals were calculated by weighing each uPN-type by the number of pairs it participates in. We further computed the null distribution of PC1 readout angles using the generalized beta distribution. The resulting null distributions are shown in *SI Appendix*, Fig. S8 and described in detail in *SI Appendix*, §2.D.

2.5. Analysis of AL-to-LH connectivity at the level of individual neurons. We describe the analysis of AL-to-LH connectivity at the level of individual uPNs and LHNs, to gain enhanced statistical power. As in *SI Appendix*, §2.A, we neglected neurons that did not participate in any AL-to-LH connections, resulting in a

connectivity matrix with 136 uPNs and 1442 LHNs. We describe in detail the implementation of Pearson's χ^2 test of independence, which was used to rule out the null hypothesis that readout weights of coupled ORNs are statistically independent (Fig. 3 *D* and *E* and *SI Appendix*, Fig. S9). We also describe the implementation of the Kolmogorov–Smirnov goodness-of-fit test to show that PC1 readout angles do not follow a null distribution. The *P*-values for both tests were obtained by simulations, as detailed in *SI Appendix*, §2.E.

2.6. Compound morphometric ratio. We computed a ratio of measured values for coupled negative- and positive-valence ORNs for each morphometric quantity measured in ref. 14. We then calculated the correlation of the resulting morphometric ratios with the PC1 readout-angles of the corresponding uPN-pairs (Fig. 3*G* and *SI Appendix*, Fig. S10). To test the significance of the correlations, we repeated the above procedure for randomly chosen pairs of positive- and negative-valence uPNs. The pairs were generated by randomly pairing uPNs of *opposite valence*, with the constraint that each uPN participates in the same number of pairs as in the original data (repeated 10^5 times). We found that inner-dendrite volume, outer-dendrite surface area, and dendrite branch number ratios had a significant correlation with readout-angles compared to the controls (*P* values computed as distribution percentiles). We further constructed a *compound* morphometric ratio using a linear combination of the aforementioned three dendritic size measurements (Fig. 3*G*). The coefficients in the linear combination were chosen to maximize the correlation between the compound morphometric ratio and PC1 readout-angles. The optimal coefficients for each morphometric ratio are 0.64 (inner dendrite volume), 0.15 (outer dendrite surface area), and 0.21 (dendritic branch number).

2.7. Analysis of secondary PCs. We identified two additional PCs, PC4, and PC36 that showed a significant difference between the average coefficients for positive- and negative-valence uPNs relative to shuffles. However, neither PC4 nor PC36 readout-angles were significantly correlated with *any* morphometric ratios of the corresponding ORNs (*SI Appendix*, Fig. S11).

3. Model of Stimulus Classification in the AL-to-MB Pathway, Related to Fig. 4.

3.1. Snapshot model of AL responses. We analyzed ORN responses at a particular snapshot in time. Here, we computed the response of each pair of coupled ORNs as follows. Let odor stimuli for one pair of coupled ORNs be denoted by $S_A,\ S_B$, with amplitude $S = (S_A^2 + S_B^2)^{1/2}$ and angle $\phi_S = \tan^{-1}(S_B/S_A)$. The amplitude ($R$) and angle ($\phi_r$) of the transformed response were computed as,

$$R(S, \phi_S) = \frac{S}{1 + AS^2 \cos[2(\phi_S - \frac{\pi}{4})]},$$

$$\phi_R(S, \phi_S) - \phi_S = AL^2 \left(\phi_S - \frac{\pi}{4}\right) e^{-L|\phi_S - \frac{\pi}{4}|}$$

Here, the parameters $A,\ L$ denote the amplitude and range (in stimulus space) of the ephaptic interactions. An illustration of the transformation from $S_A,\ S_B$ to $R_A,\ R_B$ is shown in Fig. 4*C*, which is similar to the transformation imposed by our time-dependent model (*SI Appendix*, Fig. S12*A*). The responses $R_A,\ R_B$ can be thought of as the ORN firing rates $x_A,\ x_B$ at time t after stimulation.

3.2. Assigning labels to stimuli based on the primacy set. AL responses in our model were not clustered into groups with different labels, as in ref. 19. Instead, we assigned labels to each $N = 50$ dimensional stimulus based on its primacy set, composed of the $N/2$ most active glomeruli. Previous work has shown that the primacy-code is most informative when the size of the primacy set equals half the dimension of the stimulus space (40). Glomeruli corresponding to coupled ORNs were assigned opposite valence. The label of a stimulus was set to $+1/-1$ if the majority of glomeruli in the primacy set had positive/negative-valence, respectively. Notably, the contribution of each glomerulus' valence to the label of the stimulus is small (~2%), since only half of the glomeruli participate in the primacy set, and within the primacy set each glomerulus' counts as $2/N$ of the vote.

