Convergent motifs of early olfactory processing are recapitulated by layer-wise efficient coding

Juan Carlos Fernández del Castillo,[1,2] Farhad Pashakhanloo,[1]
Venkatesh N. Murthy,[1,3,4,*] and Jacob A. Zavatone-Veth[1,5,*]

[1] *Center for Brain Science, Harvard University, Cambridge, MA, 02138, USA*
[2] *Program in Biophysics, Harvard University, Cambridge, MA, 02138, USA*
[3] *Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, 02138, USA*
[4] *Kempner Institute for the Study of Natural and Artificial Intelligence, Harvard University, Cambridge, MA, 02138, USA*
[5] *Society of Fellows, Harvard University, Cambridge, MA, 02138, USA*

The architecture of early olfactory processing is a striking example of convergent evolution. Typically, a panel of broadly tuned receptors is selectively expressed in sensory neurons (each neuron expressing only one receptor), and each glomerulus receives projections from just one neuron type. Taken together, these three motifs—broad receptors, selective expression, and glomerular convergence—constitute "canonical olfaction," since a number of model organisms including mice and flies exhibit these features. The emergence of this distinctive architecture across evolutionary lineages suggests that it may be optimized for information processing, an idea known as efficient coding. In this work, we show that by maximizing mutual information one layer at a time, efficient coding recovers several features of canonical olfactory processing under realistic biophysical assumptions. We also explore the settings in which noncanonical olfaction may be advantageous. Along the way, we make several predictions relating olfactory circuits to features of receptor families and the olfactory environment.

I. INTRODUCTION

Chemosensation is our oldest sensory modality, and for most animals it is the primary means of sensing the environment. The chemoreceptors underlying this process have a rich evolutionary history which mirrors the specialization for diverse habitats across species [1–5]. Chemoreceptors evolve with dizzying speed, evincing their position at the interface between the organism and an ever-changing chemical landscape [6, 7]. Surprisingly, however, the olfactory circuit in which these receptors are embedded has deep similarities across vertebrates and invertebrates [8, 9]. In both lineages, organisms have evolved a large repertoire of receptors, many of them broadly tuned (though some receptors are specialists for a given ligand of high importance) [10, 11]. Each primary sensory neuron typically expresses just one receptor, and only neurons expressing the same receptor converge onto a given olfactory glomerulus.

The transcriptional and wiring mechanisms for achieving this circuit organization vary widely across animals. This diversity suggests that these three motifs (broadly-tuned receptors, one neuron-one receptor, and glomerular convergence) are the result of strong selective pressure, rather than evolutionary chance or molecular constraints. Why might this architecture be optimal?

To answer this question, we must specify an objective. Organisms rely on olfaction for many of their basic needs, including mating, feeding, and avoiding predation. From a computational perspective, this means that animals need to solve a wide array of tasks, including discriminating odor identity [11–13], segmenting odor landscapes [14–18], and matching single odorants to a known mixture ("pattern completion") [19]. Complicating the picture further are the rich temporal dynamics of olfactory stimuli, which induce correspondingly rich dynamics even in the earliest levels of sensory processing [14, 20, 21].

In light of this complexity, we consider a task-agnostic objective: maximizing mutual information. The idea that early sensory processing is organized to maximize mutual information between stimulus and neural representation is known as efficient coding [22, 23]. Of course it is true that in practice, animals benefit from discarding irrelevant information. But critically, the relevance of information is often context- and task-dependent. Since higher areas of the animal brain can only access the olfactory stimulus through the glomeruli, it seems likely that up to the glomerular layer, information maximization is a plausible approximate objective. Many previous works have applied efficient coding principles to olfaction under this reasoning [24–29].

In this paper, we model olfactory stimuli and receptor activity with widely-used existing models that capture the relevant biophysics while remaining computationally tractable [24, 26]. To allow tractable analysis, we ignore temporal response dynamics in this work (but see the Discussion). Within this framework, we leverage recent technical advances from the domain of unsupervised representation learning to numerically maximize a proxy for the mutual information between the stimulus and the corresponding neural response. We do this with respect to three matrices: the receptor by odorant sensing matrix W, the neuron by receptor expression matrix E, and the

* To whom correspondence may be addressed: vnmurthy@fas.harvard.edu (V.N.M.) or jzavatoneveth@fas.harvard.edu (J.A.Z.-V.).

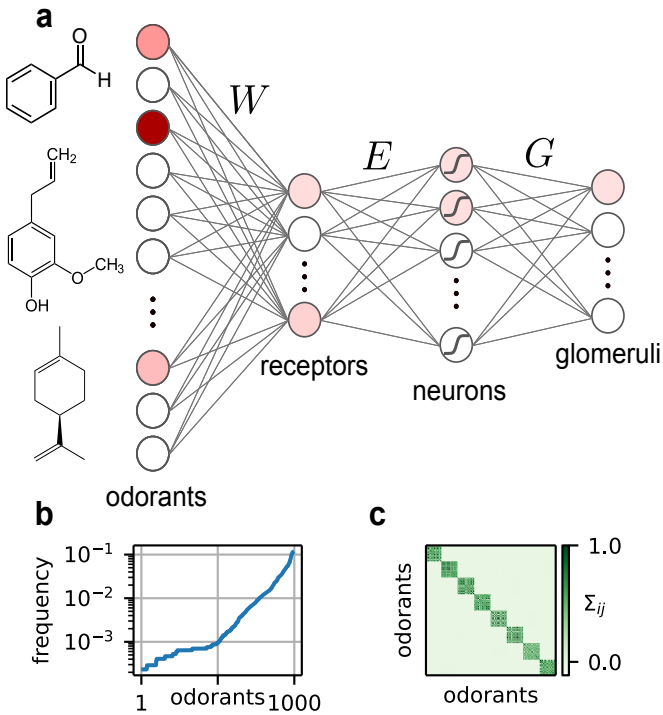

FIG. 1. Overview of the model. (a) The structure of the stimulus and the network. (b) The frequency of each odorant across samples. (c) An example covariance matrix of the binarized odorants. The number of blocks (each representing a source) is a parameter we vary—see Section III B.

glomerulus by neuron connectivity matrix G. In order to better respect the different timescales at play in the evolution and development of the circuit, we optimize one layer at a time. We find that efficient coding recovers the motifs of olfactory architecture that have now been described in several species separated by hundreds of millions of years of evolutionary time [9, 30–33]. Thus, this striking example of convergent evolution can be understood through the lens of optimality for information processing.

II. SETUP

A. Stimulus and encoding model

We model the olfactory stimulus $c \in \mathbb{R}^N$ as a sparse vector in the space of concentrations of N monomolecular odorants (Figure 1a). We first generate a binary vector $c_{bin} \in \{0,1\}^N$ using a Gaussian copula. This allows us to tune the mean and covariance of c_{bin}. Means are drawn from a Gamma distribution, so that some odorants are frequent, but most are rare (Figure 1b). We set the covariance of c_{bin} to be a block matrix, where each of the k blocks represents a mixture emitting correlated odorants (Figure 1c). A more realistic covariance matrix would have a nested, hierarchical structure rather than merely a single level of blocks [34], but it is less obvious how to parametrize such matrices—see Section III B and the Discussion.

We then assign an independent and identically distributed log-normal concentration with variance σ_c^2 to each odorant in the sample to generate $c \in \mathbb{R}^N$. A version of this sparse, log-normal model for c was used by Qin *et al.* [24], and we extend their model by incorporating varying frequencies, structured covariance, and a degree of sparsity that changes across samples. We include full details of the statistical model in the Methods.

Note that while we preserve the presence/absence statistics across samples (so that odorant α can occur more frequently than odorant β), we do not preserve odorant concentration ratios across samples (odorant α cannot occur at a typically higher concentration than odorant β). Such differences in concentration can be a critical aspect of identity coding [35]. However, this structure is liable to disruption by the turbulent transport of volatile molecules to the sensory epithelium, as well as differences in molecular diffusivity [35, 36]. Neglecting this structure greatly reduces the number of tunable parameters in our stimulus model.

Given the stimulus, the encoding model is

$$r = \varphi(EWc) + \xi = \frac{1}{1 + (\frac{1}{EWc})^n} + \xi \tag{1}$$

where $r \in \mathbb{R}^L$ is the firing rate of the neurons, $W \in \mathbb{R}_{>0}^{M \times N}$ is the receptor by odorant affinity matrix, $E \in \mathbb{R}_{\geq 0}^{L \times M}$ is the neuron by receptor expression matrix, $\varphi(x) = 1/(1 + x^{-n})$ is a Hill function with coefficient n, applied element-wise, and $\xi \sim \mathcal{N}(0, \sigma_0)$. This form of nonlinearity is consistent with measurements in the fly larva by Si *et al.* [37], in which the authors found that a Hill coefficient of $n = 1.46$ best fit their data. The coefficient is typically higher in vertebrates (on average $\mathbb{E}[n] \approx 2$), where olfactory receptors are G protein-coupled receptors rather than ligand-gated ion channels [38, 39] (but see Grosmaitre *et al.* [40] for $n \approx 1$ in vitro). We confirmed that the qualitative features of the optimal W and the optimal E did not depend on this coefficient (*SI Appendix: Fig. S1*). We discuss limitations of this model, as well as possible improvements, in the Discussion section. While both W [24, 26] and E [25] have been studied using efficient coding, most works have not considered the interplay between the two (with the exception of Lienkaemper *et al.* [41]).

Finally, the neural activity $r \in \mathbb{R}^L$ is integrated in regions of neuropil known as glomeruli $g \in \mathbb{R}^M$, with $M \ll L$. For simplicity, we match the number of glomeruli to the number of receptors, although in reality there could be more glomeruli (as in mice [42]) or fewer (as in mosquitoes [43]). Here the vector g consists of one representative dendrite per glomerulus. In paired recordings from pre- and post-synaptic neurons at the glomerulus in *Drosophila melanogaster*, Bhandawat *et al.* [44] found that the transfer function was sigmoid in shape, so we

parametrize it using the hyperbolic tangent with varying gain α:

$$g = G \tanh(\alpha r) \tag{2}$$

B. Layer-wise efficient coding

Given this model, we perform three separate optimizations. The first is over receptor sensitivities:

$$\sup_{W|E_{can}} \widehat{MI}(r, c) \tag{3}$$

By E_{can}, we indicate that we are plugging in a canonical (one neuron-one receptor) E matrix into (1) to optimize over W. Thus we are modeling the ongoing evolution of receptors in the context of the widespread one neuron-one receptor motif, rather than the joint emergence of the two phenomena. When optimizing W, we constrain its elements to be positive (see Methods for details of the optimization procedures for all three layers).

The next optimization we perform is

$$\sup_{E|W_{opt}} \widehat{MI}(r, c) \tag{4}$$

in which we first optimize W as in (3), then optimize over E. We also shuffle W to ensure that our conclusions are robust to the details of the receptor matrix. Varying the details of W and the parameters of the environment allows us to probe why canonical expression is so common, as well as exploring the phenomenon of noncanonical expression, as recently characterized in the *Aedes aegypti* mosquito [43, 45]. We constrain E to be positive and to sum row-wise to unity, since neurons must allocate a finite budget of receptor expression [43, 45–48].

Finally, we optimize the glomerular layer:

$$\sup_{G,\alpha|E_{opt},W_{shuffle}} \widehat{MI}(g, c) \tag{5}$$

where $W_{shuffle}$ is a shuffled version of W_{opt}. This is likely to be a more realistic model of true receptors than W_{opt} due to the "sloppy" tuning of real receptors (something we discuss at length in Section III B). Again we constrain G to be positive and to sum row-wise to unity, since glomeruli receive excitatory inputs from finitely many olfactory receptor neurons [8].

We emphasize that we do not jointly optimize over all three matrices, but rather treat one layer at a time, starting with W. We consider the trade-offs of this layer-wise approach, its alternatives, and its biological interpretation in the Discussion.

C. Maximizing mutual information by proxy

The primary challenge in any efficient coding analysis is the estimation of mutual information. To enable analytical progress, strong assumptions on the stimulus and encoding are required [24, 25, 41, 49–51]. Most simply, if one assumes linear processing of a Gaussian stimulus, the mutual information can be computed in terms of the log determinants of the resulting covariance matrices. However, it is challenging to analytically compute mutual information in more realistic non-Gaussian models like the one we adopt here.

Numerical estimation of mutual information in high dimensions is also fraught with difficulties [50, 52]. Here, inspired by work on deep representation learning, we adopt a conservative approach based on the principle that finding an optimal encoding model does not require an estimate of the precise value of the mutual information. Instead, it is sufficient to maximize a reliably-estimable proxy objective function [53, 54]. Committing to this approach limits our ability to compare different optimized models, but it allows for a much more robust and efficient optimization procedure.

As detailed in the Methods, we leverage a variational formulation for mutual information maximization first proposed by Nowozin *et al.* [55]. Specifically, we maximize a bound on the Jensen-Shannon divergence (JSD) between the joint and marginal distributions of the stimulus and the encoding. This was inspired by previous works showing that the JSD can be estimated more reliably than the Kullback–Leibler divergence (KLD) that defines the mutual information [53, 54]. We describe the trade-offs of this approach in the Methods. We handle constraints on W, E, and G using the framework of mirror descent; see Methods for details.

III. RESULTS

A. Receptors

We first optimized the mutual information over the sensing matrix W. Each row of this matrix represents a receptor, and each column an odorant. Thus W_{ij} is the sensitivity of the i-th receptor to the j-th odorant. It is helpful to consider that chemoreceptors are both ancient (predating neurons, for example), and fast to evolve—indeed, they are some of the fastest evolving proteins in many organisms [2, 6, 7, 56]. This may be because their conserved function is less critical for survival than that of metabolic or structural proteins. Furthermore, there is very direct evolutionary pressure to modify chemoreceptors in response to changes in the organism's chemical environment.

Thus it seems that the matrix W is relatively tunable (notwithstanding biophysical constraints—see Discussion), and so we asked: to what extent can efficient coding recover the qualitative features of receptor affinities measured in biological circuits? To guide our analysis, we turned to published data on the well-characterized olfactory system of the *Drosophila melanogaster* larva, which has just 21 olfactory receptor neurons, each identified by the expression of a unique receptor. The tractable size of

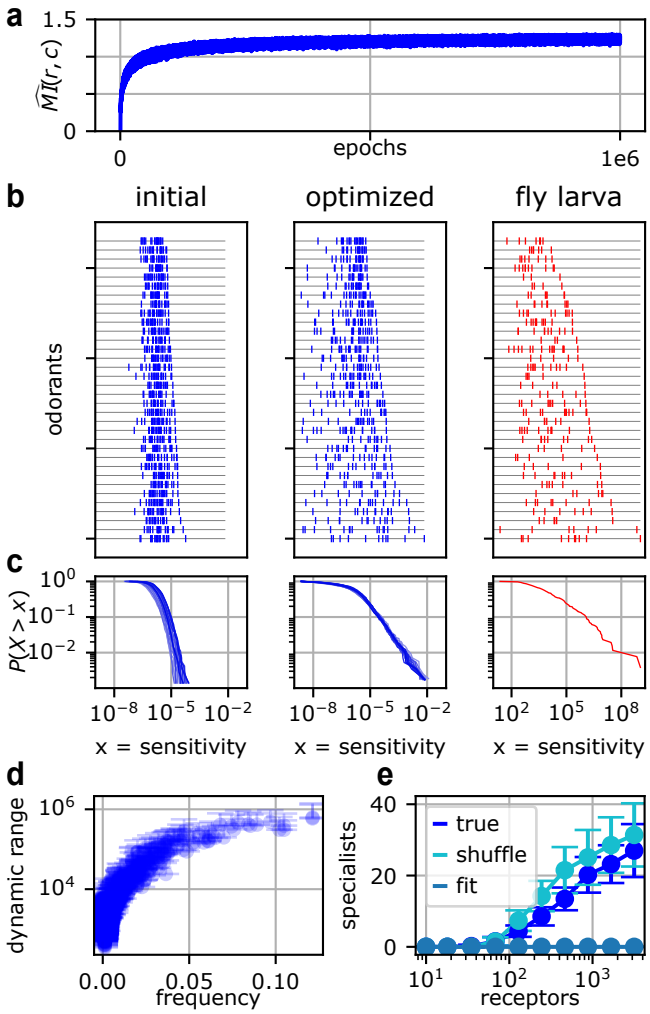

FIG. 2. Optimizing over W for parameters typical of the fly larva, given canonical expression. (a) The trajectory of mutual information over the course of the optimization. (b) Sensitivity per odorant in the optimization (blue) and fly larva (red). Each tick denotes the sensitivity of one receptor for that odorant. In both cases, there are typically many receptors per odorant, and their sensitivities span several orders of magnitude. Odorants are sorted by their maximum sensitivity within each panel. (c) Distribution of W_{ij}. For the optimizations, 20 runs with different random seeds are shown. (d) More frequent odorants are afforded greater dynamic range. Error bars denote standard deviation over 20 runs (only the upper error bars are shown). (e) For fixed $N = 1000$, the number of "specialist receptors" grows with the number of receptors. "Fit" is an analytic control obtained by fitting a log normal distribution to the optimized W matrix. Experimental data in panels (b) and (c) is from Si *et al.* [37].

this system permits comprehensive interrogation of the complete receptor repertoire, which was conducted by Si *et al.* [37]. Plugging in a canonical expression pattern E as in equation (3), we optimized mutual information between neural activity r and the stimulus c, using parameters that matched the fly larva circuit.

The resulting W_{opt} had a non-negligible degree of sparsity ($\approx 40\%$ entries $= 0$; see discussion in *SI Appendix A*), and its non-zero elements were well-fit by a log-normal distribution—see *SI Appendix: Figs. S2 and S3*. Both of these features have been derived analytically in previous theoretical work [24, 26] in the limit of vanishing neural noise. Here we confirmed that these results still approximately hold when neural noise is non-negligible. We next sought to compare our optimized W to experimental results in a more detailed and biology-focused analysis, as shown in Figure 2.

We found three qualitative matches between our optimized W and the fly larva W. First, the distribution of sensitivities across all receptors and odorants is broad, spanning approximately six orders of magnitude (see Figure 2c). This heavy-tailed distribution is the object of much analysis by Si *et al.* [37], who argue for the computational advantages of such a code. In our results, it is difficult to confirm a precise power law to the exclusion of other heavy-tailed distributions. But the qualitative agreement shown in Figure 2c may further support the optimality of this distribution for the entries of W.

Second, each odorant typically has at least several receptors tuned to it, as shown in Figure 2b. This is necessary to transmit faithful information about the stimulus, since the two orders of magnitude spanned by each receptor's Hill function activity do not cover the full range of odorant concentrations. (For Figure 2, odorant concentrations were drawn from a log-normal distribution with $\sigma_c = 3$.)

Third, some odorants are sensed by a greater number and range of receptors than others (compare the top and bottom lines of Figure 2b). Within the context of our optimization, there is an intuitive reason for this: it is the more frequent odorants that are afforded greater dynamic range, as shown in Figure 2d. While this result may hold in experimentally measured receptor sensitivities, testing it directly would require quantitative knowledge about natural olfactory landscapes, which are notoriously difficult to characterize.

