

1  
2      **Front-end Weber-Fechner gain control**  
3      **enhances the fidelity of combinatorial odor coding**

4      Nirag Kadakia

5      *Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06511*

6      Thierry Emonet\*

7      *Department of Molecular, Cellular, and Developmental Biology*

8      *Department of Physics, Yale University, New Haven, CT 06511 and*

\*Corresponding Author

In a previous paper (Gorur-Shandilya et al 2017), we showed that *Drosophila* olfactory receptor neurons (ORNs) expressing the co-receptor Orco scale their gain inversely with mean odor intensity, according to the Weber-Fechner law of psychophysics. Here we investigate the implications of this front-end mechanism for odor coding capacity in natural environments, where the intensity and timescales of odor signals can span several orders of magnitude, and odors can mix together. We find that ORN adaptation promotes the reconstruction of odor identity from dynamic odor signals, even in the presence of confounding background odors and rapid intensity fluctuations. These enhancements are further aided by known downstream transformations in the antennal lobe and mushroom body. Our results, which are applicable to various odor classification and reconstruction schemes, stem from the fact that ORN adaptation is not intrinsic to the identity of the receptor involved. Instead, a feedback mechanism adjusts receptor sensitivity based on the activity of the receptor-Orco complex, in accordance with the Weber-Fechner law. This common scaling of the gain across Orco-expressing ORNs may be one of the features of ORN adaptation that helps preserve combinatorial odor codes in naturalistic landscapes.

9      Keywords: Insect olfaction | Adaptation | Compressed sensing | Olfactory receptor neurons | Combinatorial  
10     coding | Orco

11     Animals identify and discriminate odors using olfactory receptors (Ors) expressed in olfactory receptor neurons (ORNs) [1–4]. Individual ORNs, which typically express a single Or, respond to many odorants, while individual odorants activate many distinct ORNs [5–8]. Odors are thus encoded by the combinatorial patterns of activity they elicit in the sensing periphery [5–7, 9–11], patterns decoded downstream into behavioral response [12, 13]. Still, ethologically-relevant odors are often mixed with background ones [14, 15] and intensity can vary widely and rapidly as odors are carried by the wind [16–19]. How are odors recognized reliably despite these confounds? In *Drosophila melanogaster*, ORN dose response curves exhibit similar Hill coefficients but distinct power-law distributed activation thresholds [6, 20], which together with inhibitory odorants enhance coding capacity [6, 20–22]. In antennal lobe (AL) glomeruli, mutual lateral inhibition normalizes population response, reducing the dependency of activity patterns on odor concentration [23, 24]. Further downstream, sparse connectivity to the mushroom body (MB) helps maintain neural representations of odors, and facilitates compressed sensing and associative learning schemes [25–28]. Finally, temporal features of neural responses contribute to concentration-invariant representations of odor identity [29–32].

37     Here we examine how short-time ORN adaptation at the very front-end of the insect olfactory circuit contributes to the fidelity of odor encoding. Our theoret-

40     ical study is motivated by the recent discovery of in-  
41     variances in the signal transduction and adaptation dy-  
42     namics of ORNs expressing the co-receptor Orco. ORN  
43     response is initiated upon binding of odorant molecules  
44     to olfactory receptors (Ors), opening the ion channels  
45     they form with the co-receptor Orco [33, 34]. Because  
46     of differences in odor-receptor affinities, the responses  
47     of ORNs to diverse odorants of the same concentra-  
48     tion differ widely [6, 35, 36]. In contrast, downstream  
49     from this input nonlinearity, signal transduction and  
50     adaptation dynamics exhibit a surprising degree of in-  
51     variance with respect to odor-receptor identity: reverse-  
52     correlation analysis of ORN response to fluctuating stim-  
53     uli produces highly stereotyped, concentration-invariant  
54     response filters [20, 37, 38].

55     These properties stem in part from an apparently in-  
56     variant adaptive scaling law in ORNs: gain varies in-  
57     versely with mean odor concentration according to the  
58     Weber-Fechner Law of psychophysics [39, 40], irrespec-  
59     tive of the odor-receptor combination [38, 41, 42]. This  
60     invariance can be traced back to adaptative feedback  
61     mechanisms in odor transduction, upstream of ORN fir-  
62     ing [38, 41–43], which depend on the activity of the sig-  
63     naling pathway rather than on the identity of its re-  
64     ceptor [43]. The generality of the adaptive scaling sug-  
65     gests it could be mediated by the highly conserved Orco  
66     co-receptor [34, 44–46]. Indeed, phosphorylation sites  
67     have been recently identified on Orco, some being im-  
68     plicated in odor desensitization, albeit over much longer

69 timescales [46, 47].

70 While in a simpler system such as *E. coli* chemotaxis [48], adaptive feedback via the Weber-Fechner  
 71 Law robustly maintains sensitivity over concentration  
 72 changes, the implication for a multiple-channel system  
 73 – which combines information from hundreds of cells  
 74 with overlapping receptive fields – is less clear. Here  
 75 we combine a biophysical model of ORN adaptive re-  
 76 sponse and neural firing with various sparse signal de-  
 77 coding frameworks to explore how ORN adaptation with  
 78 Weber-Fechner scaling affects combinatorial coding and  
 79 decoding of odor signals spanning varying degrees of in-  
 80 tensity, molecular complexity, and temporal structure.  
 81 We find that this front-end adaptive mechanism pro-  
 82 motes the accurate discrimination of odor signals from  
 83 backgrounds of varying molecular complexity, and aids  
 84 other known mechanisms of neural processing in the ol-  
 85 factory circuit to maintain representations of odor iden-  
 86 tity across environmental changes.

88

## RESULTS

89

### Model of ORN sensing repertoire

90 To model ORN firing rates in response to time-  
 91 dependent odor signals, we extended a minimal  
 92 model [38] that reproduces the Weber-Fechner gain adap-  
 93 tation and firing rate dynamics measured in individual  
 94 *Drosophila* ORNs in response to Gaussian and natural-  
 95 istic signals (code available on GitHub[49]).

96 We consider a repertoire of  $M = 50$  ORN types that  
 97 each express one type of Or together with the co-receptor  
 98 Orco [33]. Within ORNs of type  $a = 1, \dots, M$ , Or-  
 99 Orco complexes form non-selective cation channels [34]  
 100 (Fig. 1a) that switch between active and inactive confor-  
 101 mations, while simultaneously binding to odorants  $i$  with  
 102 affinity constants,  $K_{ai}^*$  and  $K_{ai}$ , respectively [38, 43]. For  
 103 simplicity we only consider agonists, i.e.  $K_{ai}^* > K_{ai}$ , and  
 104 assume receptors can only bind one odorant at a time.  
 105 The analysis can easily be extended to include inhibitory  
 106 odorants, which increases coding capacity [21]. Dissocia-  
 107 tion (inverse affinity) constants are chosen from a power  
 108 law distribution ( $\alpha = 0.35$ ) recently found across ORN-  
 109 odor pairs in *Drosophila* larvae [20]. For a handful of  
 110 ORNs, we choose a very large value for one of the  $K_{ai}^*$   
 111 to mimic high responders to private odorants relevant to  
 112 innate responses [36]. These private odors do not affect  
 113 the general findings.