We carried out two additional sets of simulations where the geometrical structure imposed on AL responses by ephaptic interactions is conserved, but the label-assignment scheme is modified: in *SI Appendix*, Fig. S14, the stimulus label was not based on a valence majority in the primacy set, but rather on a random (predetermined) split of glomeruli into two groups (triangles vs. squares); in *SI Appendix*, Fig. S15, AL responses were randomly rotated prior to the assignment of stimulus labels, such that the response of each virtual glomerulus is a linear combination of both positive and negative-valence real glomeruli responses. The virtual glomeruli then decide the stimulus label based on primacy (triangles vs. squares).

3.3. Bhattacharyya distance. We compared the structure of AL responses transformed by ephaptic coupling to that of responses obtained under the assumption of clustering (19). For both cases, we computed Euclidean distances (in $N = 50$ dimensional response space) between stimuli with the same label, and stimuli with different labels. In the case of clustered AL responses, the distributions of the two distances were significantly different (Fig. 4*A*). This is because when responses are clustered, nearby stimuli always have the same label. In contrast, the two distance distributions are very similar for AL responses generated by our model (Fig. 4*C*). This is quantified by the Bhattacharyya distance between the distributions (*SI Appendix*, Fig. S12 *B* and *C*). We estimated the degree of effective clustering imposed by ephaptic coupling by matching Bhattacharyya distances computed based on AL responses generated by our model and by the clustered model (*SI Appendix*, Fig. S12*D*).

3.4. Classification following random expansion and sparsification. We used the snapshot model to generate responses to sets of $P = 2{,}500$ N-dimensional stimuli. We then evaluated the classification performance based on stimuli labels using a SVM classifier (Fig. 4) and a perceptron classifier (*SI Appendix*, Fig. S13). At the level of the AL, we repeated these simulations 100 times for each combination of the parameters $A,\ L$. We compared these results with classification performance in the MB. We computed responses of $N_M = 2{,}000$ KCs in the MB as follows. For each simulation, we considered a $N_M \times N$ random matrix, which represents random expansion from the AL to MB. Each element in the matrix is sampled i.i.d. from a standard normal distribution. The KC responses to each stimulus were then generated by multiplying the N dimensional vector of AL response with the random matrix. The sparsity of KC responses was determined by the parameter f. We modeled KC responses to be binary (19), such that a neuron was active if it was among the fN_M neurons with the strongest input. We varied the sparsity level f systematically and found that $f \approx 0.2$ led to optimal classification performance. The simulations were repeated 100 times for each parameter combination ($A,\ L,\ f$).

We additionally evaluated classification performance in a scenario where stimuli were initially clustered as in ref. 19, with labels determined by cluster identity. The snapshot model of ephaptic coupling was then used to transform these stimuli to AL responses, which were then randomly expanded giving the sparse MB responses (*SI Appendix*, Fig. S12 *E*–*G*). We used a completely random expansion to transform AL to MB responses, instead of the partly structured and partly random connectivity shown in ref. 19 to give better performance. The reason is that the structured part of the connectivity in ref. 19 depends on the cluster *location* which in our case was shifted by ephaptic coupling.

3.5. Classification in scenario of mixtures of innately meaningful and neutral labels. We explored the possibility that ephaptic interactions can facilitate odor classification during associative learning. In such a scenario, the vote of a glomerulus in the primacy set for determining stimulus labels may be independent of the innate valence of the presynaptic ORNs. We performed simulations where a fraction of glomeruli voted based on a random partition into predetermined groups (*SI Appendix*, Fig. S14). A consequence of this randomization is that glomeruli corresponding to ephaptically coupled ORNs may not have the opposite vote. We varied the fraction of randomly partitioned glomeruli, and examined the effect of ephaptic interactions on classification performance in the MB relative to the scenario with no interactions (*SI Appendix*, Fig. S14*E*).

4. Supplementary Discussion on a Potential Functional Role of Secondary PCs. This section includes additional discussion on the possibility that secondary patterns of structured connectivity in the AL-to-LH pathway could facilitate initiation context-dependent behaviors.

Data, Materials, and Software Availability. Previously published data were used for this work (12–14, 17).