More fundamentally, however, this last finding underscores one limitation of the efficient coding framework. Mutual information is a statistical measure of the dependence between two distributions and does not privilege any particular component of the stimulus for reasons of ecology or behavior. Thus, when optimized, our model circuit simply devotes greater resources to more common odorants. In natural settings, by contrast, animals may have evolved receptors to sense an infrequent but critically important odorant (such as a toxin or pheromone) with high sensitivity across a wide range of concentrations, as in Sakurai *et al.* [57]. Thus, although the trend in Figure 2d may hold on average, it is likely to admit important exceptions in biological circuits.

The latter observations consider the circuit's ability to report information about a given odorant (studying columns of W). The complementary viewpoint is to begin with a receptor and study its tuning across odorants (rows of W). For example, much of the recent experimental work on olfactory receptors has emphasized the promiscu-

ity of their binding. Rather than a "lock and key" pairing, ligands were found to fit loosely in the binding pocket of the insect receptor *Mh*OR5 [58]. This strategy is the natural result of the odorant to receptor bottleneck ($N \gg M$), and under our model, most receptors indeed exhibit broad tuning (*SI Appendix: Fig. S2*). However, this bottleneck argument does not explain why some receptors seem to have narrow tuning for one or a handful of odorants [10], especially in organisms with greater numbers of receptors (such as in mice, where $M \approx 1100$).

To explore this question, we varied the degree of the bottleneck (keeping N fixed and increasing M) and inspected the resulting optimized receptor profiles. As we increased the number of receptors, the circuit acquired an increasing number of "specialist receptors" (see Figure 2e). We formalized this using the following criterion: if the maximum sensitivity of a receptor for an odorant was at least two orders of magnitude greater than the 99th percentile sensitivity, then we counted it as a specialist for that odorant. This corresponds to testing a panel of 100 odorants, and finding that a receptor has 100x greater sensitivity for its specialized odorant than for all other odorants in the panel.

In Figure 2e, we plot two important controls. One is a shuffled W. For this control, the trend still holds, which indicates that specialist receptors are developed by pushing up the sensitivity of the specialist receptor-ligand pair, rather than pushing down the sensitivities of the specialist receptor for other ligands. Therefore, a specialist in our model is a receptor with mostly typical sensitivities, and one very high outlier.

Our other control is an analytic fit. Here we fit a log-normal distribution to the values W_{ij} for each value of M, then sample from that distribution to generate a W, and count specialists in that W. While the log-normal is heavy-tailed and fits $p(W_{ij})$ well for the bulk of the distribution (*SI Appendix: Fig. S3*), it does not generate the extreme outliers needed to produce any specialist receptors under our very stringent criterion. This demonstrates that the effect is not driven simply by drawing more samples from the same distribution.

These findings suggest that specialist receptors may "look normal" until their target ligand is found. Conversely, given knowledge about a specialist receptor, it may be worth testing other odorants against that receptor to see if they are detected at reasonable concentrations (such as in Meyerhof *et al.* [59]).

Interestingly, despite the allocation of greater dynamic range to higher frequency odorants (Figure 2d), we did not find that they were more likely to be targeted by specialized receptors (*SI Appendix: Fig. S4*). This may reflect the fact that, when specialist receptors are saturated in the presence of their target odorant, they effectively exacerbate the bottleneck faced by the rest of the circuit ($\frac{N-1}{M-1} > \frac{N}{M}$). This suggests that there may be slight pressure against developing specialist receptors for common odorants.

In this way, optimal sensing matrices W may leverage structure in the stimulus, as well as promiscuous binding, to get useful information from a specialist receptor when the target odorant is not present. Some support for this idea can be found in studies [60, 61] which showed that background odors can compromise the detection of a ligand by its cognate specialist receptor. Conversely, host plant volatiles were shown to synergistically improve pheromone detection in the silkmoth *Bombyx mori* at the receptor neuron level [62].

Finally, one important check on our numerical procedure was to confirm that the distribution of the optimal W does not depend on initialization. In Figure 2, the initial W is a scaled log-normal: $W_{init} = 1/\mathbb{E}[||c||]e^{Z}$, where $\mathbb{E}[||c||]$ is the mean magnitude of the stimulus vector (see Methods). We confirmed that scaling the log-normal to the minimum sensitivity instead, below which W_{ij} is effectively 0, did not change the results in Figure 2. (We derive this value and discuss its implications in *SI Appendix A*.)

Initialization had virtually no effect on the distribution of W_{opt} (as in panels b,c) or the relationship between frequency and dynamic range (panel d). The one instance of dependence on initialization was seen in the high M regime ($M > 500$) in panel e. With so many receptors, the bulk of W_{ij} could remain close to initialization without degrading performance. This increased the number of specialists that emerged by roughly a factor of 2 (maximum 57 rather than 27 as in Figure 2), but did not change the trend. We plot Figure 2 for this alternative initialization in the *SI Appendix: Fig. S5*.

To summarize, the W_{opt} which results from our optimization procedure shares a number of properties with the experimental W measured in the fly larva. The distribution of values $p(W_{ij})$ spans six orders of magnitude and is well fit by a log-normal density. Receptor tuning is surprisingly broad, given that we do not explicitly model the imperfect binding of real ligand-receptor pairs. Lastly, increasing M permits the emergence of specialists in circuits with larger receptor families.

B. Expression

Perhaps the most striking feature of canonical olfaction is the "one neuron-one receptor" rule. From a computational perspective, it is not immediately clear why this should be optimal, and yet it has emerged in various organisms including flies [31], mice [32], and ants [33]. Notably, each of these organisms has developed an entirely different mechanism to generate this pattern of expression.

In flies, a core set of 15-20 transcription factors and cis-regulatory elements acts combinatorially to establish the expression of a single odorant receptor (OR) in a given cell type [63]. In mice, on the other hand, the OR loci from 18 chromosomes come together in a hub that chooses a single receptor allele for expression and silences all the rest [32]. Very recent work has demonstrated how a complex gene expression program whose usage varies tightly with

position in the olfactory epithelium orchestrates this OR choice [47, 48]. In ants, many odorant receptors are transcribed, but only the most upstream gene is exported out of the nucleus [33, 64].

These and other disparate mechanisms [63] suggest that the one neuron-one receptor rule may confer a strong fitness advantage. Accordingly, when we numerically maximized our proxy mutual information with respect to E in (1) for numbers typical of the fly larva, we found that a clear pattern of canonical expression emerged, as shown in Figure 3. This was despite random initialization (Figure 3b).

Immediately, however, we are confronted with a "chicken-or-egg" problem. In (3), we optimize W given canonical E, then in (4) we optimize E with the resulting W_{opt}. This could in principle bias the optimization over E to converge on canonical expression. To account for this possibility, in our subsequent analysis we optimize E given five plausible alternative models for W. We find that each of these models favors canonical expression (see analysis below, and *SI Appendix: Fig. S11*), which indicates that the result is not merely an artifact of our layer-wise approach.

Given that the solution is robust, then why is canonical expression optimal? In the limit of low neural noise, one simple way to increase mutual information is simply to decorrelate activity, when a distribution of activity is computed over the stimulus distribution $p(c)$. This is a standard prediction of many efficient coding analyses [23, 65]. Accordingly, we find that the distribution of variances explained by the principal components of activity is flatter under canonical expression than under random expression (Figure 3d). The best way to decorrelate activity in this way is to place neurons at the corners of the simplex in gene expression space, since total receptor expression per neuron must sum to unity. This corresponds precisely to the one neuron-one receptor rule. (See Lienkaemper *et al.* [41] for a related analysis in the case of one neuron.)

We next sought to understand if the same result would hold in a larger system. We chose a scale comparable to the adult fly, with $M = 60$ receptors and $L = 1260$ olfactory neurons. This is a natural focus, since olfactory processing in the fly is well characterized (see for example Hallem and Carlson [11]), and these numbers permit reasonable numerics. We first optimized W, and found a qualitative match to the measurements in Hallem and Carlson [11], with most receptors exhibiting broad tuning and a handful of receptors exhibiting narrow tuning (*SI Appendix: Fig. S6*).

Surprisingly, when we optimized over E using this W matrix, a sloppier expression matrix emerged, in which a handful of receptors were typically coexpressed per neuron, as shown in Figure 4a. Interestingly, recent work in the *Aedes aegypti* mosquito has discovered this qualitative pattern of expression [43, 45]. Reanalysis of existing data by Adavi *et al.* [45] further characterized occasional exceptions to the one neuron-one receptor rule in *Drosophila*. The extent of this coexpression is controversial [63, 66]

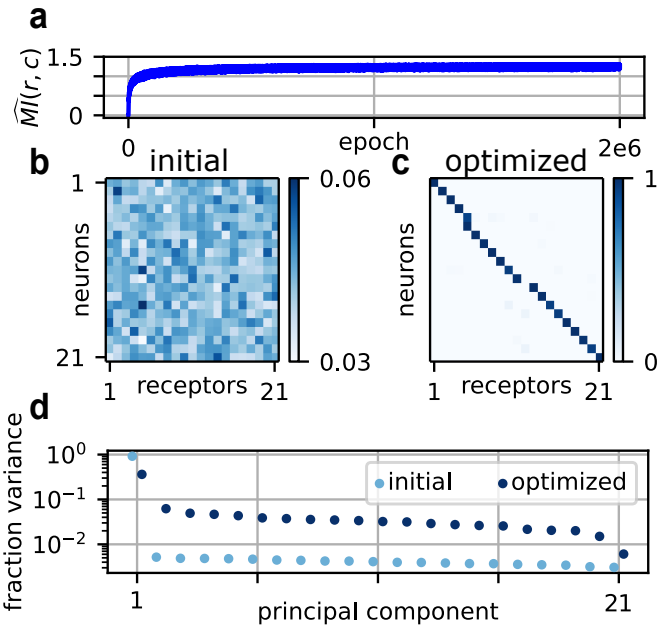

FIG. 3. Optimization over E for numbers typical of the fly larva. (a) The trajectory of mutual information over the course of the optimization. (b) The initial expression matrix. (c) The optimized expression matrix. The optimization is over E in (1), plugging in the optimal W from Figure 2. (d) The fraction of variance in neural activity r explained by each of the top principal components, given the initial expression from panel (b) and the optimized expression from panel (c). Statistics are computed over 1000 samples.

but some unambiguous cases have been known for many years [67, 68].

As we interpret this result, it is is important to note that, in both *Aedes aegypti* and *Drosophila*, the majority of coexpressed chemoreceptors are disproportionately close to each other in genomic and phylogenetic space [45]. Adavi *et al.* [45] term this "coexpression by descent." In only a minority of cases did the authors find that coexpressed receptors were far from each other ("coexpression by co-option.") This suggests that much of the phenomenon may be a by-product of shared gene regulation. Since families of chemoreceptors frequently expand via tandem duplications (the "birth and death model" of gene evolution) [7, 69], there are many examples of neighboring ORs with similar sequences whose coexpression might be driven by cis-regulatory elements [63].

These molecular mechanisms likely have little to do with optimal information processing. However, the fact remains that different insects (like *Drosophila* and *Aedes aegypti*) exhibit strikingly different levels of coexpression. Our results suggest that similarity between receptor affinity profiles is one factor in understanding the functional significance of these differences, as shown in Figure 4c and d. Since phylogenetically related receptors typically have similar affinity profiles [70, 71], this is consistent with the analysis of Adavi *et al.* [45]. Under this reasoning, coexpression is primarily driven by shared regulatory

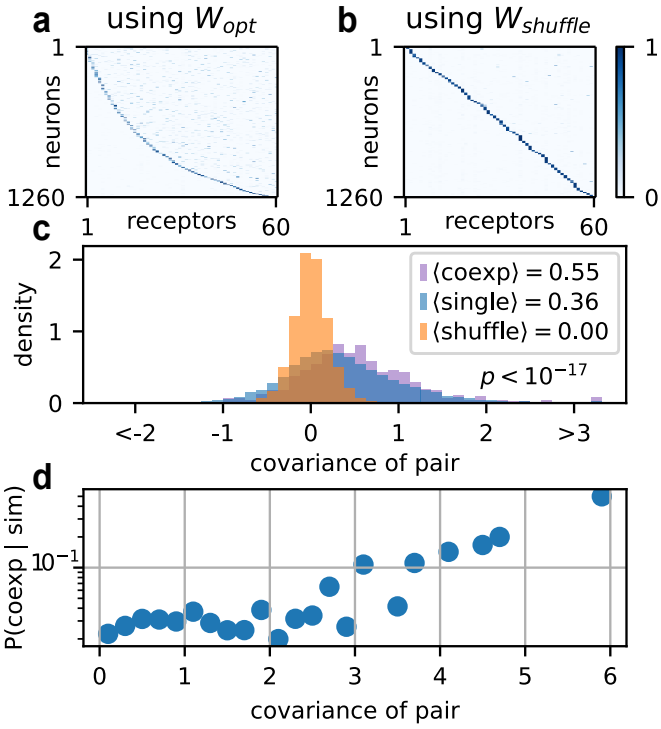

FIG. 4. Optimization over E for numbers typical of adult *Drosophila*. (a) The optimized expression E obtained from plugging in the optimized sensing matrix W_{opt}. (b) The optimized E obtained from plugging in a shuffled version of W_{opt}. (c) Covariance between W_{opt} receptor pairs which are coexpressed in at least three neurons ("robust coexpression", purple), W_{opt} receptor pairs which are singly expressed (blue), and $W_{shuffle}$ receptor pairs (orange). Covariances are computed on log sensitivities. (d) Probability of robust coexpression given binned covariances. Data in (c) and (d) are across 21 runs with varying environmental parameters σ_c and Σ_c, given by the bottom three rows of the phase diagram in Figure 5. The p-value for the difference in distribution between robustly coexpressed pairs (purple, $n_c = 963$) and singly expressed pairs (blue, $n_s = 20978$) is obtained using the Mann-Whitney U test.

machinery, but once it emerges, its fitness effect depends on the similarity of the coexpressed receptors, with coexpression of more similar receptors being less detrimental.

Given these subtleties, why did we obtain such a clean canonical expression profile in our analysis of the larval fly (Figure 3)? The answer may be in the number of receptors M. Sweeping over this parameter, we found that receptor matrices with small M were much closer to full rank (*SI Appendix: Fig. S7*), indicating that their receptors were more differentiated. It may be possible that organisms with only a handful of evolutionarily mature receptors cannot afford similar affinity profiles. Conversely, circuits with many receptors are "over-parameterized" and can find many solutions to the problem, coexpressing without penalty.

To ensure that these results did not depend on specific parameter values, we next swept over the two important parameters of our odorant model: the structure of the covariance matrix Σ_c and the log-normal noise parameter σ_c. We parameterized Σ_c by giving it a block structure and varying the number of blocks (see Figure 1 and the Methods). Ecologically, each block of correlated odorants corresponds to a source in the organism's environment. Precise quantification of the fly's natural olfactory environment is elusive (see Yang *et al.* [72] and Zhou *et al.* [34] for recent efforts in this direction), but it is plausibly contained in the large space we explore.

One trend in the resulting phase diagram (Figure 5a) is that limited noncanonical olfaction is favored for increasing environmental noise σ_c. This agrees with very recent work by Lienkaemper *et al.* [41], who approach the problem using a linear-Gaussian theory. In further accord with their analysis, we find that having fewer sources, each emitting more correlated odorants, supports more noncanonical olfaction, at least for low levels of environmental noise σ_c (see bottom two rows of Figure 5a). The magnitudes of these effects, however, are small (Figure 5c, d). Note that our model is not directly comparable with theirs, due to differences in the formulation of stimulus versus noise (see Discussion).

Since our primary focus is on understanding why canonical olfaction is so common, in Figure 5 we have initialized expression matrices to be noncanonical: each neuron expresses 3-7 receptors at roughly equal levels, for a "canonical score" of 0.74 (see Figure 5 caption for definition). This was done in order to determine the degree to which canonical expression is strongly favored compared to receptor coexpression. It is clear that most of the optimizations in Figure 5a barely move, indicating that this noncanonical pattern at initialization is already locally optimal. However, when we initialize expression to be canonical, we also see virtually no change (*SI Appendix: Fig. S8*). Therefore, neither canonical nor noncanonical expression is strongly favored, regardless of environmental structure. Mathematically, both expression patterns represent solutions to the optimization.

To better understand this puzzling result, we inspected the optimized matrices W_{opt}. We found a distinctive low rank structure in W_{opt} across environmental parameters that was especially pronounced for higher values of σ_c (*SI Appendix: Fig. S9*). This is driven by our frequency model since, as indicated in Figure 1b, only a small subset of odorants occur with high frequencies, while the bulk occur at low frequencies. Flattening the frequency distribution abolished the low rank structure (*SI Appendix: Fig. S10*). This degeneracy in W_{opt} permits either canonical or noncanonical olfaction, since the precise details of expression do not matter when W is low rank. Importantly, this finding is likely to depend on the details of our biophysical model (1), especially the placement of the nonlinearity (see Discussion).

Our modeling is therefore consistent with the hypothesis that if organisms need to sense only relatively few odorants in their environment, then their receptors may share much of their tuning for those odorants. This in

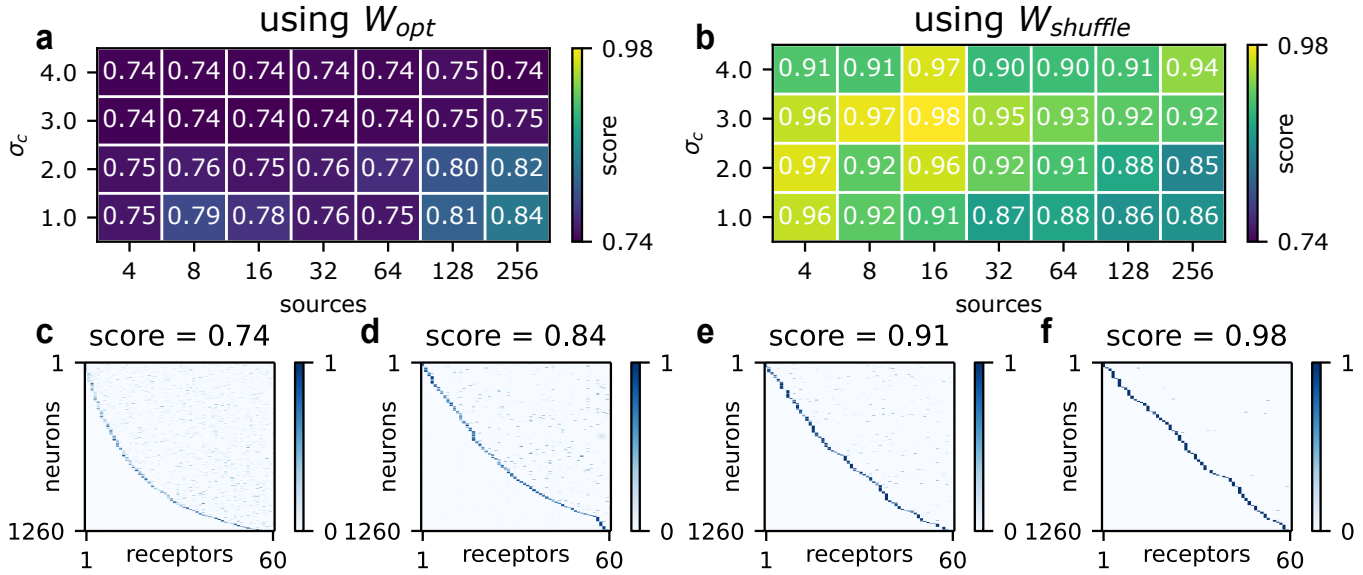

FIG. 5. The degree of canonical expression in optimized E matrices while varying environmental parameters σ_c and Σ_c. (a) The phase diagram using optimized W. (b) The phase diagram using shuffled W. (c) through (f): example expression matrices taken from the above phase diagrams. The score is $1 - \mathbb{E}[H(E_i)]$, where $H(E_i)$ measures the entropy of the i-th row of E, when the expression is viewed as a probability distribution over receptors, and the expectation is taken over the rows of E.

turn permits both canonical and noncanonical expression without strong pressure in either direction.