114 Assuming that odorant binding and conformation  
 115 changes are faster than other reactions in the signaling  
 116 pathway, the fraction of channels of type  $a$  that are active  
 117 at steady state is:

$$A_a(t) = \frac{C_a^* + C_a^* \mathbf{K}_a^* \cdot \mathbf{s}(t)}{C_a^* + C_a^* \mathbf{K}_a^* \cdot \mathbf{s}(t) + C_a + C_a \mathbf{K}_a \cdot \mathbf{s}(t)}. \quad (1)$$

118  $C_a$  and  $C_a^*$  represent unbound channels in the inactive  
 119 and active conformation. Here,  $\mathbf{K}_a \cdot \mathbf{s}(t) = \sum_i^K K_{ai} s_i(t)$ ,  
 120 where  $s_i(t)$  is the time-dependent concentration of the  
 121  $i$ -th monomolecular component of the odor signal  $\mathbf{s}(t)$  at  
 122 time  $t$  (Fig. 1b).  $N = 150$  is the size of the molecu-  
 123 lar odorant space (Fig. 1b). Eq. 1 can be rearranged as  
 124 (derivation in Methods):

$$A_a(t) = \left[ 1 + \exp \left( \epsilon_a(t) + \ln \left( \frac{1 + \mathbf{K}_a \cdot \mathbf{s}(t)}{1 + \mathbf{K}_a^* \cdot \mathbf{s}(t)} \right) \right) \right]^{-1}. \quad (2)$$

125 The two terms in the exponential represent the change  
 126 in the channel’s free energy due to the binding of odorant  
 127  $i$ , and the free energy difference  $\epsilon_a$  between the unbound  
 128 states  $C_a$  and  $C_a^*$ , in units of  $k_B T$ . Because  $K_{ai}^* > K_{ai}$ ,  
 129 a sudden increase in the concentration of excitatory odor  
 130 results in an increase in activity  $A_a$ .

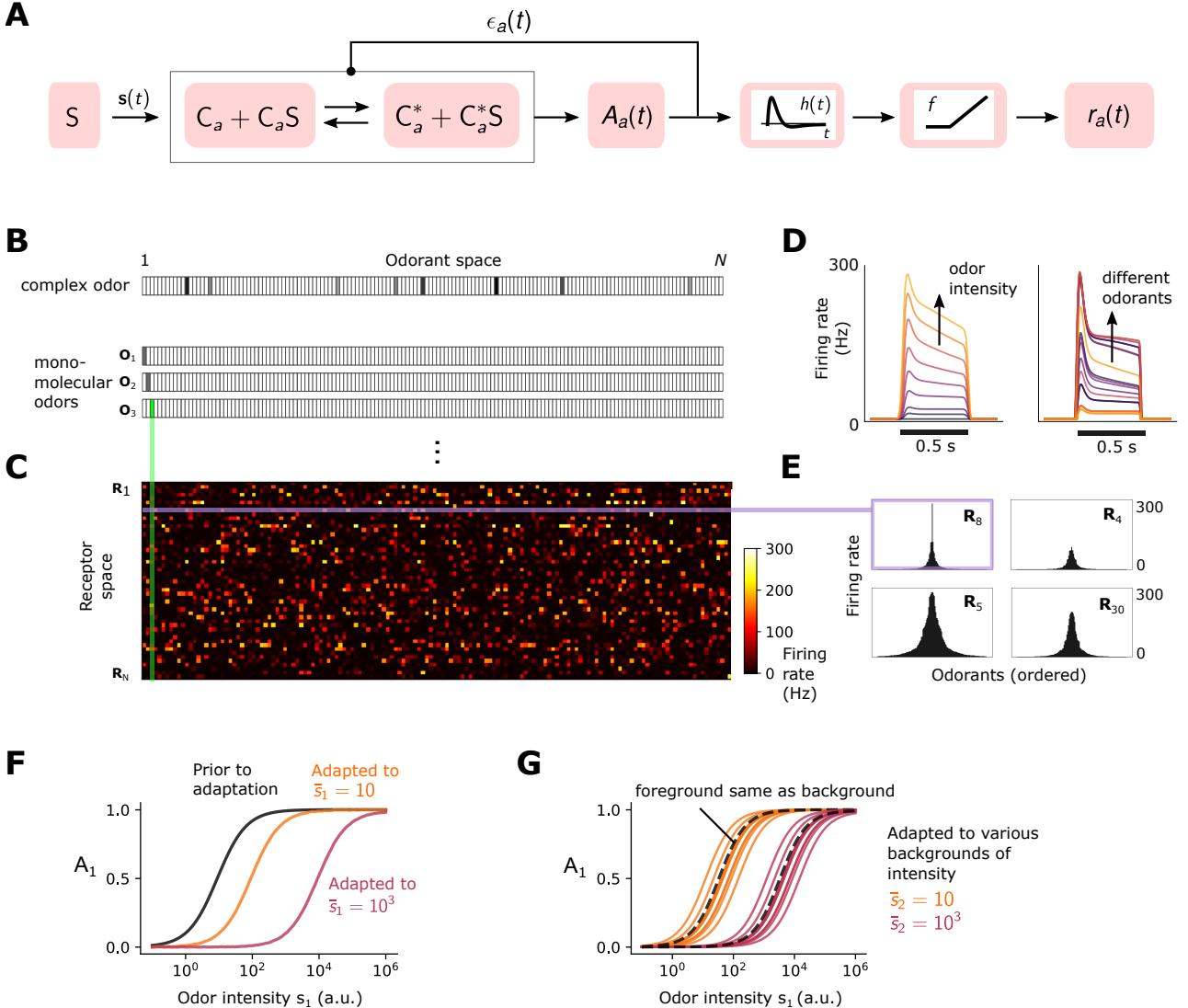
131 Upon prolonged stimulation, ORNs adapt. At least  
 132 one form of adaptation, which takes place over short time  
 133 scale,  $\tau \simeq 250$  ms [38], involves a negative feedback of  
 134 the Or-Orco channel activity onto the channel sensitiv-  
 135 ity [38, 43]. To model this adaptation process, we assume  
 136 that inward currents elicited by activating Or-Orco chan-  
 137 nels eventually result in an increase of the free energy  
 138 difference  $\epsilon_a(t)$ , possibly via a feedback onto Orco [34]:

$$\tau \frac{d\epsilon_a(t)}{dt} = A_a(t) - A_{0a}, \quad (3)$$

139 where  $\epsilon_{L,a} < \epsilon_a(t) < \epsilon_{H,a}$ . The lower bound  $\epsilon_{L,a}$  de-  
 140 termines the spontaneous activity of the channel. The  
 141 higher bound  $\epsilon_{H,a}$  determines the concentrations of odors  
 142 at which adaptation is unable to keep up and saturation  
 143 occurs [38]. Through these dynamics,  $\epsilon_a(t)$  can com-  
 144 pensate for changes in free energy due to ligand binding (see  
 145 Eq. 2), returning the activity  $A_a$  towards an adapted  
 146 level  $A_{0a}$  above the spontaneous activity. Since  $\epsilon_a$  is  
 147 bounded below, a minimum amount of signal intensity  
 148 is needed for adaptation to kick in. Finally, the firing  
 149 rate is modeled by passing the activity  $A_a(t)$  through  
 150 the derivative-taking bi-lobed filter  $h(t)$  and a rectifying  
 151 nonlinearity  $f$  [38]:

$$r_a(t) = f(h(t) \otimes A_a(t)), \quad (4)$$

152 where  $\otimes$  is convolution. When deconvolved from stim-  
 153 ulus dynamics, the shapes of the temporal kernels of  
 154 *Drosophila* ORNs that express Orco tend to be stereo-  
 155 typed for many odor-receptor combination [20, 37, 38]