ACKNOWLEDGMENTS. We thank J.W. Wang and D. Kleinfeld for useful discussions and comments on the manuscript. This work was supported by DARPA grant D21AP10162-00 (J.A.), DOE grant DE-SC0022042 (J.A.), NIH grants R01NS135853 (J.A.), R21AI169343, R21DC020536, R01DC016466 (C.-Y.S.), an Innovative Research Grant from the Kavli Institute for Brain and Mind (UC San Diego) (awarded jointly to C.-Y.S., J.A.). P.P. thanks the UC San Diego Friends of the International Center, and UC San Diego GPSA for support.

Author affiliations: [a]Department of Physics, University of California, San Diego, La Jolla, CA 92093; and [b]Department of Neurobiology, University of California, San Diego, La Jolla, CA 92093

1. A. Celani, E. Villermaux, M. Vergassola, Odor landscapes in turbulent environments. *Phys. Rev. X* **4**, 041015 (2014).
2. G. Reddy, V. N. Murthy, M. Vergassola, Olfactory sensing and navigation in turbulent environments. *Annu. Rev. Condens. Matter Phys.* **13**, 191–213 (2022).
3. J. Merel, M. Botvinick, G. Wayne, Hierarchical motor control in mammals and machines. *Nat. Commun.* **10**, 5489 (2019).
4. J. H. Siegle *et al*., Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* **592**, 86–92 (2021).
5. C. Y. Su, K. Menuz, J. Reisert, J. R. Carlson, Non-synaptic inhibition between grouped neurons in an olfactory circuit. *Nature* **492**, 66–71 (2012).
6. S. D. Roper, N. Chaudhari, Taste buds: Cells, signals and synapses. *Nat. Rev. Neurosci.* **18**, 485–497 (2017).
7. C. Schnaitmann *et al*., Color processing in the early visual system of *Drosophila*. *Cell* **172**, 318–330 (2018).
8. R. Ng, S. T. Wu, C. Y. Su, Neuronal compartmentalization: A means to integrate sensory input at the earliest stage of information processing? *BioEssays* **42**, 2000026 (2020).
9. E. A. Hallem, M. G. Ho, J. R. Carlson, The molecular basis of odor coding in the *Drosophila* antenna. *Cell* **117**, 965–979 (2004).
10. A. Couto, M. Alenius, B. J. Dickson, Molecular, anatomical, and functional organization of the *Drosophila* olfactory system. *Curr. Biol.* **15**, 1535–1547 (2005).
11. M. De Bruyne, T. C. Baker, Odor detection in insects: Volatile codes. *J. Chem. Ecol.* **34**, 882–897 (2008).
12. Y. Zhang *et al*., Asymmetric ephaptic inhibition between compartmentalized olfactory receptor neurons. *Nat. Commun.* **10**, 1–16 (2019).
13. S. T. Wu *et al*., Valence opponency in peripheral olfactory processing. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2120134119 (2022).
14. C. Nava Gonzales *et al*., Systematic morphological and morphometric analysis of identified olfactory receptor neurons in *Drosophila melanogaster*. *eLife* **10**, e69896 (2021).
15. S. J. Caron, V. Ruta, L. Abbott, R. Axel, Random convergence of olfactory inputs in the *Drosophila* mushroom body. *Nature* **497**, 113–117 (2013).
16. J. M. Jeanne, M. Fişek, R. I. Wilson, The organization of projections from olfactory glomeruli onto higher-order neurons. *Neuron* **98**, 1198–1213 (2018).
17. L. K. Scheffer *et al*., A connectome and analysis of the adult *Drosophila* central brain. *eLife* **9**, e57443 (2020).
18. P. Schlegel *et al*., Information flow, cell types and stereotypy in a full olfactory connectome. *eLife* **10**, e66018 (2021).
19. B. Babadi, H. Sompolinsky, Sparseness and expansion in sensory representations. *Neuron* **83**, 1213–1226 (2014).
20. A. Litwin-Kumar, K. D. Harris, R. Axel, H. Sompolinsky, L. Abbott, Optimal degrees of synaptic connectivity. *Neuron* **93**, 1153–1164 (2017).
21. S. Dasgupta, C. F. Stevens, S. Navlakha, A neural algorithm for a fundamental computing problem. *Science* **358**, 793–796 (2017).
22. G. Heimbeck, V. Bugnon, N. Gendre, A. Keller, R. F. Stocker, A central neural circuit for experience-independent olfactory and courtship behavior in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 15336–15341 (2001).