While potentially informative, this result also reflects a limitation of our mutual information objective. In our results, receptor tuning is almost uniformly high for the highest frequency odorants (*SI Appendix: Fig. S6*), whereas in reality, organisms may need to detect common and rare odorants with equal sensitivity. Additionally, biophysical constraints limit the number of ligands a receptor can bind. Indeed, Si *et al.* [37] found that covariance in olfactory sensory neuron activity is primarily driven by the geometric structure of odorants. This geometry constitutes a significant constraint on receptor affinities that is not accounted for in our model, and thus real receptors are very likely to be sloppier and less finely tuned than our W_{opt} matrices.

In light of these limitations, we next studied expression using shuffled versions $W_{shuffle}$ of the optimized W matrices, as in Figure 4b. Shuffling allowed us to preserve the scale and distribution of W_{opt}, while breaking the low rank structure. It also accounts, albeit imperfectly, for the suboptimal tuning of real receptors. Using $W_{shuffle}$, we found that canonical expression emerged robustly across almost all combinations of environmental parameters (as shown in Figure 5b), despite noncanonical initialization. Again, we found only a weak dependence on environmental parameters, and the trends were reversed compared to Figure 5a. We tested two other unstructured models for W: shuffling W_{opt} within rows (in order to preserve mean receptor tuning), and fitting a log normal distribution to W. Both of these favored canonical expression when E was then optimized (see *SI Appendix: Fig. S11*).

Real W matrices are most likely best modeled by something between W_{opt} and $W_{shuffle}$. They have some correlation structure, due to the development of new odorant receptors from existing ones, but are unlikely to be extremely low rank. We therefore implemented two more models for W: a log normal analytic W with block covariance structure, to model families of related receptors, and a log normal W with Toeplitz covariance structure, to model a range of similarities across receptors. We initialized expression to simulate coexpression by descent; that is, coexpression of correlated receptors. However, we found that canonical expression was still largely favored regardless of receptor details (see *SI Appendix: Fig. S11*).

Finally, we sought to understand how the level of neural noise σ_0 affected these results. We swept across σ_0 for a representative pair of environmental parameters (blocks $= 64$, $\sigma_c = 2$) (*SI Appendix: Fig. S12*). We found that at extremely low noise levels ($\sigma_0 = 0.01$), both solutions were equally favored, but for more realistic levels ($\sigma_0 \in [0.1, 1.0]$), canonical olfaction was favored as shown above. All of the above results were run for $\sigma_0 = 0.1$, a value we chose because it sets the noise to be on the order of the mean activity (see *SI Appendix: Fig. S13* and Glomerular convergence). Increasing σ_0 beyond this point further favored canonical olfaction. This trend is consistent with the theory of Lienkaemper *et al.* [41].

Taken together, these results suggest that receptor co-tuning, rather than environmental statistics, is decisive in determining optimal expression. In our model, only extremely low-rank W matrices permit significant levels of noncanonical olfaction. Biologically, such a set of receptors would have highly correlated tuning across many odorants. A number of more plausible models for W support largely canonical olfaction for realistic levels of

neural and environmental noise. This may explain why the one neuron-one receptor rule has emerged in such distantly related organisms, despite the fact that these organisms sense different chemical environments with receptor families that are accordingly divergent.

C. Glomerular convergence

Olfactory sensory neurons converge onto regions of neuropil known as glomeruli [73]. In the canonical model, only neurons expressing the same receptor converge onto a given glomerulus [74]. This "glomerular convergence" confers an intuitive advantage: given canonical expression, the signals from each receptor can be averaged across neurons without mixing across receptors. We plugged in a shuffled version of W_{opt}, the corresponding E_{opt}, and optimized our proxy for mutual information over G and α as in (6). We confirmed that glomerular convergence is recapitulated in our model across a wide range of environmental parameters (see Figure 6a-c, and *SI Appendix: Fig. S14* for sweep).

But glomeruli do not just serve to denoise olfactory neuron responses. Careful recording of both the pre-synaptic olfactory receptor neuron (ORN) and the post-synaptic projection neuron (PN) in *Drosophila* has characterized other aspects of the transformation [44]. Chief among these is "histogram equalization," in which low ORN activity is amplified but higher activity is not. Since ORN activity is clustered near 0, this induces a more balanced distribution of activity in the post-synaptic neuron, which in turn enhances information transmission. This is a classical prediction from the early days of efficient coding [75].

We first confirmed that our activity clustered near 0 as in Bhandawat *et al.* [44] (see Figure 6d, bottom axis, and *SI Appendix: Fig. S13* for sweep). This is likely necessary due to the log-normally distributed stimulus; if the optimal W is to account for the rare, highest concentration odorants, then the bulk of stimuli will generate responses closer to baseline. Accordingly, we found that as we increased σ_c, the distribution of activity was more sharply peaked around 0 (*SI Appendix: Fig. S13*).

Optimizing over the gain parameter α in (5) resulted in a qualitative match to data from adult *Drosophila*, in which the post-synaptic activity was more evenly distributed across the dynamic range of the neuron (Figure 6d). The resulting histograms were not perfectly equalized, but this may reflect the fact that increasing the entropy of the representation $H(g)$ addresses just one term out of two in the expression for mutual information:

$$MI(g, c) = H(g) - H(g|c) \tag{6}$$

In sum, our results are consistent with the idea that glomerular convergence allows for denoising and histogram equalization of olfactory receptor neuron responses when expression is canonical. As we previously observed that canonical expression is largely favored across a range of environmental parameters, we did not investigate what convergence patterns emerge when multiple receptors are expressed in each sensory neuron. However, as long as neural noise is a limiting factor, convergence is the likely solution for denoising.

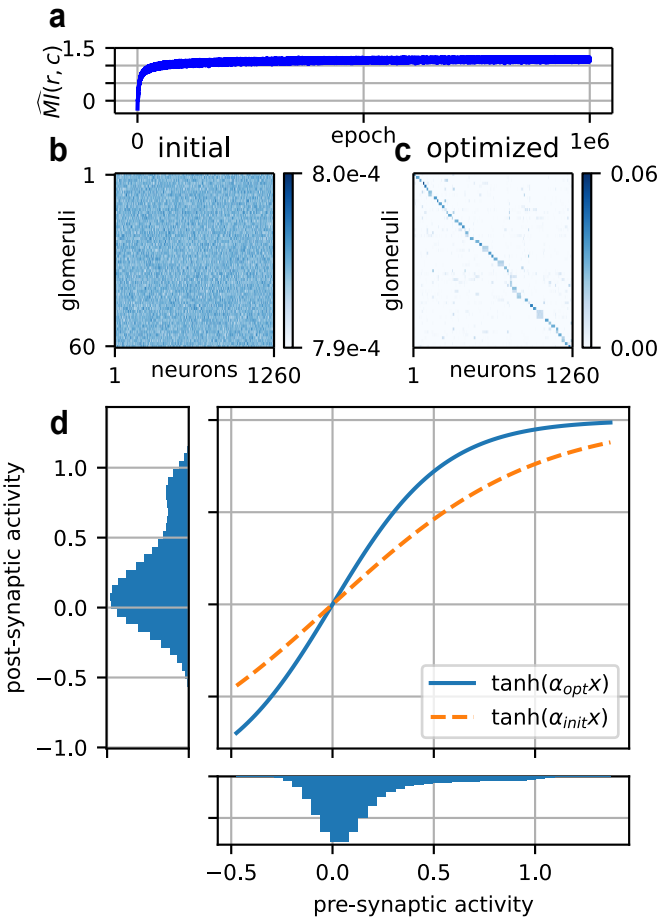

FIG. 6. An example optimization over the glomerular layer for 64 sources and $\sigma_c = 3$. (a) The course of the mutual information estimate over the optimization. (b) The initial random connectivity. (c) The optimized connectivity. Neurons are sorted by their highest expressed receptor. (d) The transformation at the glomerular layer. A higher α_{opt} serves to flatten the distribution of activity at the glomerulus.

IV. DISCUSSION

In this work, we set out to understand why odor coding exhibits deep similarities across vertebrates and invertebrates. Using efficient coding, we recovered three motifs of a widely shared olfactory logic: broad receptors, single receptor expression, and glomerular convergence.

A. Receptors

The non-zero sparsity of the optimal W and the broad distribution of its values have been derived in previous theoretical works that approximate the mutual information $MI(r, c)$ as the entropy $H(r)$, an approximation which becomes exact in the limit of no neural noise [24, 26]. Our analysis first recovers these results in the regime where neural noise is non-negligible, then builds on them by studying more granular features of W_{opt} such as receptor tuning, odorant sensitivity across receptors, and specialists. These details allow us to make new predictions for experiment, particularly regarding specialist receptors.

In our model, receptor tuning is much more promiscuous (most receptors respond to most odorants) than strictly necessary to map all N odorants in an approximate labeled-line scheme. This correspondence was surprising, since broadly tuned receptors in biology might simply be a byproduct of ligand-receptor biophysics. If this were the case, then such receptors would not emerge in our unconstrained model when we initialize to zero.

Our results suggest instead that broad tuning confers an information processing advantage. The rich theory of compressed sensing may shed light on this advantage, but it is difficult to apply classical results from this field directly to the problem at hand, since a linear measurement model is usually assumed [76]. Recent progress in nonlinear compressed sensing may enable theoretical understanding of the bottleneck problem in olfaction, but much remains unknown [24, 27, 28, 77, 78].

These broadly tuned receptors enable a combinatorial coding scheme in which many more odorants can be represented than there are receptors [79]. To the extent that specialists do emerge in our optimizations, they have increased their sensitivity for the target ligand rather than decreasing their sensitivity for off-target ligands, and are thus still available for sensing other odorants.

One prediction of our analysis is therefore that, in the absence of the target ligand, specialist receptors may be co-opted for the sensing of other relevant molecules at realistic concentrations. This idea is indirectly supported by their diminished performance in the presence of background odorants [60, 61], which indicates that specialist receptors preserve some tuning for off-target ligands. Our results suggest that this phenomenon may be a feature, not a bug, when the specialized ligand is not present. This could be further explored simply by presenting specialist receptors with a typical panel of common odorants, ideally ones not naturally co-occuring with the target ligand.

The question of the relevant range of concentrations for odorant stimuli is critical. Wachowiak *et al.* [80] have recently argued that the concentrations used in typical experimental studies may exceed those encountered in the environment by several orders of magnitude. If this is true, than much of the broad tuning of receptors as currently characterized by experiment would be functionally inaccessible to the organism. On the other hand, careful experimental work [81, 82] (see Vickers [83] for a review), supported by models of turbulent transport [35, 36], suggests that odorant concentrations fluctuate wildly at the sensory epithelium. Animals may therefore sense odorants quickly during high-concentration bursts, rather than averaging over long timescales. This is supported by behavioral experiments in which flies can execute odor-guided behavior within 100ms [84], as well as the discovery that rodents can identify odors within a similar temporal window [12–14]. Note, however, that these turbulence arguments do not apply to the *Drosophila* larva, since the fly spends most of its larval stage inside a food source [85].

Even granting that experimental concentrations are unnaturally high, the problem is mitigated by the consideration that all studies necessarily use a limited panel of odorants. As more odorants are tested, more extremely high sensitivities will be filled in for each receptor, and the current picture of combinatorial coding might survive, even if all sensitivities are shifted upwards. Our model, which only considers the relative values of sensitivity and concentration, can say little to resolve this question, which amounts to fixing the mean of the stimulus distribution $p(c)$. But our results do suggest that the variance of $p(c)$ is likely to be high, since this variance is what drives the spread in the optimized sensitivities, and these in turn are a close match to experimental data.

There are several limitations of our receptor-level analysis. First, we have only accounted for excitation of sensory neurons by odorants. This was primarily because inhibition was not detected by calcium imaging in Si *et al.* [37], which was our primary experimental comparison. In reality, however, antagonistic interactions play a crucial role in mammalian odor coding [38, 86, 87]. Inhibitory responses also emerged in previous theoretical work which optimized over W given a non-zero baseline activity in neurons [24]. Mechanistically, a minimal two step model in which the odorant first binds the receptor with some affinity and then activates it with a different, potentially uncoupled affinity, explained a number of classic observations in psychophysical experiments such as synergy and overshadowing between pairs of odorants [88]. In future work, we hope to optimize over the parameters of this more realistic biophysical model and revisit the above analyses (initial attempts to do so proved numerically unstable within our framework).

Second, lurking in any biophysical model is the question of where to place the nonlinearity. In (1), we have placed the expression matrix E inside of the Hill function. Therefore we are using the Hill function as an empirical fit to ORN activity in the presence of multiple odorants, not as a mechanistic model for coöperativity in odorant receptor signaling. A plausible extension of our model would place E outside of the nonlinearity, then include another nonlinearity for neural activity. To our knowledge, this issue has not been considered in previous works on efficient coding in olfaction, either because they assumed linear processing of the stimulus [25, 41] or canonical olfaction [24, 26] (i.e., a one-hot E whose placement therefore has

no effect).

Third, and most fundamentally, biological evolution is a much more constrained and local process than our numerical optimization procedure. For example, real receptors likely cannot change their affinity for one odorant without changing their affinities for others, especially since odorant geometry is a primary determinant of receptor specificity (see Si *et al.* [37] and del Mármol *et al.* [58]). Furthermore, organisms develop new receptors from existing ones through a birth-and-death model [7, 69], rather than optimizing a set of randomly initialized receptors as we have done. Any attempt to encode these evolutionary dynamics in our optimization would be computationally challenging, but a principled approach could be fruitful.

With these constraints in mind, it was perhaps surprising that efficient coding could predict any aspects of receptor tuning. Our optimization process becomes more realistic as we compare to more evolutionarily mature receptors that diverged from each other long ago, so the correspondence we found may indicate that the receptors of the larval *Drosophila* are reasonably mature and tunable.

B. Expression

One key takeaway from our work is that something very close to the one neuron-one receptor rule appears optimal across a broad range of environmental parameters. In any normative model, a chief concern is the extent to which such results depend on specific details of the setup. This is why we have swept over a broad range of parameter values for Σ_c, σ_c, and σ_0, and inserted multiple models for W. The fairly general emergence of canonical expression in these analyses is likely due in part to the decorrelation of activity (see Figure 3d) that occurs when neurons "spread out" in gene expression space by choosing one receptor. A complete answer, however, would require a more developed theory for the nonlinear setting, which we leave for future work.

Critically, we do not find evidence that receptor coexpression should be finely tuned to either the statistics of receptors or the environment. On the contrary, we find that one neuron-one receptor is basically optimal, and that exceptions to this rule should have varying effects depending on receptor similarity. Stronger claims would need to be weighed against the recent work suggesting that many coexpressed pairs are recent duplicates whose coexpression is driven simply by shared regulatory factors [45, 63]. Ramdya and Benton [89] raise the possibility that such instances may reflect a transient evolutionary state, in which recent duplicates are coexpressed only until regulatory machinery has "caught up" to the duplication event by creating new cell types.

In this respect, our findings differ from a very recent theoretical analysis by Lienkaemper *et al.* [41], who characterized the dependence of optimal expression on environmental noise and the "signal correlation" $W\Sigma_c W^\top$ using a linear-Gaussian theory. Discrepancies may be due to the nonlinearity we include in the receptor activity or the log-normality of our odor model. It is also important to note that our formulation differs slightly from theirs: whereas they add noise to a Gaussian stimulus $c = c' + \xi$ and compute $MI(r, c')$, we compute $MI(r, c)$ where c is log normal with variance σ_c^2. Thus, any trends that depend on environmental noise cannot be directly compared. On the other hand, in accordance with their theory, we do find that increasing neural noise favors canonical olfaction (see *SI Appendix: Fig. S12*). Another important difference is that Lienkaemper *et al.* [41] focus on the setting where the number of neurons L is less than the the number of receptors M, forcing neurons into coexpression if no receptors are to be neglected. In our setting, $L > M$, which permits neurons to spread out their activity by singly expressing without losing information.

C. Glomeruli

At the glomerular level, the optimized circuit averages across neurons expressing the same receptor, and uses the transfer function to spread out post-synaptic activity. This has already been understood in terms of efficient coding (see Bhandawat *et al.* [44]), and it is the most intuitive of the three layers. In this way, it serves as a useful sanity check of our computational approach. Because we found canonical expression to be largely robust, we did not attempt to make predictions about glomerular convergence in the case of noncanonical expression.

D. Layer-wise efficient coding

The key structural choice we made in this work was to adopt a layer-wise optimization approach: we first optimized receptor affinities W assuming canonical expression E, then optimized E given fixed W, and finally optimized glomerular connectivity G given fixed E and W. Layer-wise optimization is both computationally convenient, and—as we will argue below—biologically interpretable.

The risk of such a procedure is that the results could depend on the order of the optimizations. As mentioned before, the analysis of canonical expression represents a chicken-or-egg problem. If we optimize W given canonical E, then optimize E given the resulting W, it might be expected that canonical expression would emerge, if this structure is somehow encoded in the optimal W. Such a dependence would severely compromise the generality of our results. We therefore sought to check that the finding of generally-canonical expression was robust to the details of the optimization procedure by optimizing expression for different plausible models of the affinity matrix. For these alternative models, optimization resulted in strictly greater levels of canonical expression than the optimized W. This suggests that the key takeaway—canonical expression is typically optimal—is not an artifact of our

optimization procedure.

From a biological perspective, layer-wise optimization may be a reasonable model for the evolution of the convergent motifs we consider. Organisms are not presented with the opportunity to optimize an entire sensory circuit *ab initio*. Instead, W, E, and G are tuned by different mechanisms, on different timescales, and in different contexts. We therefore had to approximate the constrained setting in which each motif evolved given limited experimental knowledge.

Chemoreceptor families are evolutionarily ancient and continually evolving [2, 6, 7, 30, 56], so we chose to optimize over W first. Since approximate canonical expression is widespread, we plugged in canonical E for this optimization. This also permitted direct comparison to the experimental results presented by Si *et al.* [37]. Next, it seems reasonable to ask which gene expression programs are optimal given a broad, mature set of receptors. In the mouse, for example, the relative abundance of olfactory receptors can be adapted over the timescale of just a few hours [46, 90], although this process does not generate coordinated non-canonical expression. This framing also addresses the emergence of different expression programs across organisms which each possess mature receptor families. Hence we optimized over E given fixed W.