**FIG. 1: Simple ORN model.** [38] **A** Or/Orco complexes of type  $a$  switch between active  $C_a^*$  and inactive conformations  $C_a$ . Binding an excitatory odorant ( $S$  in the diagram) favors the active state. The active fraction is determined by the free energy difference between inactive and active conformations of the Or/Orco complex in its unbound state,  $\epsilon_a(t)$  (in units of  $k_B T$ ), and by odorant binding with affinity constants  $\mathbf{K}_a^* = (K_{a1}^*, \dots, K_{ai}^*, \dots, K_{aN}^*)$  and  $\mathbf{K}_a$  for the active and inactive conformations, respectively (Eqs. 1-2). Adaptation is mediated by a negative feedback [43] from the activity of the channel onto the free energy difference  $\epsilon_a(t)$  with timescale  $\tau$ . ORN firing rates  $r_a(t)$  are generated by passing  $A_a(t)$  through a linear temporal filter  $h(t)$  and a nonlinear thresholding function  $f$ . **B** Odors are represented by  $N$ -dimensional vectors  $\mathbf{s} = (s_1, \dots, s_i, \dots, s_N)$ , whose components  $s_i$  are the concentrations of the individual molecular constituents of  $\mathbf{s}$ . **C** Step-stimulus firing rate of 50 ORNs to the  $N=150$  possible monomolecular odorants  $\mathbf{s} = s_i$ , given power-law distributed affinity constants [20]. **D** Temporal responses of a representative ORN to a pulse stimulus, for a single odorant at several intensities (left), or to many odorants of the same intensity (right). **E** Representative ORN tuning curves (a single row of the response matrix in C, ordered by magnitude). Tuning curves are diverse, mimicking measured responses [6]. **F** Dose-response of an ORN before (black) and after adaptation to either a low (yellow) or high (magenta) odor concentration. **G** Same, but the ORN was allowed to first adapt to one of various backgrounds of differing identities, before the foreground (same as in F) was presented. Also shown is the specific case when the foreground and background have the same identity (dashed lines).

(although there are known exceptions such as super-sustained responses [35]). Moreover, adaptation is not intrinsic to the receptor [43]. Accordingly, for simplicity  $\tau$ ,  $h(t)$ , and  $f$  are assumed independent of receptor and odorant identities.

This minimal model reproduces the essential features

of ORN response to odorant pulses [37, 42, 43]. In the absence of stimulus, ORNs fire spontaneously at rates (1-10 Hz) [6] set by the lower free energy bound  $\epsilon_{L,a}$ , which we choose from a normal distribution (Fig. 1d) [37, 43]. For sufficiently strong stimuli, adaptation causes  $\epsilon_a$  to increase, compensating for the drop in free energy differ-

ence due to ligand binding. This gradually reduces the firing rate to a steady state level  $r(A_{0a}) \simeq 30\text{-}40 \text{ Hz}$  [38] (Fig. 1d). The diversity of temporal firing responses and tuning curves measured experimentally [6, 29–31, 35] arise naturally in the model due to the distribution of chemical affinity constants and the nonlinearity of Eq. 2 (Figs. 1b-1e).

The model also reproduces Weber-Fechner scaling of the gain with the inverse of the mean odorant intensity  $\bar{s}_i$  [38, 42]. For small fluctuations  $\Delta s_i$  around  $\bar{s}_i$ , we have from Eq. 2 that  $\Delta A_a / \Delta s_i \simeq A_a(\bar{s}_i) (1 - A_a(\bar{s}_i)) / \bar{s}_i$ , whereby Weber's Law is satisfied provided  $A_a(\bar{s}_i)$  is approximately constant (derivation in Methods). In our model, since the rate of adaptation depends only on the activity of the ion channel (right hand-side of Eq. 3), then in the adapted state we have  $A_a(\bar{s}_i) \simeq A_{0a}$ , ensuring that the gain scales like  $1/\bar{s}_i$ . This process adjusts the sensitivity of the ORN by matching the dose responses to the mean signal concentration, while maintaining their log-slopes (Fig. 1f). However, for foreground odors mixed with background odors to which the system has adapted, the dose response curves now exhibit background-dependent shifts (Fig. 1g).

While this phenomenological model could be extended to include further details – e.g. we could relax the quasi-steady-state assumption in Eq. 2 and use a more complex model for channel adaptation and neural firing [38] – this minimally-parameterized form captures the key dynamical properties of Orco-expressing ORNs relevant to our study: receptor-independent adaptation [43] with Weber-Fechner scaling [38, 41, 42] that maintains response time independent of mean stimulus intensity [37, 38], along with a diversity of temporal firing patterns in response to a panel of monomolecular odorants [6, 29–31, 35] (Fig. 1d-1e).

### Front-end Weber-Fechner adaptation preserves odor coding among background and intensity confounds.

To investigate how front-end adaptation and its scaling according to Weber-Fechner's law affect the representations of odor identity within the repertoire of ORN response, we quantified the similarity between the responses  $\mathbf{r}_1$  and  $\mathbf{r}_2$  of the ORN repertoire to different stimuli  $\mathbf{s}_1$  and  $\mathbf{s}_2$  by measuring the Euclidean distance between  $\mathbf{r}_1$  and  $\mathbf{r}_2$ . Since it is not possible to visualize these 50-dimensional vectors, we projected them onto a lower-dimensional space using t-distributed stochastic neighbor embedding (t-SNE) [51]. Like principle component analysis (PCA), t-SNE preserves similarity between objects (Fig. 2a), but is more suitable than PCA for objects that are related nonlinearly – in our case, the dependency of firing rates on odor concentrations (Eq. 2).