23. G. S. Suh *et al*., A single population of olfactory sensory neurons mediates an innate avoidance behaviour in *Drosophila*. *Nature* **431**, 854–859 (2004).
24. H. H. Lin, L. A. Chu, T. F. Fu, B. J. Dickson, A. S. Chiang, Parallel neural pathways mediate CO_2 avoidance responses in *Drosophila*. *Science* **340**, 1338–1341 (2013).
25. J. L. Semmelhack, J. W. Wang, Select *Drosophila* glomeruli mediate innate olfactory attraction and aversion. *Nature* **459**, 218–223 (2009).
26. T. T. Hayashi *et al*., Mushroom body input connections form independently of sensory activity in *Drosophila melanogaster*. *Curr. Biol.* **32**, 4000–4012 (2022).
27. G. C. Turner, M. Bazhenov, G. Laurent, Olfactory representations by *Drosophila* mushroom body neurons. *J. Neurophysiol.* **99**, 734–746 (2008).
28. M. Stopfer, V. Jayaraman, G. Laurent, Intensity versus identity coding in an olfactory system. *Neuron* **39**, 991–1004 (2003).
29. J. W. Wang, A. M. Wong, J. Flores, L. B. Vosshall, R. Axel, Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain. *Cell* **112**, 271–282 (2003).
30. K. I. Nagel, R. I. Wilson, Biophysical mechanisms underlying olfactory receptor neuron dynamics. *Nat. Neurosci.* **14**, 208–216 (2011).
31. C. A. Anastassiou, C. Koch, Ephaptic coupling to endogenous electric field activity: Why bother? *Curr. Opin. Neurobiol.* **31**, 95–103 (2015).
32. A. Arvanitaki, Effects evoked in an axon by the activity of a contiguous one. *J. Neurophysiol.* **5**, 89–108 (1942).
33. M. Vergassola, E. Villermaux, B. I. Shraiman, 'Infotaxis' as a strategy for searching without gradients. *Nature* **445**, 406–409 (2007).
34. G. Bondanelli, S. Ostojic, Coding with transient trajectories in recurrent neural networks. *PLoS Comput. Biol.* **16**, 1–36 (2020).
35. O. Mazor, G. Laurent, Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons. *Neuron* **48**, 661–673 (2005).
36. M. J. Dolan *et al*., Neurogenetic dissection of the *Drosophila* lateral horn reveals major outputs, diverse behavioural functions, and interactions with the mushroom body. *eLife* **8**, e43079 (2019).
37. A. S. Bates *et al*., Complete connectomic reconstruction of olfactory projection neurons in the fly brain. *Curr. Biol.* **30**, 3183–3199 (2020).
38. Y. Aso *et al*., The neuronal architecture of the mushroom body provides a logic for associative learning. *eLife* **3**, e04577 (2014).
39. C. D. Wilson, G. O. Serrano, A. A. Koulakov, D. Rinberg, A primacy code for odor identity. *Nat. Commun.* **8**, 1–10 (2017).
40. D. Zwicker, Primacy coding facilitates effective odor discrimination when receptor sensitivities are tuned. *PLoS Comput. Biol.* **15**, e1007188 (2019).
41. E. A. Hallem, J. R. Carlson, Coding of odors by a receptor repertoire. *Cell* **125**, 143–160 (2006).
42. Sy. Takemura *et al*., A visual motion detection circuit suggested by *Drosophila* connectomics. *Nature* **500**, 175–181 (2013).
43. T. Ohyama *et al*., A multilevel multimodal circuit enhances action selection in *Drosophila*. *Nature* **520**, 633–639 (2015).
44. C. Lyu, L. Abbott, G. Maimon, Building an allocentric travelling direction signal via vector computation. *Nature* **601**, 92–97 (2022).
45. S. Chung, D. D. Lee, H. Sompolinsky, Classification and geometry of general perceptual manifolds. *Phys. Rev. X* **8**, 031003 (2018).
46. R. I. Wilson, G. C. Turner, G. Laurent, Transformation of olfactory representations in the *Drosophila* antennal lobe. *Science* **303**, 366–370 (2004).
47. A. F. Silbering, C. G. Galizia, Processing of odor mixtures in the *Drosophila* antennal lobe reveals both global inhibition and glomerulus-specific interactions. *J. Neurosci.* **27**, 11966–11977 (2007).
48. A. Thomas, S. Dasgupta, T. Rosing, A theoretical perspective on hyperdimensional computing. *J. Artif. Intell. Res.* **72**, 215–249 (2021).