Glomerular convergence, which constitutes an exquisite example of wiring specificity, emerges within the lifespan of the organism by pruning during development [91, 92]. Since examples of highly adaptive connectivity abound in organisms that learn, we chose to study glomerular convergence in the setting of a mature receptor array and largely canonical expression. Hence we optimized over G given fixed, optimized W and E.

These conceptual distinctions are naturally much cleaner than the underlying biology. For example, receptors continue to evolve after a program of noncanonical expression is established [89]. However, our breakdown enables a tractable model and represents a best guess at the constraints which govern the circuit's evolution. It is of course possible that layer-wise optimization could yield sensitivities and expression patterns that do not lend themselves to glomerular convergence [93]. However, given access to at least as many neurons as receptor types and canonical expression, the simple intuition that glomerular convergence allows denoising suggests that not much is being lost.

One alternative to our approach is an unconstrained end-to-end optimization of the entire circuit. Wang *et al.* [94] recovered largely canonical expression and glomerular convergence using unconstrained optimization on a match-to-prototype classification task. This is qualitatively consistent with our results. Their work, however, differs in assuming a particular classification task, and starts with a very different assumption on the stimulus: they directly model W by assuming that the activity of each receptor is independent and uniformly distributed. This differs substantially from the statistics resulting from our optimized sensitivity matrices, and from those measured by Si *et al.* [37]. This kind of unconstrained optimization can shed light on the computation instantiated by the circuit, but cannot capture the path-dependent quirks of biological evolution [95].

Another alternative would be to formalize our context-specific model as a joint dynamical system whose dynamics are decomposed across different timescales. Most simply, this would lead to a setting where we optimize W given the expression E that is optimized for each sensitivity matrix, *i.e.*, to study the nested optimization $\sup_W\{\sup_E \widehat{MI}(r,c)\}$. This however could be too flexible—it seems that many organisms have "committed" to canonical expression, for example, and do not have the ability to smoothly modulate receptor co-expression as receptors evolve. Correspondingly, there may be some biological merit to freezing E while W is optimized.

Fleshing this out completely would require organism-specific knowledge of when and how in evolutionary history each motif emerged. In the absence of such knowledge, we approximate the distinction between the three layers in a way that seems biologically plausible, and treat each one as a separate problem. This piecewise approach limits the normative mathematical claims we can make about the circuit, but it is a closer match to evolutionary dynamics.

E. Dynamics

Our work does not address the dynamics of the odor landscape and its neural representation in the olfactory bulb. As mentioned above, turbulent transport of air to the sensory epithelium induces high-frequency fluctuations in concentration that are richly informative about identity and location [14, 35]. Here, we have ignored these dynamics in order to focus on a statistical problem of compression: namely, the representation of a high-dimensional chemical space by a finite set of receptors. The temporal challenge of encoding rapidly fluctuating odor signals is equally daunting. Recent work has begun to bridge these statistical and dynamical pictures of olfactory sensing [18, 78, 96–101], but many questions remain regarding how to efficiently encode the temporal statistics of the olfactory world. In this vein, it could be interesting to consider minimal extensions of our model that include temporal filtering [51, 102].

F. Conclusion

We reiterate in closing that mutual information is only a rough proxy for performance in olfactory processing. For example, organisms may privilege the sensing of rare but critical odorants, even at the expense of more common ones—this would not be captured by our statistical model [57, 61, 103]. Therefore some of our results are unlikely to transfer perfectly to biological circuits. That said, the great variety of olfactory tasks faced by the typical

organism suggests mutual information as a reasonable starting point [23, 24, 103].

The correspondence between our results and experimental measurements suggests that these motifs may constitute uniquely accessible solutions to the problem of processing olfactory stimuli. As chemosensation is further explored in non-model organisms [3, 104], the extent to which canonical principles hold will be an interesting point of focus. While receptors always reflect the organism's chemical environment, the other layers we consider (receptor expression and the downstream routing of receptor activity) are not obviously bound to any stereotypy. Especially interesting will be cases where departures from canonical olfaction are unambiguous, tunable, and clearly advantageous. Since our simplified modeling suggests that canonical olfaction is optimal, such departures should serve as a flag that something interesting is afoot.

V. METHODS

A. Generating odor mixtures with tunable statistics

We used a Gaussian copula to generate the binary mixture data $c_{bin} \in \{0,1\}^N$ with tunable mean and covariance. This works as follows.

1. Draw $Z \sim \mathcal{N}(0, \Sigma)$ where $\Sigma \in \mathbb{R}^{N \times N}$.
2. Set thresholds $t = \text{ppf}(1-\mu)$ where μ is the desired mean vector and ppf is the inverse CDF of the standard normal distribution.
3. Set $c_{bin} = \mathbb{1}_{Z>t}$.

Given $\Sigma_{ii} = 1$, then c_{bin} has $\mathbb{E}[c_{bin}] = \mu$. In general, c_{bin} is not guaranteed to have covariance equal to Σ, but the pairwise covariances $\mathbb{E}[c_{bin_i} c_{bin_j}]$ are monotonic, nonlinear functions of Σ_{ij}.

To model a range of frequencies, with most occurring rarely and some occurring frequently, we drew μ itself from a Gamma(α, λ) distribution. We parametrized $\Sigma = \Sigma_k$ using a block covariance structure with k blocks, as in Figure 1. Each block represented a mixture in the environment emitting correlated odorants. As we varied k below $k = 32$ (down to $k = 4$), it was necessary to manually tune α and λ in order to preserve an approximately fixed odorant frequency distribution while simultaneously achieving the desired correlation structure in the outputs. This is because not all combinations of mean and covariance can be satisfied in binary data.

For low k, we also used a "thinning" step in which, after generating the samples, c_{bin_i} was set to 0 with probability p_k. This *ad hoc* intervention was necessary to stabilize the sample-generating algorithm as $k \to 4$, ensuring that $\mathbb{E}[c_{bin}]$ was approximately constant across different values of k while achieving the desired correlation structure. Thankfully, only minor adjustments to parameters were needed, and exact settings for each k are included as parameter files in the codebase so that data can be generated anew as needed. We include plots of the covariance matrix and frequency distribution for each k in *SI Appendix: Fig. S15*.

The resulting data c_{bin} represented binarized odorant mixture samples, with typically 10 (range $1-100$) odorants per sample. We then assigned an *i.i.d.* log-normal concentration with variance σ_c^2 to each odorant that was present in a given sample. This resulted in an odorant vector $c \in \mathbb{R}^N$ which we passed into our model.

To generate samples with a flat frequency distribution (as in *SI Appendix: Fig. S10*), we repeated the above procedure, but using a constant $\mu = \mu_0$.

B. Comparison to other models

The stimulus $c \in \mathbb{R}^N$ is sensed by the neuron through a panel of m receptors. Each receptor has a vector of affinities for each odorant, which is represented by a receptor-by-odorant sensing matrix W. Our goal is to describe neural activity as a function of odorant concentration. This obviously involves W and c, but what form should it take?

One approach is to assume that each neuron expresses just one receptor, and that there is one neuron per receptor type. We will call this the case of single expression, and denote it r_{single}. If the firing rate is simply a linear sum of the (Gaussian) stimulus components

$$r_{single} = Wc + \xi$$

where $\xi \sim \mathcal{N}(0, \sigma_0)$, then $MI(r_{single}, c) = H(r_{single}) - H(r_{single} \,|\, c)$ can be computed analytically.

However, experimental [37, 38, 86, 88] and theoretical [24, 88] works have shown that saturation and other nonlinear effects play an important role in odor coding, especially when the stimulus varies over orders of magnitude. Thus a more realistic activity model is $r_{single} = \varphi(Wc) + \xi$, where φ is a Hill function applied element-wise:

$$r_{single} = \varphi(Wc) = \frac{1}{1 + (Wc)^{-n}} + \xi \tag{7}$$

We choose a Hill function because experimental work has shown, in the case of monomolecular odorants, that neural activity takes this form as a function of odorant concentration [37], with $n = 1.46$. See Setup in the main text for more details, and *SI Appendix: Fig. S1* for a sweep over Hill coefficients. Note that φ is sigmoid when c varies over a log scale.

By r_{single} in (7), we emphasize that this model is making a strong assumption on expression: just one receptor per neuron. Several prior works on efficient coding in olfaction have made this assumption, and within this setup, one can focus on the optimal properties of the sensing matrix W (as in [26], [24]) or how neurons should be allocated to each receptor [25]. This assumption is natural, since many organisms exhibit this organization. But here we wanted to target this for study as well, asking how much of canonical olfaction can be recapitulated by efficient coding theory across both layers W and E. Thus in our model, we include a neuron-by-receptor expression matrix E that allows each neuron to express a variable amount of each receptor. This yields the model (1):

$$r = \varphi(EWc) + \xi = \frac{1}{1 + (1/EWc)^n} + \xi \tag{8}$$

Here $\xi \sim \mathcal{N}(0, \sigma_0)$ is added to model the effect of neural noise. While Poisson noise would more closely model variability in OSN responses [78], differentiating through a Poisson sampler for the purposes of computing gradients was numerically challenging (requiring a continuous relaxation). For simplicity, we therefore used Gaussian noise.

For all analyses except the sweep over neural noise, we set the magnitude of the noise to be $\sigma_0 = 0.1$. This was to match the approximate mean activity $\mathbb{E}[r]$ when

computed over samples—see Glomeruli and *SI Appendix: Fig. S13* for histograms of activity.

The placement of E inside the nonlinearity does not constitute a principled mechanistic model integrating receptor signaling. Instead, we are using the Hill function as an empirical model for neural activity in the presence of mixtures.

C. Details of mutual information maximization

We leverage a variational formulation of mutual information first proposed by Nowozin *et al.* [55]. Following their presentation, we give a short overview of the theory underlying this approach, which was first presented by Nguyen et al. [105]. The *f-divergence* between two distributions P and Q defined on a domain $\mathcal{X}$ with densities $p(x)$ and $q(x)$ is defined as

$$D_f(P\|Q) = \int_{\mathcal{X}} q(x) f\left(\frac{p(x)}{q(x)}\right) \mathrm{d}x \tag{9}$$

The mutual information $MI(g, c)$ is a special case of this: namely the Kullback-Leibler divergence between the joint distribution $p(x) = p(g, c)$ and the product of marginals $q(x) = p(g)p(c)$, with $f(u) = u\log(u)$.

If f is convex and lower semi-continuous, then we can write the Fenchel conjugate

$$f^*(t) = \sup_{u \in \mathrm{dom} f} \{ut - f(u)\} \tag{10}$$

The functions f and f^* are dual in the sense that $f = f^{**}$. Then we can write

$$\begin{aligned} &D_f(P\|Q) \\ &= \int_{\mathcal{X}} q(x) \sup_{t \in \mathrm{dom} f^*} \left\{ t\frac{p(x)}{q(x)} - f^*(t) \right\} \mathrm{d}x && (11) \\ &\geq \sup_{T \in \mathcal{T}} \left\{ \int_{\mathcal{X}} p(x)T(x)\mathrm{d}x - \int_{\mathcal{X}} q(x) f^*(T(x))\mathrm{d}x \right\} && (12) \\ &= \sup_{T \in \mathcal{T}} \left\{ \mathbb{E}_{x\sim P}[T(x)] - \mathbb{E}_{x\sim Q}[f^*(T(x))] \right\}, && (13) \end{aligned}$$

where $\mathcal{T}$ is an appropriately-chosen class of test functions $T : \mathcal{X} \to \mathbb{R}$. We have samples from the joint distribution $p(g, c)$, and we obtain samples from the distribution $p(g)p(c)$ by shuffling activity. Thus, we can estimate the expectations in (13).

We then simply parameterize T using a small neural network, and optimize jointly over the layer in question and the parameters of T. Tschannen *et al.* [106] gave a thorough discussion of critic architectures. Following their recommendations, we used an inner product architecture for T so that $T(g, c) = \chi(g)^\top \psi(c)$, where χ and ψ are separate two-hidden-layer networks with width = 128. To sanity-check that our results did not depend strongly on the choice of critic network, we performed spot-checks using other critic architectures suggested by Tschannen *et al.* [106].

Such approaches are not without controversy. McAllester and Stratos [52] have argued that, for the Kullback-Leibler divergence (KLD), the expectations in (13) cannot be estimated without exponentially many samples. This is ultimately due to the fact that the KLD has $f^*(t) = \exp(t - 1)$. This theoretical result prompted investigation into why such estimators are still useful in applications such as representation learning [106].

Here, we sidestep this problem by using the Jensen-Shannon divergence (JSD), a symmetrized version of the Kullback-Leibler divergence (KLD) which has $f^*(t) = -\log(2 - \exp(t))$, thus avoiding the need for exponentially many samples. The greater stability and robustness of this Jensen-Shannon-based estimator compared to the exact KL formulation were empirically established by Hjelm *et al.* [53] and Huszár [54]. These variational approaches have gained considerable popularity in the machine learning community, to the point that maximizing mutual information is no longer considered an intractable problem [50, 53, 54].

D. Constrained optimization using mirror descent

When optimizing over W, E, and G, we must account for the constraints on each matrix: W must have positive elements, and both E and G must have non-negative elements and have all row sums equal to unity (*i.e.*, each row lies in a simplex). To handle these constraints, we use the framework of mirror descent [107, 108]. For a pedagogical introduction to mirror descent, we direct the reader to Vlad Niculae's blog post [109].

This approach requires specifying a ψ function which maps from the primal (constrained) space to the dual (unconstrained space), and ϕ function which maps from the dual space back to the primal space. So for W, to ensure positivity, we have used $\phi(u) = \exp(u) > 0$. For E and G optimizations, we used $\phi(u) = \mathrm{softmax}(u)_i = \exp(u_i)/\sum_i \exp(u_i)$. Mirror descent is equivalent to using "natural" gradient descent [110] in the dual space, but enjoys favorable numerical properties [108]. It greatly increased the stability of our optimization procedure.

All gradients were computed via automatic differentiation using the JAX library [111]. Initializations were as discussed in the main text: for W, we initialized using a scaled log-normal. For E, we varied the initialization depending on the analysis. In Figure 3, we initialized to random expression. In *SI Appendix: Fig. S8*, we initialized to canonical expression. Otherwise, we initialized to noncanonical expression. Typically this meant 3-7 receptors per neuron with roughly equal expression, as in Figure 5. However, in *SI Appendix: Fig. S11*, we initialized to 1-5 receptors per neuron with roughly equal expression. This was to respect the block structure of W when modeling coexpression of correlated receptors. For G, we initialized to random connectivity (Figure 6).

ACKNOWLEDGMENTS

We are grateful to A. D. T. Samuel and all the authors of Si *et al.* [37] for making their data publicly available. J.C.F.C. thanks D. M. Zimmerman and I. Chandok for helpful discussions regarding the evolution and biophysics of receptors. We thank C. Lienkaemper, M. A. Younger, and G. K. Ocker for inspiring discussions regarding non-canonical expression. Moreover, we are indebted to A. D. T. Samuel and the members of his group, and to G. K. Ocker and M. A. Younger, for their thoughtful comments, which have helped us improve our manuscript.

J.C.F.C. was supported by the Harvard Biophysics Graduate Program through training grant T32GM158477 from the National Institutes of Health. F.P. was supported by the Harvard Center for Brain Science (CBS)-NTT Fellowship Program on the Physics of Intelligence. V.N.M. was supported by grant number A47994 from NTT Research, Inc. J.A.Z.-V. was supported by a Junior Fellowship from the Harvard Society of Fellows.

AUTHOR CONTRIBUTIONS

J.C.F.C., V.N.M., and J.A.Z.-V. conceptualized the project, with input from F.P. J.C.F.C. wrote code, performed simulations, prepared figures, and wrote an initial draft of the paper, with input and assistance from all authors. All authors contributed to editing the manuscript. J.A.Z.-V. and V.N.M. jointly supervised this project.

DATA AND CODE AVAILABILITY

Code to perform all optimizations and reproduce all figures is available on GitHub at `https://github.com/VNMurthyLab/olfactory-ec`.

The dataset of estimated *Drosophila* larva receptor affinities from Si *et al.* [37] that we analyzed in Figure 2 was downloaded from A. D. T. Samuel's lab GitHub, where it is available under an MIT License: `https://github.com/samuellab/Larval-ORN/blob/master/Figure3/results/MLEFit.mat`.

[1] S. Hayden, M. Bekaert, T. A. Crider, S. Mariani, W. J. Murphy, and E. C. Teeling, Ecological adaptation determines functional mammalian olfactory subgenomes, Genome Research **20**, 1–9 (2009).

[2] W. A. Valencia-Montoya, N. E. Pierce, and N. W. Bellono, Evolution of sensory receptors, Annual Review of Cell and Developmental Biology **40**, 353–379 (2024).

[3] L. R. Yohe and P. Brand, Evolutionary ecology of chemosensation and its role in sensory drive, Current Zoology **64**, 525–533 (2018).

[4] T. O. Auer, M. A. Khallaf, A. F. Silbering, G. Zappia, K. Ellis, R. Álvarez-Ocaña, J. R. Arguello, B. S. Hansson, G. S. X. E. Jefferis, S. J. C. Caron, M. Knaden, and R. Benton, Olfactory receptor and circuit evolution promote host specialization, Nature **579**, 402–408 (2020).

[5] H. M. Robertson, Molecular evolution of the major arthropod chemoreceptor gene families, Annual Review of Entomology **64**, 227–242 (2019).

[6] Y. Niimura and M. Nei, Extensive gains and losses of olfactory receptor genes in mammalian evolution, PLoS ONE **2**, e708 (2007).

[7] M. Nei, Y. Niimura, and M. Nozawa, The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity, Nature Reviews Genetics **9**, 951–963 (2008).

[8] J. G. Hildebrand and G. M. Shepherd, Mechanisms of olfactory discrimination: Converging evidence for common principles across phyla, Annual Review of Neuroscience **20**, 595–631 (1997).

[9] K. A. Fulton, D. Zimmerman, A. Samuel, K. Vogt, and S. R. Datta, Common principles for odour coding across vertebrates and invertebrates, Nature Reviews Neuroscience **25**, 453–472 (2024).

[10] H. Saito, Q. Chi, H. Zhuang, H. Matsunami, and J. D. Mainland, Odor coding by a mammalian receptor repertoire, Science Signaling **2**, 10.1126/scisignal.2000016 (2009).

[11] E. A. Hallem and J. R. Carlson, Coding of odors by a receptor repertoire, Cell **125**, 143–160 (2006).

[12] N. Uchida and Z. F. Mainen, Speed and accuracy of olfactory discrimination in the rat, Nature Neuroscience **6**, 1224–1229 (2003).

[13] C. D. Wilson, G. O. Serrano, A. A. Koulakov, and D. Rinberg, A primacy code for odor identity, Nature Communications **8**, 10.1038/s41467-017-01432-4 (2017).

[14] T. Ackels, A. Erskine, D. Dasgupta, A. C. Marin, T. P. A. Warner, S. Tootoonian, I. Fukunaga, J. J. Harris, and A. T. Schaefer, Fast odour dynamics are encoded in the olfactory system and guide behaviour, Nature **593**, 558 (2021).