We first examined how an adaptive or non-adaptive ORN repertoire encodes odor identity in an odor envi-

ronment that contains a foreground odor  $\mathbf{s}$  atop a background odor  $\bar{\mathbf{s}}$  (Fig. 2b). Both odors are sparse mixtures, with  $K \ll N$  odorants of similar concentrations, odor “identity” being the particular set of odorants in the mixture. In the adaptive case, we assume that the system has fully adapted to the background  $\bar{\mathbf{s}}$  before the foreground  $\mathbf{s}$  is presented. This is enacted by calculating the firing response to the foreground odor  $\mathbf{r}(\mathbf{s})$  only after having set the  $\epsilon_a$  in Eq. 2 to their steady state values in response to the background odor  $\bar{\mathbf{s}}$ :

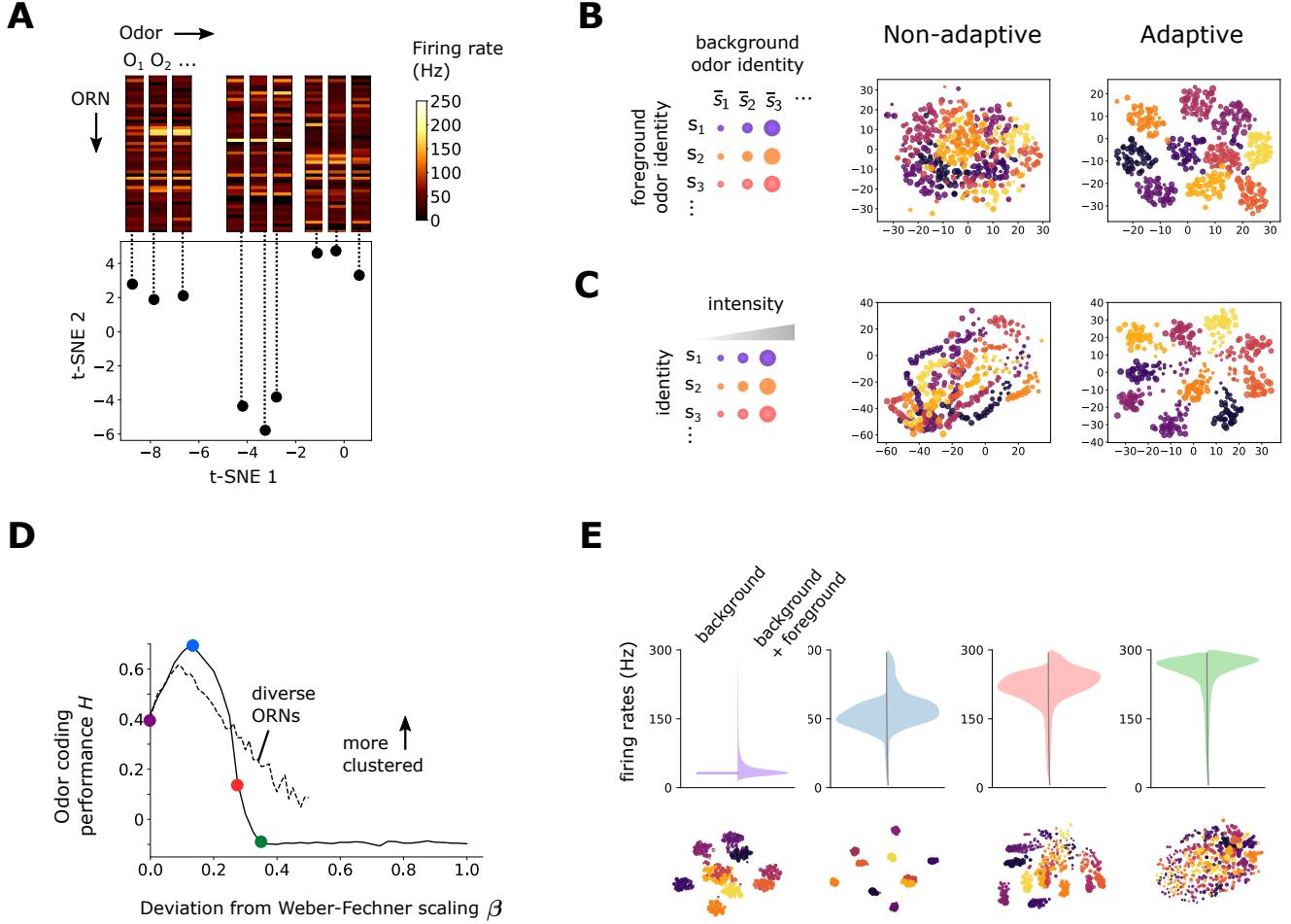
$$\epsilon_a(\bar{\mathbf{s}}) = \ln \left[ \frac{1 - A_{0a}}{A_{0a}} \right] - (1 - \beta_a) \ln \left( \frac{1 + \mathbf{K}_a \cdot \bar{\mathbf{s}}}{1 + \mathbf{K}_a^* \cdot \bar{\mathbf{s}}} \right), \quad (5)$$

where we have introduced the new parameter  $\beta_a$  to allow us to control the scaling of gain adaptation: for  $\beta_a = 0$  the system exactly follows Weber-Fechner's law, while for  $\beta_a = 1$  there is no adaptation. For small but nonzero  $\beta_a$ , the inverse gain scales sub-linearly (see Methods), and the adapted activity  $A_a(\bar{\mathbf{s}})$  increases weakly with background  $\bar{\mathbf{s}}$ . In experiments, small deviations from the strict Weber-Fechner scaling on the order of  $\beta \simeq 0.1$  are observed (see extended figures in [38]).

With Weber-Fechner's law in place for all ORNs ( $\beta_a = 0$ ) responses cluster by the identity of foreground odor, showing that the repertoire of ORNs appropriately encodes the identity of novel odors irrespective of background signals – once these backgrounds have been “adapted away” (Fig. 2b). This is the case regardless of whether  $A_{0a}$  is identical or different across neurons (Fig. 2-figure supplement 1). In contrast, when the system is non-adaptive, ( $\beta_a = 1$ ), the responses exhibit weaker separations by odor identity (Fig. 2b). Similarly, responses across different odor intensities are well separated by odor identity in the adaptive system, but less so in the non-adaptive system (Fig. 2c). Calculating the mutual information between odor and ORN response in time shows that the adaptive system retains coding capacity as it confronts novel odors (Fig. 2-figure supplement 2) whereas the non-adaptive system maintains coding capacity in a far more limited range of odor concentration.

To what extent do the benefits of front-end adaptation for odor coding depend on the precise Weber-Fechner scaling? We repeated the analysis from Fig. 2b for increasing values of  $\beta_a = \beta$  between zero (Weber's law) (perfect adaptation) and one (no adaptation). To generalize Fig. 2b, we now let the intensities range over two decades. As  $\beta$  increases, the capacity of the system to cluster responses by odor identity degrades (Fig. 2d). Introducing diversity among ORNs by distributing  $\beta_a$ 's uniformly between 0 and  $2\beta$  (so that the mean is  $\beta$ ) slightly increases performance at high  $\beta$  but reduces it at low  $\beta$  (Fig. 2d). Overall, performance of odor coding degrades with  $\beta$ , as poorly-adapting ORNs begin to saturate (Fig. 2e).

Interestingly, besides this general trend, we find that



**FIG. 2: Front-end adaptation maintains representations of odor identity across background and intensity confounds.**

**A** Example t-SNE projection of the 50-dimensional vector of ORN firing rates to 2 dimensions. Each point represents the firing response to a distinct odor. Nearby points exhibit similarities in corresponding firing rates. **B** t-SNE projection of ORN firing rates, where each point represents the response to foreground odor  $s$  (point color) on top of a background odor  $\bar{s}$  (point size). In the adaptive system,  $\epsilon_a$  are set to their steady state values given the background odor  $\bar{s}$  alone according to Eq. 5 with  $\beta = 0$ . We assumed  $A_{0a} = A_0$  for all  $a$  (we obtain similar results when  $A_{0a}$  are randomly distributed; Fig. 2—figure supplement 1). Clustering by color implies that responses cluster by foreground odor identity. **C** Similar to **B**, but now for odors whose concentrations span 4 decades (represented by point size). Here, the background odor identity is the same for all concentrations. **D** Performance of odor coding as a function of  $\beta$ , the magnitude of the deviation from Weber-Fechner's law ( $\beta = 0$ : Weber-Fechner's scaling;  $\beta = 1$ : no adaptation; see Eq. 5). Performance is quantified by the silhouette score  $H$ , which quantifies the degree by which responses cluster by foreground identity versus background identity in the t-SNE projections (1: highly clustered by foreground identity; 0 or slightly negative: not clustered) ([50] and Methods). Line: same scaling  $\beta_a = \beta$  for all ORNs. Dashed:  $\beta_a$  is uniformly distributed between 0 and  $2\beta < 1$  (i.e. has mean  $\beta$ ). **E** Distribution of ORN responses and t-SNE projections for  $\beta = 0, 0.13, 0.28, 0.35$  in **D**.

for  $\beta$  very close to zero, a small deviation from Weber-Fechner's law instead *improves* odor coding. This arises because of the nonlinearity in the onset of adaptation: adaptation kicks in only when the strength of stimulus is sufficient for the response  $A_a$  to exceed  $A_{0a}$ , so that the right hand-side of Eq. 3 is positive. The minimum background intensity  $\bar{s}$  required for this to happen is given by  $\epsilon_{L,a} = \epsilon_a(\bar{s})$ , which, according to equation Eq. 5, increases with  $\beta$ . This initial effect increases odor coding performance, as the firing rates can distribute more broadly across the dynamical range of the ORNs, before adaptation is effected (Fig. 2e).

Thus, while Weber-Fechner's Law scaling largely preserves the representation of foreground odor identity

276 amid backgrounds, in some cases it may benefit from a  
 277 slight relaxation so that the full dynamical range of the  
 278 ORNs can be exploited.

279 **Front-end adaptation enhances odor decoding in  
 280 complex environments**

281 How well does the preservation of odor coding translate  
 282 to better signal reconstruction from ORNs responses?  
 283 One potentially complicating factor is the disparity be-  
 284 tween sensor dimension and stimulus dimension: while  
 285 *Drosophila* only express  $\sim 60$  Or genes [52], the space  
 286 of odorants is far greater [27]. An  $N$ -dimensional odor  
 287 signal would naively need  $N$  sensory neurons to decode  
 288 it – one for each odorant. However, naturally-occurring  
 289 odors are sparse, typically comprised of only a few odor-  
 290 ants. Enforcing sparsity of the signal during decoding  
 291 greatly restricts the number of possible odors consistent  
 292 with a given ORN response, suggesting that such high-  
 293 dimensional signals might be inferred from less than  $N$   
 294 ORNs. Indeed, the decoding of sufficiently sparse signals  
 295 from lower-dimensional responses is rigorously guaran-  
 296 teed by the theory of compressed sensing (CS) [53, 54]. It  
 297 is unknown whether CS is implemented in the *Drosophila*  
 298 olfactory circuit [55]. Here we use this framework mainly  
 299 as a tool to quantify how front-end adaptation poten-  
 300 tially affects odor decoding, later verifying our conclu-  
 301 sions with other classification techniques that incorporate  
 302 the known architecture of the olfactory system.

303 In practice, CS is performed as a constrained linear  
 304 optimization over the components of the signal vector –  
 305 here, the odorant concentrations  $s_i$ . The constraints in  
 306 the optimization are  $\mathbf{r} = \mathbf{Ds}$ , where  $\mathbf{D}$  is a matrix and  
 307  $\mathbf{D}$  are the measurements – here,  $\mathbf{r}$  would be the vector  
 308 of ORN responses. The cost function to be minimized,  
 309  $C = \sum_i |s_i|$ , enforces sparsity by driving the estimate of  
 310 each odorant component to zero; the constraints balance  
 311 this tendency by simultaneously enforcing information  
 312 from the ORN firing rates. The result is a reconstructed  
 313 odor signal  $\hat{\mathbf{s}}$  that is as sparse as possible, consistent with  
 314 the ORN responses  $\mathbf{r}$ .

315 To incorporate the linear framework of CS into our  
 316 nonlinear odor encoding model, we treat the nonlinear  
 317 odor encoding exactly, but approximate the decoding to  
 318 first order around the background concentration. Specif-  
 319 ically, we use Eqs. 2-4 to generate ORN responses  $\mathbf{r}$   
 320 for sparse odors  $\mathbf{s}$  having  $K \ll N$  nonzero components  
 321  $s_i = \bar{s}_i + \Delta s_i$ , where the mean concentration is  $\bar{s}_i$ . To re-  
 322 construct these signals using CS, we minimize  $\sum_i |\Delta s_i|$   
 323 while enforcing the constraints  $\mathbf{r} = \mathbf{D}\Delta\mathbf{s}$ , where  $\mathbf{D}$  is  
 324 the linearization of Eq. 2 around  $\bar{s}_i$  (details in Methods).  
 325 This linearization simplifies the CS decoding – namely it  
 326 enforces a single, global minimum – but it is not critical  
 327 for our general results; see Methods and Fig. 3-figure sup-  
 328 plement 5. The matrix  $\mathbf{D}$  depends on  $\epsilon_a$ , and as above,

329 we assume precise adaptation by setting  $\epsilon_a$  to their steady  
 330 state values in response to the background odor alone (via  
 331 Eq. 5 with  $\beta = 0$ ). In the nonadaptive case,  $\epsilon_a$  are held  
 332 at their minimum values  $\epsilon_{L,a}$ .