[15] D. Rokni, V. Hemmelder, V. Kapoor, and V. N. Murthy, An olfactory cocktail party: figure-ground segregation of odorants in rodents, Nature Neuroscience **17**, 1225–1232 (2014).

[16] S. Kundu, A. Ganguly, T. S. Chakraborty, A. Kumar, and O. Siddiqi, Synergism and combinatorial coding for binary odor mixture perception in *Drosophila*, eneuro **3**, ENEURO.0056 (2016).

[17] J. A. Riffell, E. Shlizerman, E. Sanders, L. Abrell, B. Medina, A. J. Hinterwirth, and J. N. Kutz, Flower discrimination by pollinators in a dynamic chemical environment, Science **344**, 1515–1518 (2014).

[18] S. Jayakumar, N. Rigolli, M. W. Mathis, M. Vergassola, A. Mathis, and V. N. Murthy, Mice navigate scent trails using predictive policies, bioRxiv 10.1101/2025.08.27.672631 (2025).

[19] J. Chapuis and D. A. Wilson, Bidirectional plasticity of cortical pattern recognition and behavioral sensory acuity, Nature Neuroscience **15**, 155–161 (2011).

[20] C.-Y. Su, C. Martelli, T. Emonet, and J. R. Carlson, Temporal coding of odor mixtures in an olfactory receptor neuron, Proceedings of the National Academy of Sciences **108**, 5075–5080 (2011).

[21] C. Martelli, J. R. Carlson, and T. Emonet, Intensity invariant dynamics and odor-specific latencies in olfactory receptor neuron response, The Journal of Neuroscience **33**, 6285–6297 (2013).

[22] R. Linsker, Self-organization in a perceptual network, Computer **21**, 105 (1988).

[23] E. P. Simoncelli and B. A. Olshausen, Natural image statistics and neural representation, Annual Review of Neuroscience **24**, 1193 (2001).

[24] S. Qin, Q. Li, C. Tang, and Y. Tu, Optimal compressed sensing strategies for an array of nonlinear olfactory receptor neurons with and without spontaneous activity, Proceedings of the National Academy of Sciences **116**, 20286–20295 (2019).

[25] T. Teşileanu, S. Cocco, R. Monasson, and V. Balasubramanian, Adaptation of olfactory receptor abundances for efficient coding, eLife **8**, 10.7554/elife.39279 (2019).

[26] D. Zwicker, A. Murugan, and M. P. Brenner, Receptor arrays optimized for natural odor statistics, Proceedings of the National Academy of Sciences **113**, 5570–5575 (2016).

[27] K. Krishnamurthy, A. M. Hermundstad, T. Mora, A. M. Walczak, and V. Balasubramanian, Disorder and the neural representation of complex odors, Frontiers in Computational Neuroscience **16**, 10.3389/fncom.2022.917786 (2022).

[28] V. Singh, M. Tchernookov, and V. Balasubramanian, What the odor is not: Estimation by elimination, Physical Review E **104**, 10.1103/physreve.104.024415 (2021).

[29] J. D. Victor, S. D. Boie, E. G. Connor, J. P. Crimaldi, G. B. Ermentrout, and K. I. Nagel, Olfactory navigation and the receptor nonlinearity, The Journal of Neuroscience **39**, 3713–3727 (2019).

[30] R. Benton, *Drosophila* olfaction: past, present and future, Proceedings of the Royal Society B: Biological Sciences **289**, 10.1098/rspb.2022.2054 (2022).

[31] L. B. Vosshall, A. M. Wong, and R. Axel, An olfactory sensory map in the fly brain, Cell **102**, 147–159 (2000).

[32] E. Bashkirova and S. Lomvardas, Olfactory receptor genes make the case for inter-chromosomal interactions, Current Opinion in Genetics I& Development **55**, 106–113 (2019).

[33] A. Brahma, D. D. Frank, P. D. H. Pastor, P. K. Piekarski, W. Wang, J.-D. Luo, T. S. Carroll, and D. J. Kronauer, Transcriptional and post-transcriptional control of odorant receptor choice in ants, Current Biology **33**, 5456 (2023).

[34] Y. Zhou, T. F. O'Connell, M. Ghaninia, B. H. Smith,

E. J. Hong, and T. O. Sharpee, Low-dimensional olfactory signatures of fruit ripening and fermentation, bioRxiv 10.1101/2024.06.16.599229 (2024).

[35] A. Celani, E. Villermaux, and M. Vergassola, Odor landscapes in turbulent environments, Physical Review X **4**, 10.1103/physrevx.4.041015 (2014).

[36] T. Nowotny and P. Szyszka, Dynamics of odor-evoked activity patterns in the olfactory system, in *Advances in Dynamics, Patterns, Cognition* (Springer International Publishing, 2017) p. 243–261.

[37] G. Si, J. K. Kanwal, Y. Hu, C. J. Tabone, J. Baron, M. Berck, G. Vignoud, and A. D. Samuel, Structured odorant response patterns across a complete olfactory receptor neuron population, Neuron **101**, 950 (2019).

[38] J. D. Zak, G. Reddy, M. Vergassola, and V. N. Murthy, Antagonistic odor interactions in olfactory sensory neurons are widespread in freely breathing mice, Nature Communications **11**, 10.1038/s41467-020-17124-5 (2020).

[39] A. Dewan, A. Cichy, J. Zhang, K. Miguel, P. Feinstein, D. Rinberg, and T. Bozza, Single olfactory receptors set odor detection thresholds, Nature Communications **9**, 10.1038/s41467-018-05129-0 (2018).

[40] X. Grosmaitre, A. Vassalli, P. Mombaerts, G. M. Shepherd, and M. Ma, Odorant responses of olfactory sensory neurons expressing the odorant receptor mor23: A patch clamp analysis in gene-targeted mice, Proceedings of the National Academy of Sciences **103**, 1970–1975 (2006).

[41] C. Lienkaemper, M. A. Younger, and G. K. Ocker, When non-canonical olfaction is optimal, Biorxiv 10.1101/2025.02.27.640624 (2025).

[42] S. M. Potter, C. Zheng, D. S. Koos, P. Feinstein, S. E. Fraser, and P. Mombaerts, Structure and emergence of specific olfactory glomeruli in the mouse, The Journal of Neuroscience **21**, 9713–9723 (2001).

[43] M. Herre, O. V. Goldman, T.-C. Lu, G. Caballero-Vidal, Y. Qi, Z. N. Gilbert, Z. Gong, T. Morita, S. Rahiel, M. Ghaninia, R. Ignell, B. J. Matthews, H. Li, L. B. Vosshall, and M. A. Younger, Non-canonical odor coding in the mosquito, Cell **185**, 3104 (2022).

[44] V. Bhandawat, S. R. Olsen, N. W. Gouwens, M. L. Schlief, and R. I. Wilson, Sensory processing in the *Drosophila* antennal lobe increases reliability and separability of ensemble odor representations, Nature Neuroscience **10**, 1474–1482 (2007).

[45] E. D. Adavi, V. L. dos Anjos, S. Kotb, H. C. Metz, D. Tian, Z. Zhao, J. L. Zung, N. H. Rose, and C. S. McBride, Olfactory receptor coexpression and co-option in the dengue mosquito, bioRxiv 10.1101/2024.08.21.608847 (2024).

[46] T. Tsukahara, D. H. Brann, S. L. Pashkovski, G. Guitchounts, T. Bozza, and S. R. Datta, A transcriptional rheostat couples past activity to future sensory responses, Cell **184**, 6326 (2021).

[47] B. Bintu, Y. Isogai, X. Zhuang, and C. Dulac, Molecular and spatial organization of the primary olfactory system and its responses to social odors, bioRxiv 10.1101/2025.05.02.651832 (2025).

[48] D. H. Brann, T. Tsukahara, C. Tau, D. Kalloor, R. Lubash, L. Thamarai Kannan, N. Klimpert, M. Kollo, M. Escamilla-Del-Arenal, B. Bintu, T. Bozza, and S. R. Datta, A spatial code governs olfactory receptor choice and aligns sensory maps in the nose and brain, bioRxiv 10.1101/2025.05.02.651738 (2025).

[49] S. Shao, M. Meister, and J. Gjorgjieva, Efficient population coding of sensory stimuli, Phys. Rev. Res. **5**, 043205 (2023).

[50] E. Abdelaleem, K. M. Martini, and I. Nemenman, Accurate estimation of mutual information in high dimensional data, arXiv 10.48550/ARXIV.2506.00330 (2025).

[51] N. Y. Jun, G. Field, and J. Pearson, Efficient coding, channel capacity, and the emergence of retinal mosaics, in *Advances in Neural Information Processing Systems*, Vol. 35, edited by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Curran Associates, Inc., 2022) pp. 32311–32324.

[52] D. McAllester and K. Stratos, Formal limitations on the measurement of mutual information, in *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, Vol. 108, edited by S. Chiappa and R. Calandra (PMLR, 2020) pp. 875–884.

[53] R. D. Hjelm, A. Fedorov, S. Lavoie-Marchildon, K. Grewal, P. Bachman, A. Trischler, and Y. Bengio, Learning deep representations by mutual information estimation and maximization, in *International Conference on Learning Representations* (2019).

[54] F. Huszár, How (not) to train your generative model: Scheduled sampling, likelihood, adversary?, arXiv 10.48550/ARXIV.1511.05101 (2015).

[55] S. Nowozin, B. Cseke, and R. Tomioka, f-gan: Training generative neural samplers using variational divergence minimization, in *Advances in Neural Information Processing Systems*, Vol. 29, edited by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc., 2016).

[56] D. M. Bear, J.-M. Lassance, H. E. Hoekstra, and S. R. Datta, The evolving neural and genetic architecture of vertebrate olfaction, Current Biology **26**, R1039–R1049 (2016).

[57] T. Sakurai, H. Mitsuno, A. Mikami, K. Uchino, M. Tabuchi, F. Zhang, H. Sezutsu, and R. Kanzaki, Targeted disruption of a single sex pheromone receptor gene completely abolishes in vivo pheromone response in the silkmoth, Scientific Reports **5**, 10.1038/srep11001 (2015).

[58] J. del Mármol, M. A. Yedlin, and V. Ruta, The structural basis of odorant recognition in insect olfactory receptors, Nature **597**, 126–131 (2021).

[59] W. Meyerhof, C. Batram, C. Kuhn, A. Brockhoff, E. Chudoba, B. Bufe, G. Appendino, and M. Behrens, The molecular receptive ranges of human tas2r bitter taste receptors, Chemical Senses **35**, 157–170 (2009).

[60] P. Pregitzer, M. Schubert, H. Breer, B. S. Hansson, S. Sachse, and J. Krieger, Plant odorants interfere with detection of sex pheromone signals by male heliothis virescens, Frontiers in Cellular Neuroscience **6**, 10.3389/fncel.2012.00042 (2012).

[61] M. Renou, V. Party, A. Rouyar, and S. Anton, Olfactory signal coding in an odor background, Biosystems **136**, 35–45 (2015).

[62] S. Namiki, S. Iwabuchi, and R. Kanzaki, Representation of a mixture of pheromone and host plant odor by antennal lobe projection neurons of the silkmoth bombyx mori, Journal of Comparative Physiology A **194**, 501–515 (2008).

[63] K. Mika and R. Benton, Olfactory receptor gene regulation in insects: Multiple mechanisms for

singular expression, Frontiers in Neuroscience **15**, 10.3389/fnins.2021.738088 (2021).

[64] G. L. Glotzer, P. D. H. Pastor, and D. J. C. Kronauer, Transcriptional interference gates monogenic odorant receptor expression in ants, bioRxiv 10.1101/2025.08.21.671318 (2025).

[65] B. A. Olshausen and D. J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images, Nature **381**, 607–609 (1996).

[66] D. Task, C.-C. Lin, A. Vulpe, A. Afify, S. Ballou, M. Brbic, P. Schlegel, J. Raji, G. S. Jefferis, H. Li, K. Menuz, and C. J. Potter, Chemoreceptor co-expression in *Drosophila melanogaster* olfactory neurons, eLife **11**, 10.7554/elife.72599 (2022).

[67] A. L. Goldman, W. Van der Goes van Naters, D. Lessing, C. G. Warr, and J. R. Carlson, Coexpression of two functional odor receptors in one neuron, Neuron **45**, 661–666 (2005).

[68] S. Lebreton, F. Borrero-Echeverry, F. Gonzalez, M. Solum, E. A. Wallin, E. Hedenström, B. S. Hansson, A.-L. Gustavsson, M. Bengtsson, G. Birgersson, W. B. Walker, H. K. M. Dweck, P. G. Becher, and P. Witzgall, A drosophila female pheromone elicits species-specific long-range attraction via an olfactory channel with dual specificity for sex and food, BMC Biology **15**, 10.1186/s12915-017-0427-x (2017).

[69] A. Sánchez-Gracia, F. G. Vieira, and J. Rozas, Molecular evolution of the major chemosensory gene families in insects, Heredity **103**, 208–216 (2009).

[70] K. A. Adipietro, J. D. Mainland, and H. Matsunami, Functional evolution of mammalian odorant receptors, PLoS Genetics **8**, e1002821 (2012).

[71] B. Malnic, P. A. Godfrey, and L. B. Buck, The human olfactory receptor gene family, Proceedings of the National Academy of Sciences **101**, 2584–2589 (2004).

[72] J.-Y. Yang, T. F. O'Connell, W.-M. M. Hsu, M. S. Bauer, K. V. Dylla, T. O. Sharpee, and E. J. Hong, Restructuring of olfactory representations in the fly brain around odor relationships in natural sources, bioRxiv 10.1101/2023.02.15.528627 (2023).

[73] K. Mori, H. Nagao, and Y. Yoshihara, The olfactory bulb: Coding and processing of odor molecule information, Science **286**, 711–715 (1999).

[74] P. Mombaerts, Molecular biology of odorant receptors in vertebrates, Annual Review of Neuroscience **22**, 487–509 (1999).

[75] S. Laughlin, A simple coding procedure enhances a neuron's information capacity, Zeitschrift für Naturforschung C **36**, 910 (1981).

[76] R. Baraniuk, Compressive sensing [lecture notes], IEEE Signal Processing Magazine **24**, 118–121 (2007).

[77] T. Blumensath, Compressed sensing with nonlinear observations and related nonlinear optimization problems, IEEE Transactions on Information Theory **59**, 3466 (2013).

[78] J. Zavatone-Veth, P. Masset, W. Tong, J. D. Zak, V. Murthy, and C. Pehlevan, Neural circuits for fast Poisson compressed sensing in the olfactory bulb, in *Advances in Neural Information Processing Systems*, Vol. 36, edited by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Curran Associates, Inc., 2023) pp. 64793–64828.

[79] B. Malnic, J. Hirono, T. Sato, and L. B. Buck, Combinatorial receptor codes for odors, Cell **96**, 713–723 (1999).

[80] M. Wachowiak, A. Dewan, T. Bozza, T. F. O'Connell, and E. J. Hong, Recalibrating olfactory neuroscience to the range of naturally occurring odor concentrations, Journal of Neuroscience **45**, 10.1523/JNEUROSCI.1872-24.2024 (2025).

[81] J. Murlis and C. D. Jones, Fine-scale structure of odour plumes in relation to insect orientation to distant pheromone and other attractant sources, Physiological Entomology **6**, 71–86 (1981).

[82] K. R. Mylne and P. J. Mason, Concentration fluctuation measurements in a dispersing plume at a range of up to 1000 m, Quarterly Journal of the Royal Meteorological Society **117**, 177–206 (1991).

[83] N. J. Vickers, Mechanisms of animal navigation in odor plumes, The Biological Bulletin **198**, 203 (2000).

[84] F. van Breugel and M. Dickinson, Plume-tracking behavior of flying *Drosophila* emerges from a set of distinct sensory-motor reflexes, Current Biology **24**, 274–286 (2014).

[85] S. A. Kreher, J. Y. Kwon, and J. R. Carlson, The molecular basis of odor coding in the *Drosophila* larva, Neuron **46**, 445–456 (2005).

[86] P. Pfister, B. C. Smith, B. J. Evans, J. H. Brann, C. Trimmer, M. Sheikh, R. Arroyave, G. Reddy, H.-Y. Jeong, D. A. Raps, Z. Peterlin, M. Vergassola, and M. E. Rogers, Odorant receptor inhibition is fundamental to odor encoding, Current Biology **30**, 2574 (2020).

[87] L. Xu, W. Li, V. Voleti, D.-J. Zou, E. M. C. Hillman, and S. Firestein, Widespread receptor-driven modulation in peripheral olfactory coding, Science **368**, 10.1126/science.aaz5390 (2020).

[88] G. Reddy, J. D. Zak, M. Vergassola, and V. N. Murthy, Antagonism in olfactory receptor neurons and its implications for the perception of odor mixtures, eLife **7**, 10.7554/elife.34958 (2018).

[89] P. Ramdya and R. Benton, Evolving olfactory systems on the fly, Trends in Genetics **26**, 307–316 (2010).

[90] S. W. Santoro and C. Dulac, The activity-dependent histone variant h2be modulates the life span of olfactory neurons, eLife **1**, e00070 (2012).

[91] W. Hong and L. Luo, Genetic control of wiring specificity in the fly olfactory system, Genetics **196**, 17–29 (2014).

[92] D.-J. Zou, P. Feinstein, A. L. Rivers, G. A. Mathews, A. Kim, C. A. Greer, P. Mombaerts, and S. Firestein, Postnatal refinement of peripheral olfactory projections, Science **304**, 1976–1979 (2004).

[93] G. J. Gutierrez, F. Rieke, and E. T. Shea-Brown, Nonlinear convergence boosts information coding in circuits with parallel outputs, Proceedings of the National Academy of Sciences **118**, 10.1073/pnas.1921882118 (2021).

[94] P. Y. Wang, Y. Sun, R. Axel, L. Abbott, and G. R. Yang, Evolving the olfactory system with machine learning, Neuron **109**, 3879 (2021).

[95] P. Cisek and B. Y. Hayden, Neuroscience needs evolution, Philosophical Transactions of the Royal Society B: Biological Sciences **377**, 20200518 (2022).

[96] N. Kadakia and T. Emonet, Front-end Weber-Fechner gain control enhances the fidelity of combinatorial odor coding, eLife **8**, e45293 (2019).

[97] K. Choi, W. Rosenbluth, I. R. Graf, N. Kadakia, and T. Emonet, Bifurcation enhances temporal information encoding in the olfactory periphery, PRX Life **2**,

10.1103/prxlife.2.043011 (2024).

[98] H. Giaffar, S. Shuvaev, D. Rinberg, and A. A. Koulakov, The primacy model and the structure of olfactory space, PLOS Computational Biology **20**, 1 (2024).

[99] F. X. P. Bourassa, P. François, G. Reddy, and M. Vergassola, Manifold learning for olfactory habituation to strongly fluctuating backgrounds, bioRxiv 10.1101/2025.05.26.656161 (2025).

[100] N. M. Chapochnikov, C. Pehlevan, and D. B. Chklovskii, Normative and mechanistic model of an adaptive circuit for efficient encoding and feature extraction, Proceedings of the National Academy of Sciences **120**, e2117484120 (2023).