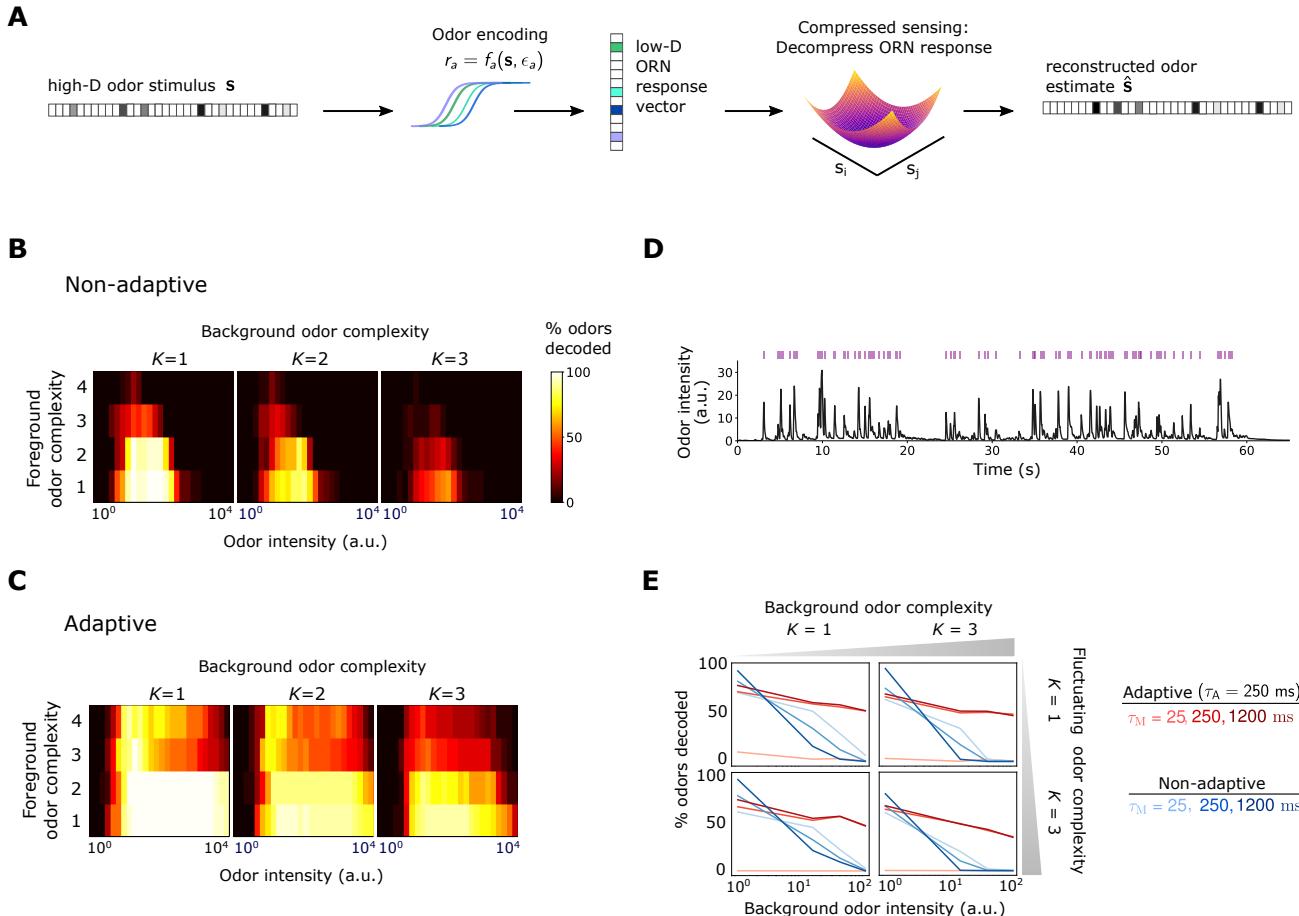
333 We first examine how foreground odors are recognized  
 334 when mixed with background odors of a distinct identity  
 335 but similar intensities, quantifying decoding accuracy as  
 336 the percentage of odors correctly decoded within some  
 337 tolerance (see Methods). Without adaptation, accuracy  
 338 is maintained within the range of receptor sensitivity for  
 339 monomolecular backgrounds, but is virtually eliminated  
 340 as background complexity rises (Fig. 3b). The range  
 341 of sensitivity is broader in the adaptive system, and is  
 342 substantially more robust across odor concentration and  
 343 complexity (Fig. 3c).

344 In realistic odor environments, the concentration and  
 345 duration of individual odor whiffs vary widely [18]. We  
 346 wondered how well a front-end adaptation mechanism  
 347 with a single timescale  $\tau$  could promote odor identity  
 348 detection in such environments. As inputs to our cod-  
 349 ing/decoding framework, we apply a naturalistic stimulus  
 350 intensity recorded using a photo-ionization detector [38]  
 351 (Fig. 3c) to which we randomly assign sparse identities  
 352 from the  $N$ -dimensional odorant space. To mimic back-  
 353 ground confounds, we combine these signals with static  
 354 odor backgrounds, and then calculate the percentage of  
 355 decoded whiffs. We assume the decoder has short-term  
 356 memory: detected odor signals are only retained for  $\tau_M$   
 357 seconds in the immediate past, bounding the amount of  
 358 past information utilized in signal reconstruction.

359 Without ORN adaptation, sufficiently strong back-  
 360 grounds eliminate the ability to reconstruct the identity  
 361 of individual odor whiffs, irrespective of the complexity  
 362 of either the foreground or background odor (Fig. 3d,  
 363 blue lines). In the adaptive system, this is substantially  
 364 mitigated (red lines in Fig. 3d), provided the memory du-  
 365 ration  $\tau_M$  is at least as long as the adaptation timescale  
 366  $\tau$  (darker red lines). Because this short-term adaptation  
 367 depends on the activity of the Or-Orco channel rather  
 368 than on the identity of the receptor [37, 38, 43], the val-  
 369 ues of  $\tau$  and  $A_0$  were assumed the same for all ORNs; still,  
 370 our results hold if these invariances are relaxed (Fig. 3-  
 371 figure supplement 1-Fig. 3-figure supplement 2).

372 **Front-end adaptation enhances primacy coding**

373 The primacy coding hypothesis has recently emerged  
 374 as an intriguing framework for combinatorial odor  
 375 coding. Here, odor identity is encoded by the set  
 376 (but not temporal order) of the  $p$  earliest responding  
 377 glomeruli/ORN types, known as primacy set of order  
 378  $p$  [32]. If the activation order of ORNs were invari-  
 379 ant to the strength of an odor step or pulse, primacy  
 380 sets would in principle form concentration-invariant rep-  
 381 resentation of odor identity. Though our coding frame-

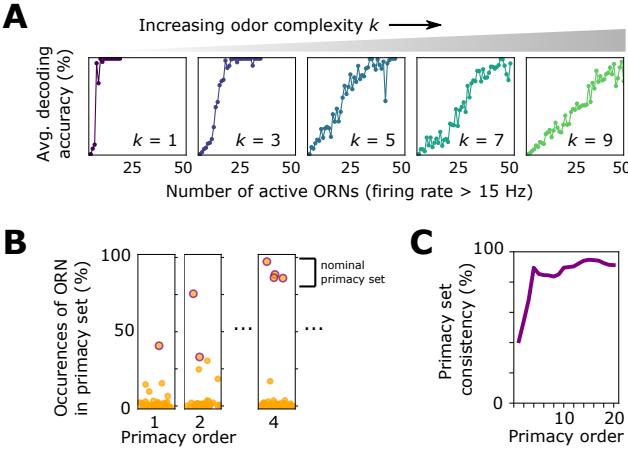


**FIG. 3: Front-end adaptation promotes accurate odor decoding in static and naturalistic odor environments.** **A** Odor stimuli produce ORN responses via odor-binding and activation and firing machinery, as described by Eqs. 2-4. Odors are then decoded using compressed sensing optimization. Odors are assumed sparse, with  $K$  nonzero components,  $K \ll N$ . **B** Decoding accuracy of foreground odors in the presence of background odors, for a system without Weber Law adaptation. **C** Same as **B**, with Weber Law adaptation. **D** Recorded trace of naturalistic odor signal; whiffs (signal  $> 4$  a.u.) demarcated by purple bars. This signal is added to static backgrounds of different intensities and complexities. **E** Individual plots show the percent of accurately decoded odor whiffs as a function of background odor intensity, for the non-adaptive (blue) and adaptive (red) systems, for different  $\tau_M$  (line shades).

work uses the full ORN ensemble in signal reconstruction, some of these responses may contain redundant information, and a smaller primacy subset may suffice. To examine this, we apply our model to a sigmoidal stimulus that rises to half-max in 50 ms, calculating decoding accuracy in time. Since ORNs activate sequentially, the primacy set is defined by the ORN subset active when the odor is decoded. For simple odors, a limited set of earliest responding neurons fully accounts for the odor identity (Fig. 4a), in agreement with primacy coding. As expected for more complex odor mixtures, the full repertoire is required for accurate decoding. Primacy coding also predicts that for stronger stimuli, responses occur earlier, since the primacy set is realized quicker, which our framework replicates (Fig. 4—figure supplement 1).

Beyond mere consistency, however, front-end adapta-

tion might also enhance primacy coding in different environments, such as background odors, which could scramble primacy sets. To investigate this, we considered again a sigmoidal odor step (odor A), now atop a static background (odor B) to which the system has adapted. We compared the primacy sets of odor A for 1000 different choices of odor B, finding that primacy sets are highly consistent across background confounds for all but the smallest primacy orders (Fig. 4b-4c). This also holds true for backgrounds of different concentrations (Fig. 4—figure supplement 1), suggesting a central role for front-end adaptation in reinforcing primacy codes across differing environmental conditions.



**FIG. 4: Effect of front-end adaptation on primacy coding.** **A** Decoding accuracy as a function of the number of active ORNs, for different odor complexities. The primacy set consists of those ORNs required to be active for accurate decoding. **B** Frequency of particular ORNs in primacy sets of an odor placed atop different backgrounds. Individual plots show, for given primacy order  $p$ , the percentage of backgrounds for which the primacy set of odor A contains a given ORN (dots). Those with purple borders are the  $p$  most highly occurring – i.e. a nominal background-invariant primacy set for odor A. Points are jittered horizontally for visualization. **C** Consistency of primacy sets across backgrounds, as a function of  $p$ . Consistency is defined as the likelihood that an ORN in the nominal primacy set appears in any of the individual background-dependent primacy sets, averaged over the nominal set (average of the y-values of the purple dots in B). 100% consistency means that for all backgrounds, the primacy set of odor A is always the same  $p$  ORNs.

#### 411 Contribution of front-end adaptation for odor 412 recognition within the *Drosophila* olfactory circuit

413 Signal transformations in the sensing periphery are  
414 propagated through the remainder of the olfactory cir-  
415 cuit. How does front-end adaptation interact with  
416 these subsequent neural transformations? ORNs express-  
417 ing the same OR converge to a unique AL glomeru-  
418 lus, where they receive lateral inhibition from other  
419 glomeruli [23, 56]. This inhibition implements a type of  
420 divisive gain control [24], normalizing the activity of out-  
421 put projections neurons, which then synapse onto a large  
422 number of Kenyon cells (KCs) in the mushroom body.  
423 To investigate how odor representations are affected by  
424 interactions between front-end ORN adaptation and this  
425 lateral inhibition and synaptic divergence, we extended  
426 our ORN encoding model by adding uniglomerular con-  
427 nections from ORNs to the antennal lobe, followed by  
428 sparse, divergent connections to 2500 KCs [25, 26, 57].  
429 Inhibition was modeled via divisive normalization, with  
430 parameters chosen according to experiment [24]. We  
431 quantified decoding accuracy by training and testing a  
432 binary classifier on the KC activity output of sparse odors

433 of distinct intensity and identity, randomly categorized as  
434 appetitive or aversive. For simplicity, odor signals of the  
435 same identity but differing intensity were assigned the  
436 same valence. We trained the classifier on  $N_{ID}$  sparse  
437 odor identities at intensities chosen randomly over 4 or-  
438 ders of magnitude, then tested the classifier accuracy on  
439 the same odor identities but of differing concentrations.

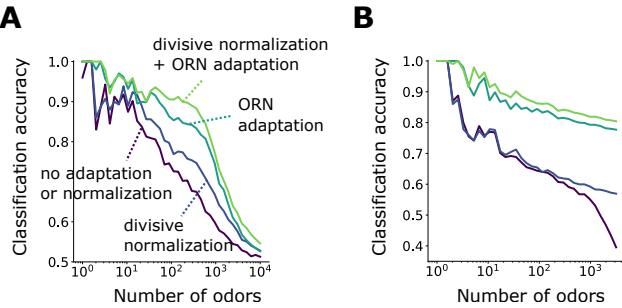
440 Classification accuracy degrades to chance level as  $N_{ID}$   
441 becomes very large (Fig. 5a). When acting alone, either  
442 divisive normalization or ORN adaptation can help, al-  
443 though the effect of ORN adaptation is stronger. When  
444 both are active, accuracy improves further, suggesting  
445 that these distinct adaptive transformations may act  
446 jointly at different stages of neural processing in preserv-  
447 ing representations of odor identity. As expected, these  
448 gains mostly vanish for the same odors chosen from a nar-  
449 rower range of concentrations (Fig. 5—figure supplement  
450 1).

451 If we train the classifier to distinguish odors by iden-  
452 tity rather than valence, the benefits conferred by divi-  
453 sive normalization do not appear until  $N_{ID}$  is substantial,  
454 with accuracy below 65% for  $N_{ID} > 50$  (Fig. 5b). On  
455 the other hand, with ORN adaptation, accuracy remains  
456 above 85% for more than 1000 odor identities, strongly  
457 implicating front-end adaptation as a key player in main-  
458 taining odor identity representations, before signals are  
459 further processed downstream.

460 We note that previous simulation results have shown  
461 that divisive normalization aids identity decoding from  
462 PN response to a stronger degree than we find here [24].  
463 The discrepancy may arise from the differences in odor  
464 classification from PN responses versus KC responses. It  
465 likely also arises from the fact that we are decoding odor  
466 mixtures rather than pure odorants, so the combinatorics  
467 may play a larger role. Finally, the divisive normaliza-  
468 tion model is a simple one in which glomeruli are all  
469 mutually inhibiting. A more complex model in which  
470 each glomerulus inhibits only a subset of other glomeruli  
471 through local neurons might produce a larger contribu-  
472 tion.

## DISCUSSION

473  
474 Weber-Law adaptation at the very front-end of the in-  
475 sect olfactory circuit [38, 41, 42] may contribute signif-  
476 icantly to the preservation of neural representations of  
477 odor identity amid confounding odors and intensity fluc-  
478 tuations. Drawing on experimental evidence for a num-  
479 ber of ORN-invariant response features [20, 22, 37, 38,  
480 43], we have found that this mechanism of dynamic adap-  
481 tation confers significant benefits in coding fidelity, with-  
482 out the need for ORN-specific parameterizations. Still,  
483 our results hold when these invariances such as adapta-  
484 tion timescale or baseline activity are relaxed (Fig. 3—  
485 figure supplement 1—Fig. 3—figure supplement 2). In the



**FIG. 5: Front-end adaptation enhances odor recognition by the *Drosophila* olfactory circuit.** **A** Accuracy of binary classification by odor valence, as a function of the number of distinct odor identities classified by the trained network (concentrations span 4 orders of magnitude), in systems with only ORN adaptation, only divisive normalization, both or neither. **B** Same as (A) but now classifying odors by identity.

olfactory periphery, front-end Weber Law adaptation therefore appears fairly robust, a consequence of controlling gain via feedback from channel activity [38, 43, 48], rather than through intrinsic, receptor-dependent mechanisms. Our results also suggest that a slight breaking of Weber scaling may aid combinatorial coding, by spreading firing rates more fully over the ORN dynamic range, while still preventing saturation. The degree of this breaking would manifest as a correction to the Weber scaling exponent,  $\sim (1/s)^1 \rightarrow \sim (1/s)^{1-\beta}$ , which could in principle be measured experimentally for individual ORNs. Such small deviations from the strict Weber-Fechner scaling have been observed (see extended Figures in [38]).