[101] G. Reddy, B. I. Shraiman, and M. Vergassola, Sector search strategies for odor trail tracking, Proceedings of the National Academy of Sciences **119**, e2107431118 (2022).

[102] S. Golkar, J. Berman, D. Lipshutz, R. M. Haret, T. Gollisch, and D. B. Chklovskii, Neuronal temporal filters as normal mode extractors, Physical Review Research **6**, 013111 (2024).

[103] S. C. L. Durian, K. Bojanek, O. Marre, and S. E. Palmer, Preserving predictive information under biologically plausible compression, bioRxiv 10.1101/2025.03.12.642864 (2025).

[104] L. van Giesen, P. B. Kilian, C. A. H. Allard, and N. W. Bellono, Molecular basis of chemotactile sensation in octopus, Cell **183**, 594 (2020).

[105] X. Nguyen, M. J. Wainwright, and M. Jordan, Estimating divergence functionals and the likelihood ratio by penalized convex risk minimization, in *Advances in Neural Information Processing Systems*, Vol. 20, edited by J. Platt, D. Koller, Y. Singer, and S. Roweis (Curran Associates, Inc., 2007).

[106] M. Tschannen, J. Djolonga, P. K. Rubenstein, S. Gelly, and M. Lucic, On mutual information maximization for representation learning, in *International Conference on Learning Representations* (2020).

[107] A. Beck and M. Teboulle, Mirror descent and nonlinear projected subgradient methods for convex optimization, Operations Research Letters **31**, 167 (2003).

[108] G. Raskutti and S. Mukherjee, The information geometry of mirror descent, IEEE Transactions on Information Theory **61**, 1451 (2015).

[109] V. Niculae, Optimizing with constraints: reparametrization and geometry, online blog post (2020), URL: `https://vene.ro/blog/mirror-descent.html`.

[110] S.-i. Amari, Natural gradient works efficiently in learning, Neural Computation **10**, 251 (1998).

[111] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: composable transformations of Python+NumPy programs (2018).

Supporting Information

Appendix A: Estimating the minimum effective sensitivity

In our analysis, with the exception of Figure 2, we have initialized the sensitivities W_{ij} to the scale W_{ij}^{min} below which W_{ij} is effectively 0. This allows us to confirm, for example, that broad receptor tuning is not an artifact of initialization (see *SI Appendix: Figs. S2, S6*). Knowing W_{ij}^{min} is also necessary to understand the range of biologically meaningful tuning, which differs from the range of the optimized W_{ij}. For example, many of the optimized elements W_{ij} are far below working machine precision ($\sim 10^{-16}$), but this does not mean that the optimal W spans more than 16 orders of magnitude. Even biologically, for example, $W_{ij} = 10^{-14}$ is equivalent to $W_{ij} = 0$, since none of our odor concentrations are close to 10^{14}. To correctly interpret our results, we must understand the minimum effective value for W_{ij}. We discuss below how that minimum relevant value is set by the maximum concentration c_j.

Assuming canonical expression for simplicity, the mean activity of a neuron expressing the i-th receptor is given by

$$\varphi(W_i^\top c) = \frac{1}{1 + (W_i^\top c)^{-n}}, \tag{A1}$$

where W_i is the vector of sensitivities for the i-th receptor. To determine the minimum relevant sensitivity, we would like to estimate the smallest W_{ij} such that changing the concentration c_j produces a detectable change in $\varphi(W_i^\top c)$. We do so by freezing all components of the concentration vector but the j-th, and write

$$W_i^\top c = a_0 + W_{ij} c_j \quad \text{for} \quad a_0 = \sum_{j' \neq j} W_{ij'} c_{j'}. \tag{A2}$$

Then, we linearize in $W_{ij} c_j$ around the half-occupancy point $a_0 = 1$, which is where the Hill function is maximally sensitive to changes in inputs:

$$\varphi(W_i^\top c) \approx \frac{1}{2} + \frac{n}{4} W_{ij} c_j. \tag{A3}$$

In the absence of neural noise, any contribution above working precision results in an numerically meaningful change in the activity. But for non-zero neural noise levels σ_0, we must estimate a minimum distinguishable difference. For example, in our modeling of the adult fly, $\sigma_0 = 0.1$, and we have approximately 21 neurons per receptor. So the effective noise is approximately $\frac{0.1}{\sqrt{21}}$. Therefore, a rough estimate for a meaningful difference in mean activity compared to the noise is to take $W_{ij} c_j$ to scale with the standard deviation of the effective noise, *i.e.*,

$$W_{ij} c_j \sim \frac{0.1}{\sqrt{21}} \sim \frac{1}{50}. \tag{A4}$$

Other choices for this value simply result in a different prefactor in the final expression for W_{ij}^{min}. To obtain the lowest possible W_{ij}^{min}, we set $c_j = c_{max}$. This results in

$$W_{ij}^{min} = \frac{1}{50 c_{max}}. \tag{A5}$$

This reasoning ensures that $W_{ij} < W_{ij}^{min}$ is functionally equivalent to $W_{ij} = 0$. In other words, the spread of W_{ij} below W_{ij}^{min} is not biologically meaningful. This allows us to make reasonable claims about the distribution of W_{ij}, as in Fig. 2. Finally, recall that c is log-normal, so the maximum c_j is on the order of $c_{max} = \exp(\mu + 3\sigma)$. In this way, the value of W_{ij}^{min} (and thus, the definition of the sparsity of W) depends on the environmental noise.

This argument is purely heuristic, and only intended to guide our interpretation of numerical results. We do not use it as the basis for any biological claims. In particular, we have not conducted any analyses that depend on a precise estimate of W_{ij}^{min}. For example, Qin *et al.* [24] report that as σ_c increases, W_{opt} becomes less sparse (more entries are non-zero). This is plausible, but difficult to confirm in our hands because the meaningful threshold for sparsity is also changing with σ_c. Furthermore, one could take $c_{max} = \exp(\mu + 3\sigma)$, but one might take a more typical value ($\exp(\mu + 2\sigma)$) or a more extreme one ($\exp(\mu + 4\sigma)$). Each choice for c_{max} results in a different W_{ij}^{min}.

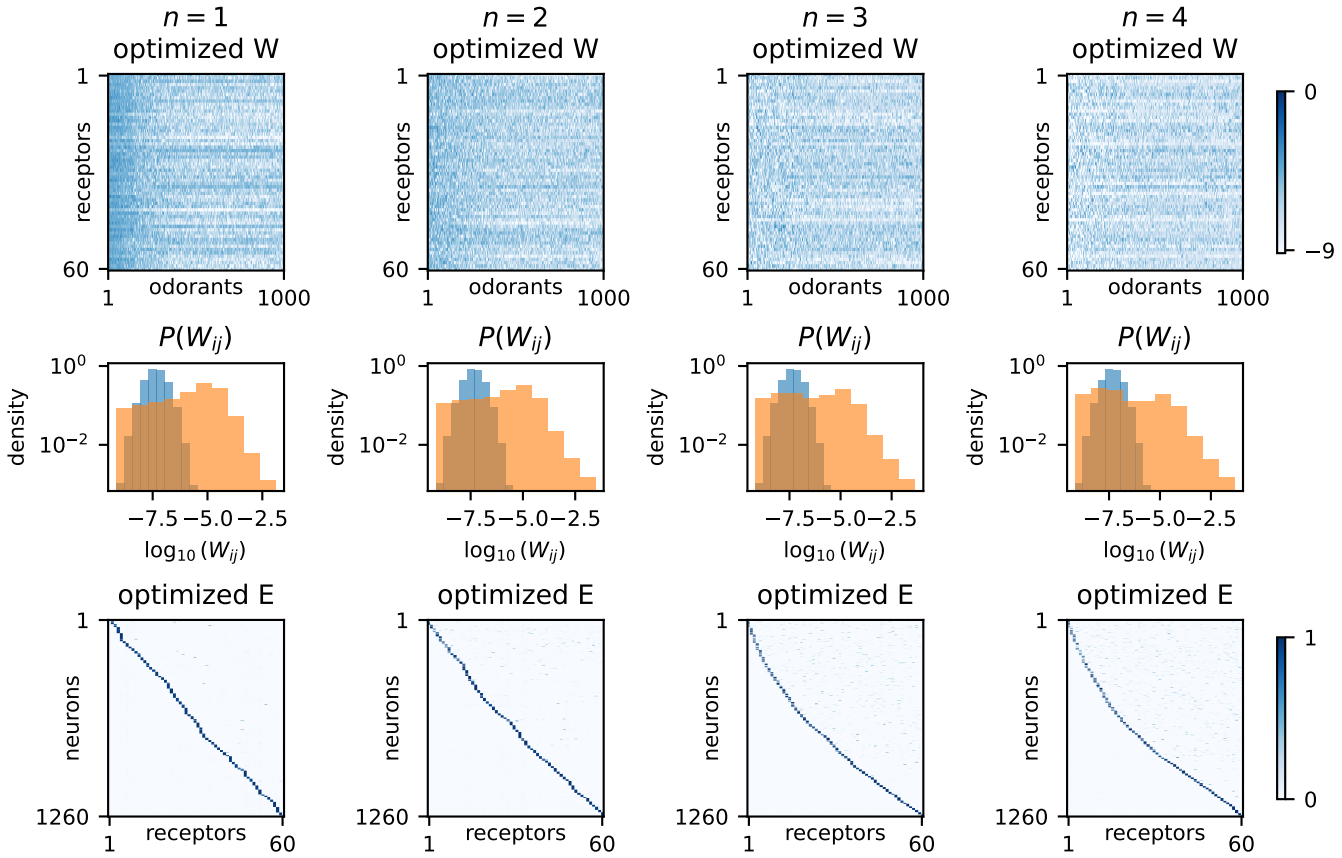

FIG. S1. Effect of the Hill coefficient n on optimized W and E. E optimizations are performed using shuffled W_{opt}. In the histograms, blue represents the initial W_{ij}, and orange the optimized W_{ij}. E was initialized to noncanonical olfaction as in Figure 5a, b.

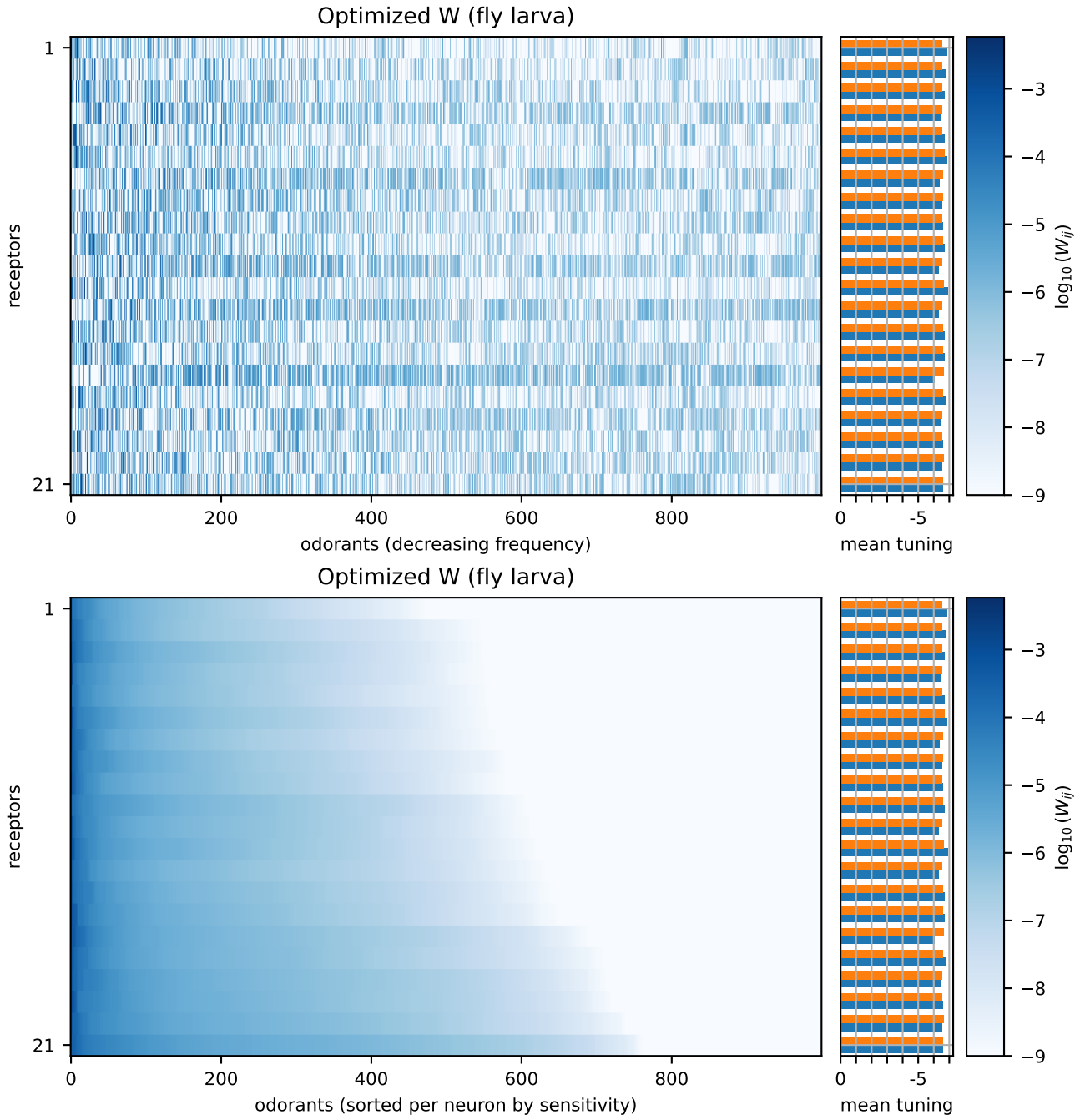

FIG. S2. Receptor tuning curves in the optimized W matrix for the fly larva. The same data is shown twice: for the top panel, odorants are sorted by frequency, whereas for the bottom panel, odorants are sorted within each receptor by sensitivity. Bars on the right indicate the mean log-sensitivity of each receptor (blue is true, orange is shuffled). The receptor mean tuning is more variable than would be expected by chance (blue bars vary more than orange bars). This is consistent with work [11] showing that some receptors are very narrowly tuned (or almost exclusively inhibitory), while others are broadly tuned.

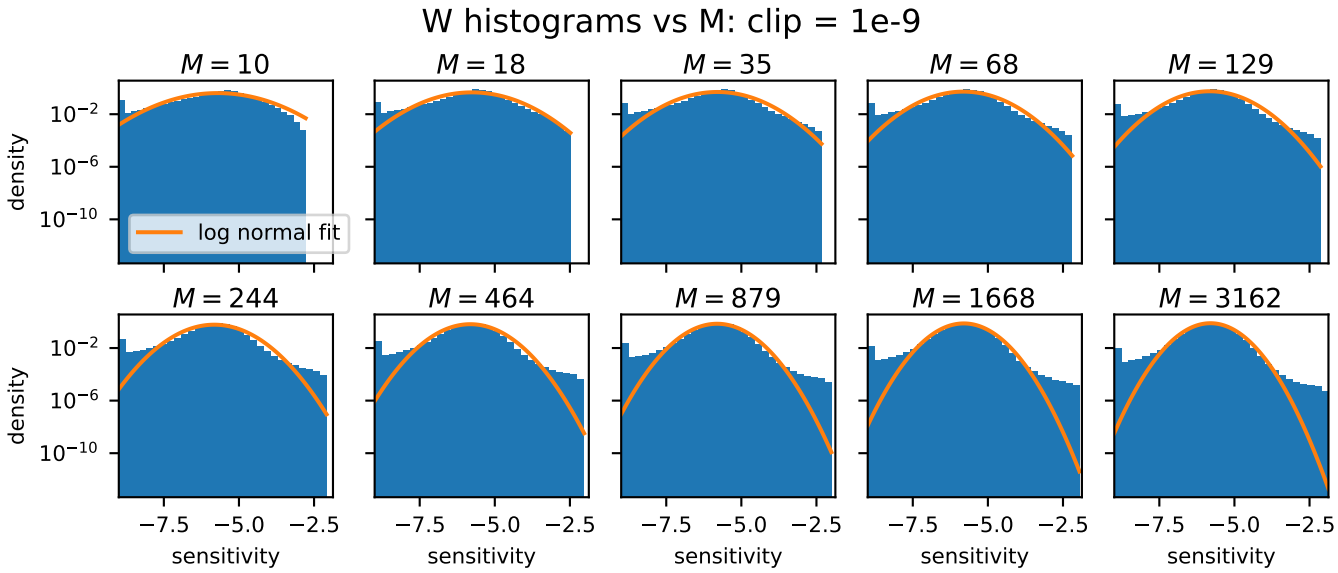

FIG. S3. Histograms of $p(\log_{10}(W_{ij}))$ after optimization, and log-normal fits. Note that the log density is plotted. The heavy right tails of the true distribution compared to the log-normal drive the emergence of specialists for large M (second row).

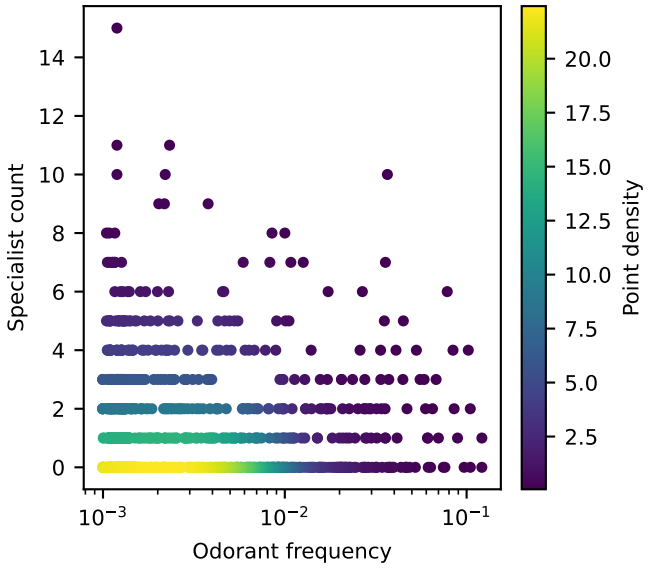

FIG. S4. Specialist receptor count vs. odorant frequency. Specialists targeting higher frequency odorants are not particularly favored.

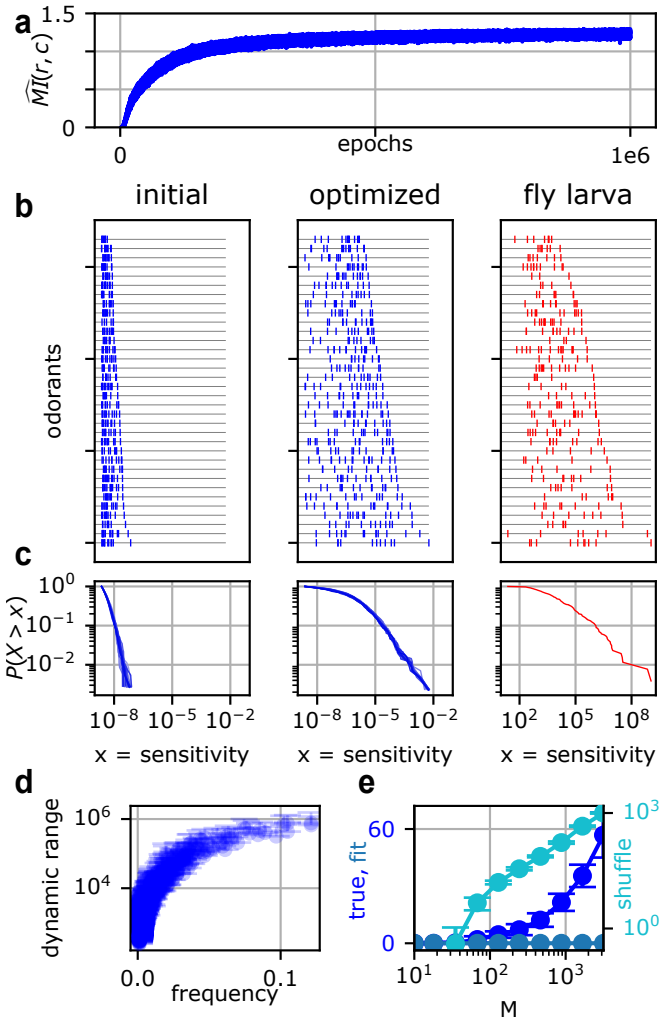

FIG. S5. Optimized W distribution under initialization to minimum W_{ij} (compare to Figure 2 in main text.) The only significant difference is in the "shuffle" curve for panel e, which is shifted dramatically upward (note log scale). This reflects the fact that, for large M, most sensitivities do not need to move under optimization. When W is shuffled, this causes many specialists to emerge. However, in the true W, the number of specialists was comparable to Figure 2e (mean 57 specialists for $M = 3162$ receptors rather than mean 27 specialists for the $1/E[||c||]$ initialization in the main text). This indicates some pressure against developing too many specialists, even in an over-parameterized ($M > N$) regime. See Appendix A for a discussion of how the minimum effective W_{ij} is computed.

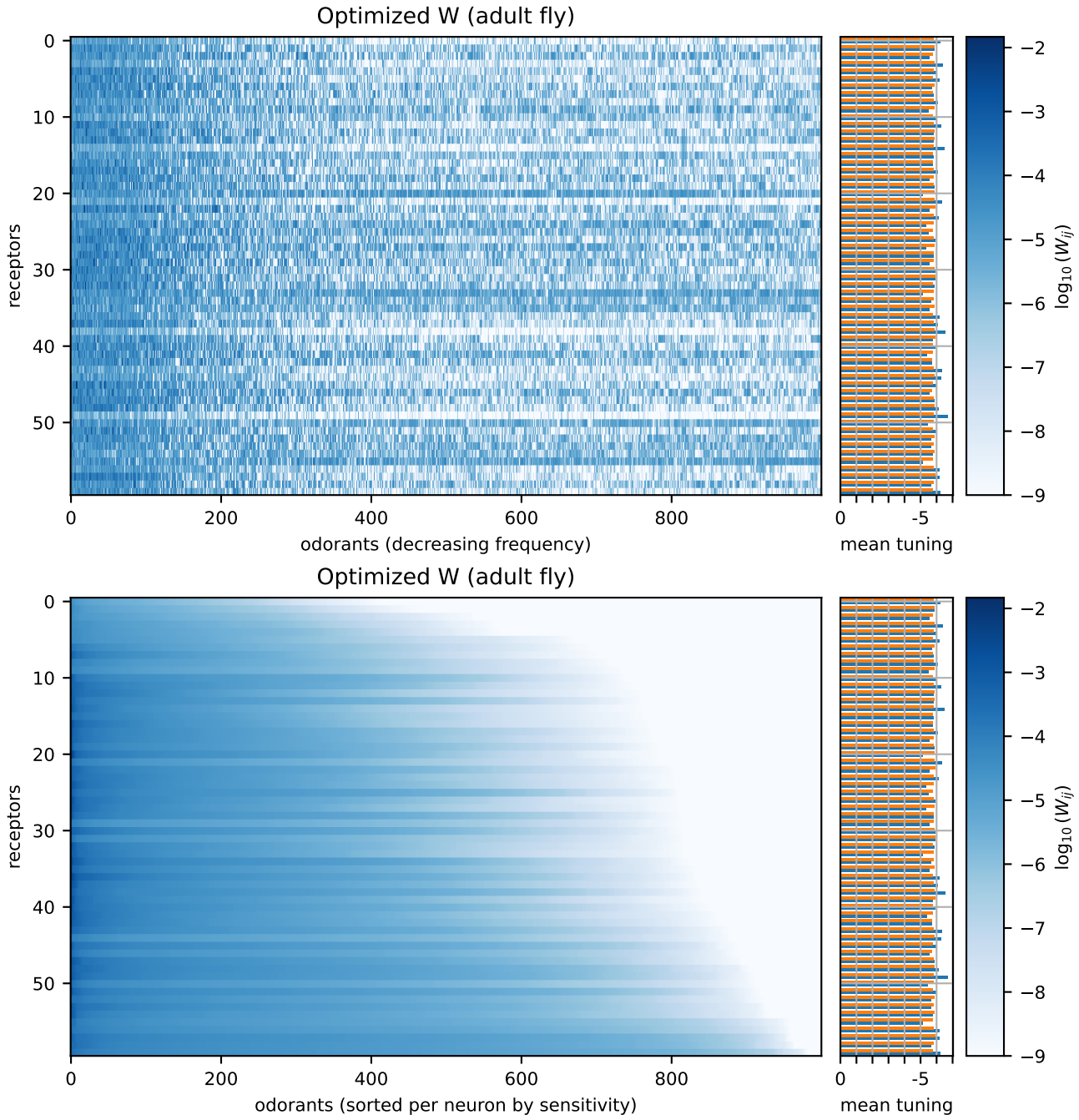

FIG. S6. Receptor tuning curves in the optimized W matrix for the adult fly. The same data is shown twice: for the top panel, odorants are sorted by frequency, whereas for the bottom panel, odorants are sorted within each receptor by sensitivity. Bars on the right indicate the mean log-sensitivity of each receptor (blue is true, orange is shuffled). The receptor mean tuning is more variable than would be expected by chance (blue bars vary more than orange bars). This is consistent with work [11] showing that some receptors are very narrowly tuned (or almost exclusively inhibitory), while others are broadly tuned.

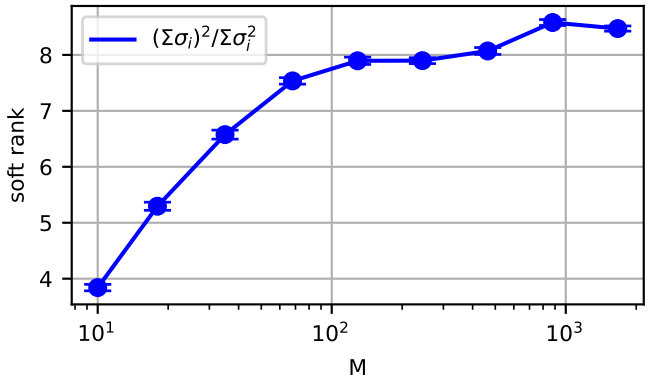

FIG. S7. Soft rank of W vs. receptor count M. The soft rank is computed using $(\sum \sigma_i)^2 / \sum(\sigma_i^2)$, where σ_i are the singular values of W. This measure is 1 when W is rank 1 and M when W is full-rank. Error bars indicate the standard deviation over 20 runs.

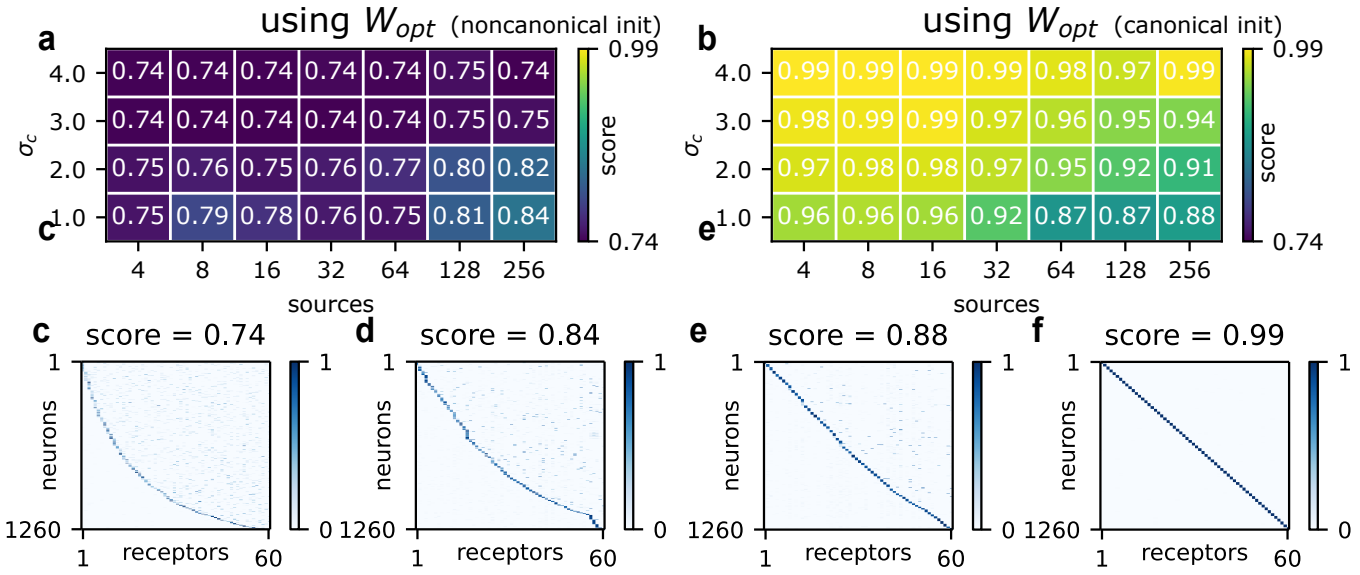

FIG. S8. Phase diagram of E optimization given W_{opt}, comparing canonical and noncanonical initialization. Only the optimizations in the bottom right corner (low noise, many sources) of the phase diagram move meaningfully from initialization.

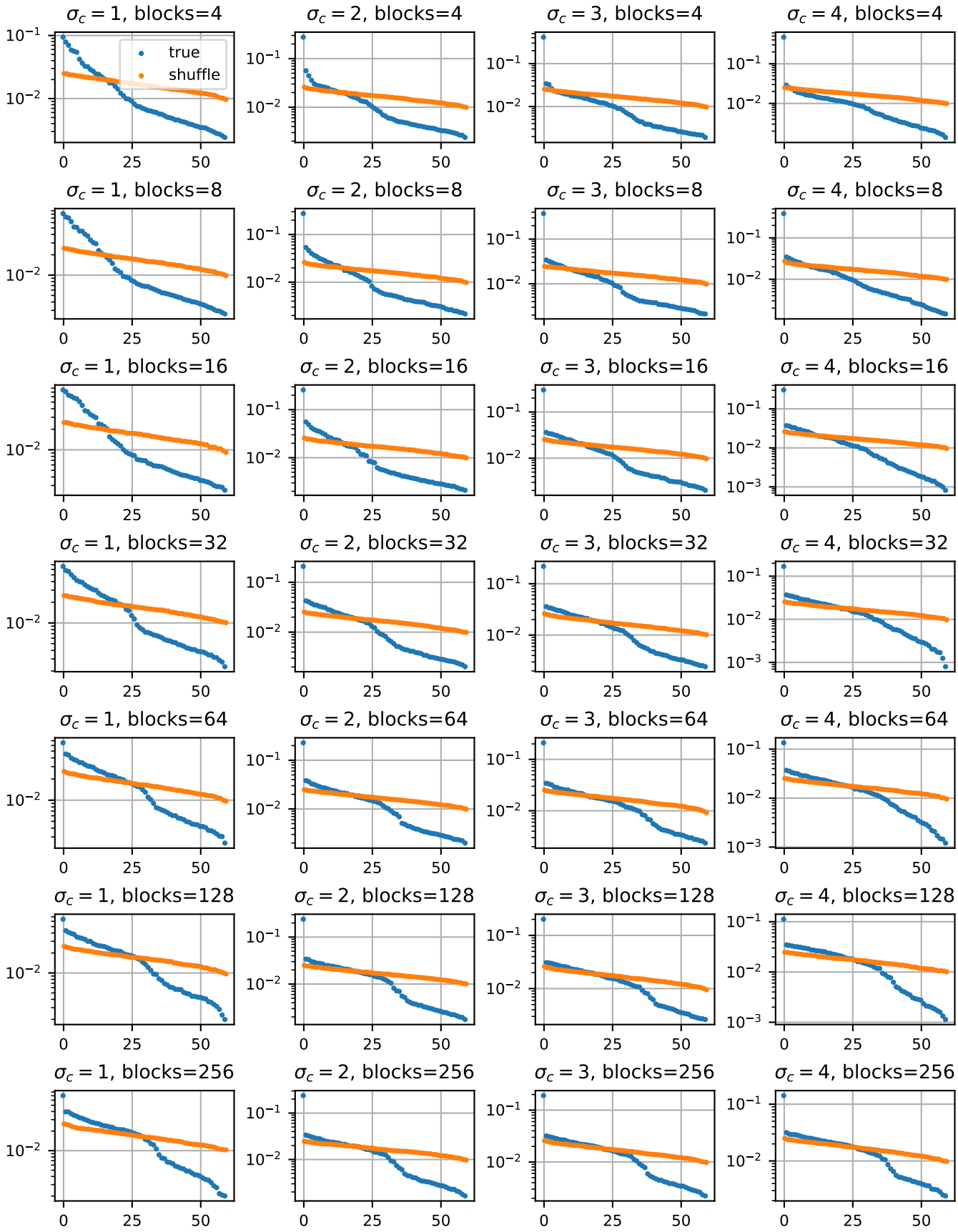

FIG. S9. Eigenvalue spectra of optimized W across environments, given the standard odorant model (most odorants are rare, some are common, as in Figure 1).

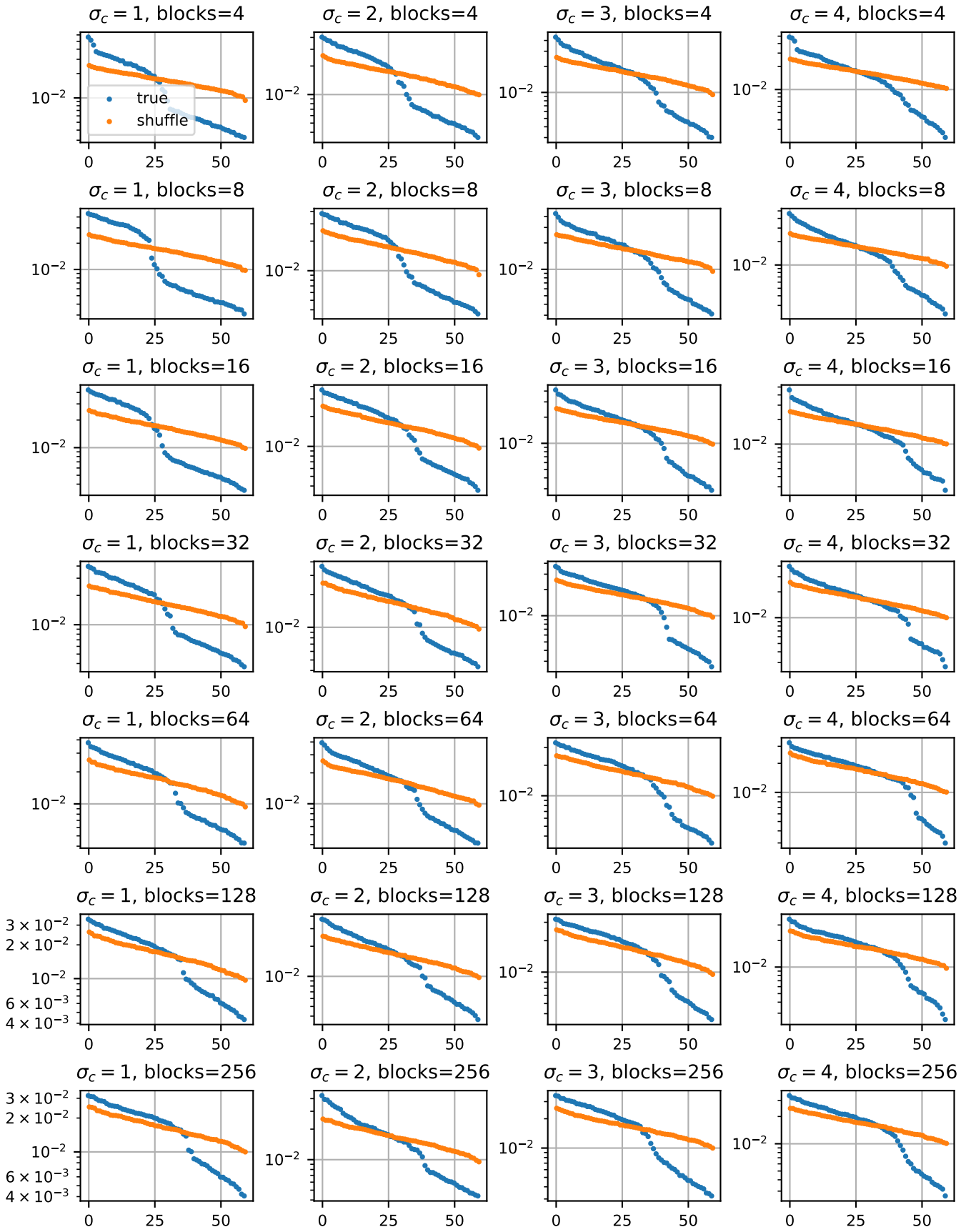

FIG. S10. Eigenvalue spectra of W under an odorant model with constant frequencies (so all $N = 1000$ odorants occur with equal probability). The low rank structure in Fig. S9 is abolished.

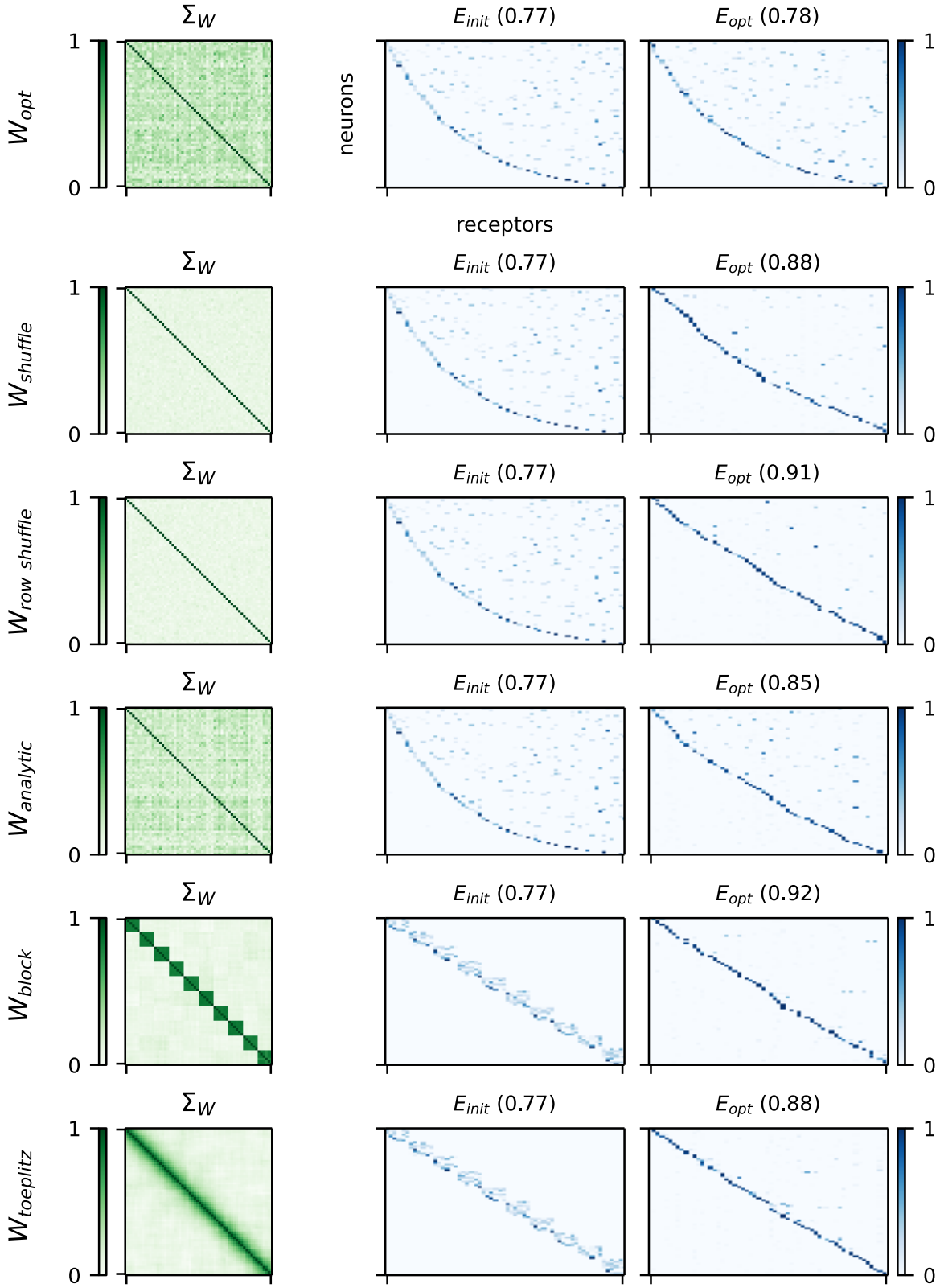

FIG. S11. Optimized expression given a variety of different W matrices. Environmental parameters are set to 64 sources and $\sigma_c = 2$. For W_{block} and $W_{toeplitz}$, the E matrices were initialized to exhibit coexpression of correlated receptors (coexpression by descent) [45]. This did not lead to more noncanonical expression than the shuffled W, in which unrelated receptors were coexpressed. $W_{analytic}$ is generated by fitting a log normal to W_{opt}, then masking by the same sparsity structure present in W_{opt}. This induces a non-negligible correlation structure in $W_{analytic}$.

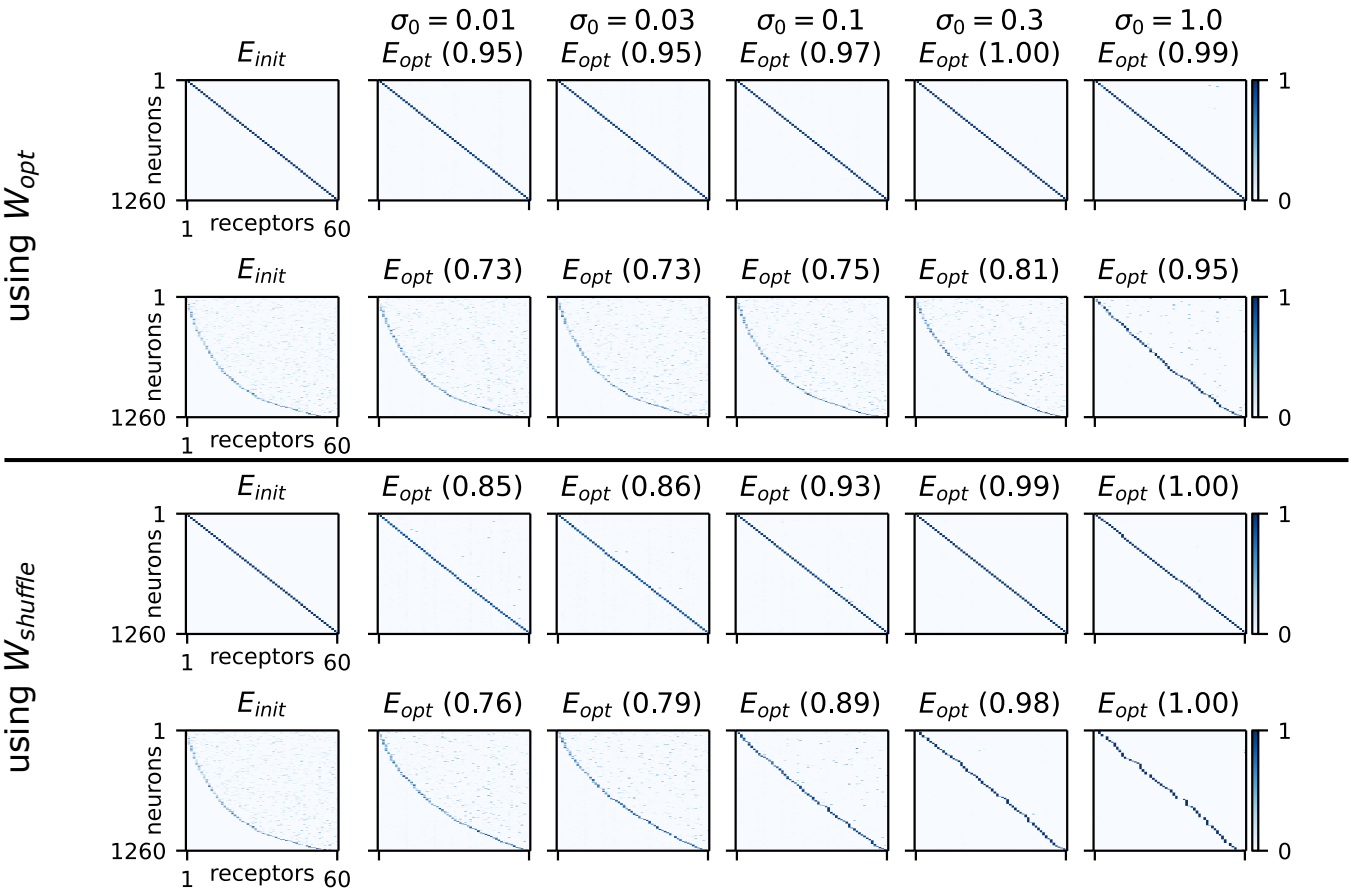

FIG. S12. Sweeping across neural noise σ_0. The canonical score (as defined in Figure 5) is shown in parentheses. Both canonical and noncanonical initializations are shown. For very low levels of neural noise, some noncanonical expression is tolerated. For $\sigma_0 \in [0.1, 1.0]$, canonical olfaction is strongly favored. Environmental parameters are set to 64 sources and $\sigma_c = 2$.

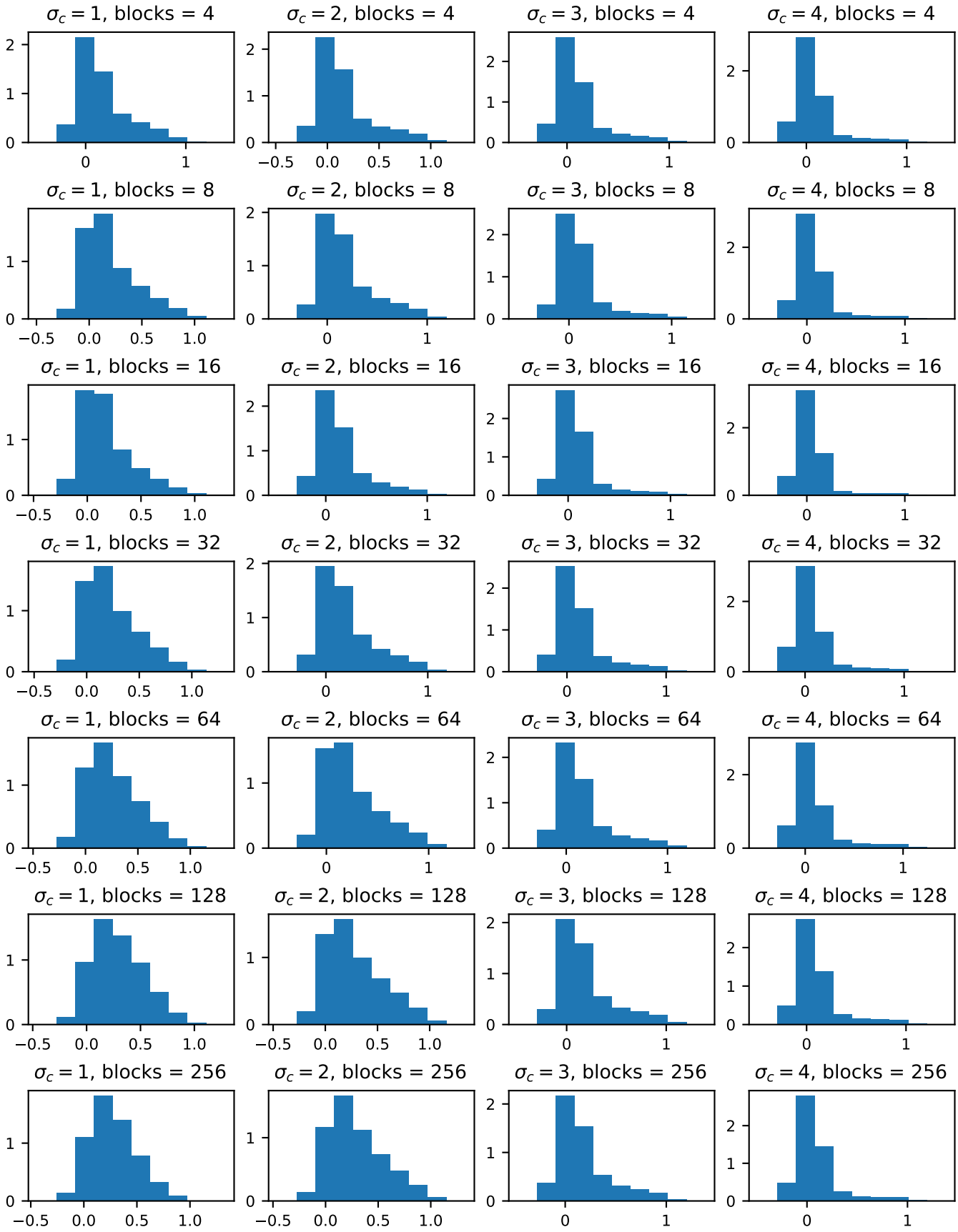

FIG. S13. ORN activity distributions across environmental parameters using shuffled W and corresponding optimized E.

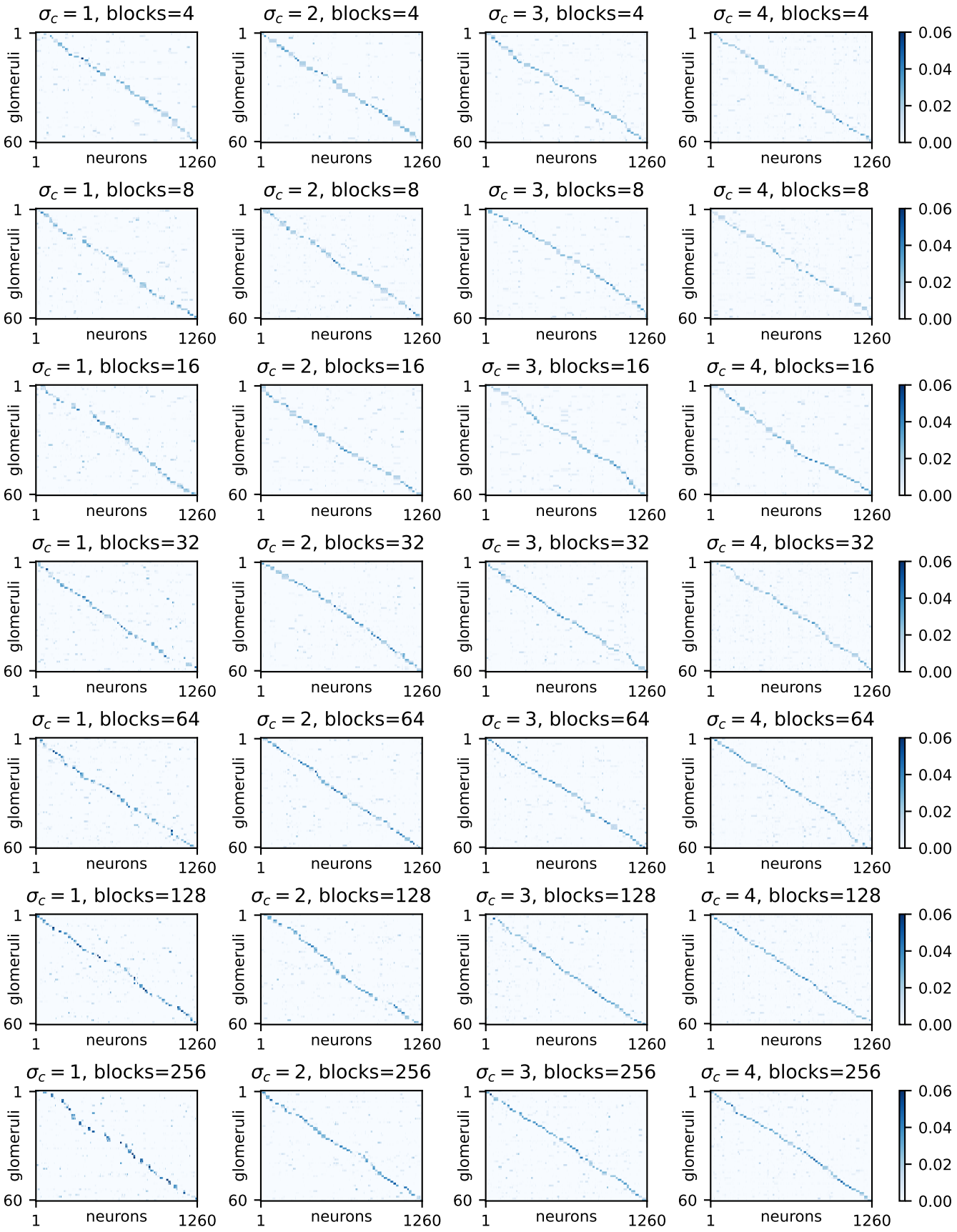

FIG. S14. Glomerular convergence across environments using shuffled W and corresponding optimized E.

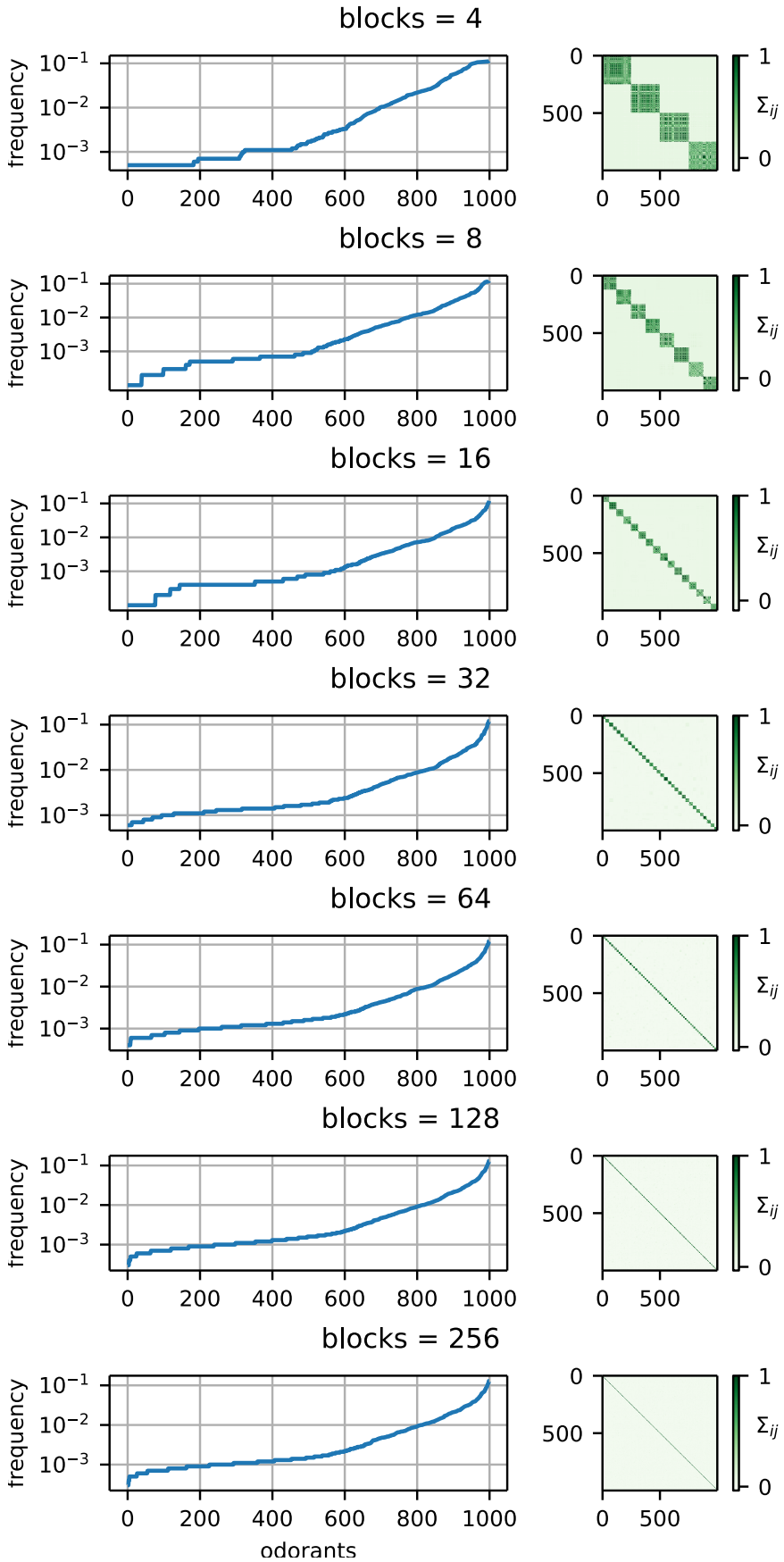

FIG. S15. Mean and covariance of the binarized odorant vector c_{bin} as the number of blocks k is varied.