While our framework incorporates many observed features of the *Drosophila* olfactory system – Weber-Law adaptation, power-law distributed receptor affinities, temporal filter invariance, connectivity topologies – it is minimal. We considered only one of the chemoreceptor families expressed in the fly antenna [1] and ignored possible contributions of odor binding proteins [58, 59], inhibitory odorants [21], and odorant-odorant antagonism [60], which could further boost coding capacity and preserve representation sparsity. Useful extensions to our nonlinear-linear-nonlinear model might incorporate ephaptic coupling between ORNs housed in the same sensillum [61], global inhibition in the mushroom body [62], and the effects of long-term adaptation [46].

Previous studies have characterized various neural mechanisms that help preserve combinatorial codes. Lateral inhibition between glomeruli helps tame saturation and boost weak signals [24]. The sparse degree of connectivity to either the olfactory bulb (vertebrates) or mushroom body (insects) may also be precisely tuned to optimize the capacity to learn associations [26]. In this work, we find that some of these downstream features act in

concert with front-end dynamic adaptation in maintaining representations of odor identity.

Other studies have implicated the unique temporal patterns of neural response as signatures of odor identity [29–31, 63]. ORN and projection neuron time traces form distinct trajectories in low-dimensional projections, and cluster by odor identity, much as we have found here for static responses at different concentrations (Fig. 2). In locusts PNs, the trajectories elicited by foreground odors when presented in distinct backgrounds exhibit some degree of overlap; though partial, these overlaps were nonetheless sufficient to maintain background-invariant decoding from Kenyon cell responses [14]. It was therefore suggested that background filtering likely occurs at the level of ORNs themselves [14]. Likewise, in our framework, temporal coding is implicit: because the input nonlinearity depends on the diversity of binding affinities, odor signals are naturally formatted into temporal patterns that are both odor- and ORN-specific (Figs. 1d-1e). Further, the short required memory timescales ( $\tau_M \sim \tau \sim 250$  ms) suggest that only brief time windows are needed for accurate odor identification, consistent with previous findings [14, 29]. Moreover, we find that front-end adaptation enhances the robustness of other combinatorial coding schemes, such as primacy coding [32], which relies on the temporal order of ORN activation but not absolute firing rate (Fig. 4).

In the well-characterized chemosensory system of bacterial chemotaxis, Weber Law adaptation is enacted through a feedback loop from the output activity of the receptor-kinase complexes onto the enzymes modifying receptor sensitivity [48]. It is interesting that some aspects of this logic are also present in ORNs: although the molecular players are different (and still largely unknown, though likely involving calcium channel signaling [42]), it has been shown that transduction activity feeds back onto the sensitivity of Or-Orco ligand-gated cation channels, enabling the Weber-Fechner relation [38, 42, 43]. That this adaptation mechanism appears to act similarly across ORNs [37, 38, 42] suggests the possible involvement of the universal co-receptor Orco, whose role in long-term adaptation has recently been reported [44–46]. Further, the identification of 4 subunits comprising the Orco-Or ion channel suggest that generic Or/Orco complexes may contain multiple odorant binding sites, which when included in our model supports our general findings (Fig 3—figure supplement 3).

Weber Law ensures that sensory systems remain in the regime of maximum sensitivity, broadening dynamic range and maintaining information capacity [64]. For a single-channel system, this requires matching the midpoint of the dose-response curve to the mean ligand concentration [65], a strategy which may fail in multi-channel systems with overlapping tuning curves: adaptation to one signal could inhibit identification of others, if the signals excite some but not all of the same sen-

578 sors, as in Fig. 1g. Our results show that this strategy  
 579 is still largely functional. In CS decoding, this can be  
 580 traced to the observation that accuracy is guaranteed  
 581 when sufficiently distinct odor identities produce suffi-  
 582 ciently distinct ORN responses, a condition known as  
 583 the restricted isometry property [54]. Indeed, the Weber-  
 584 Fechner scaling increases the likelihood that this property  
 585 is satisfied, beyond that in the non-adaptive system (SI  
 586 text and Fig. 3-figure supplement 4-Fig. 3-figure sup-  
 587 plement 5). Still, restricted isometry does not require  
 588 that response repertoires are *invariant* to environmental  
 589 changes. That is, even if the subset of active ORNs were  
 590 concentration-dependent, odors could still in principle be  
 591 fully reconstructible by CS. Nonetheless, our results in  
 592 t-SNE clustering (Fig. 2), primacy coding (Fig. 4b-4c),  
 593 and odor classification (Fig. 5) suggest that some signa-  
 594 ture of response invariance emerges as a natural byprod-  
 595 uct of front-end adaptation. Together, this implies that  
 596 Weber Law adaptation, whether required by the olfac-  
 597 tory circuit for precise signal reconstruction (as in CS)  
 598 or for developing odor associations (as in classification),  
 599 can play an integral part in maintaining combinatorial  
 600 codes amid changing environmental conditions.

## 601 ACKNOWLEDGEMENTS

602 NK was supported by a postdoctoral fellowship  
 603 through the Swartz Foundation and through an NRSA  
 604 postdoctoral fellowship through the NIH BRAIN Initia-  
 605 tive under award number 1F32MH118700. TE was sup-  
 606 ported by NIH R01 GM106189. We thank Damon Clark,  
 607 John Carlson, Mahmut Demir, Srinivas Gorur-Shandilya,  
 608 Henry Mattingly, and Ann Hermunstad for comments on  
 609 the manuscript.

## 610 AUTHOR CONTRIBUTIONS

611 N.K. and T.E. conceived the project, analyzed the re-  
 612 sults, and wrote the manuscript. N.K. performed all cal-  
 613 culations. T.E. supervised the project.

## 614 DECLARATION OF INTERESTS

615 The authors declare no competing interests.

## 616 METHODS

### 617 Adaptive ORN model

618 We model an odor as an  $N$ -dimensional vector  
 619  $\mathbf{s} = [s_1, \dots, s_N]$ , where  $s_i > 0$  are the concentrations of  
 620 individual volatile molecules (odorants) comprising the

622 odor. The olfactory sensory system is modeled as a col-  
 623 lection of  $M$  distinct Or/Orco complexes indexed by the  
 624 sub index  $a = 1, \dots, M$ , each of which can be bound with  
 625 any one of the odorant molecules, and can be either active  
 626 (firing) or inactive (quiescent). At first we assume there  
 627 is one binding site per complex; this will be generalized  
 628 to many sites. We consider the binding and activation  
 629 processes to be in equilibrium, assigning each state a cor-  
 630 responding Boltzmann weight, where the zero of energy  
 631 is set by the unbound, inactive state  $C_a$ . These weights  
 632 are:

$$\begin{aligned} C_a &= 1 \\ C_a^* &= \exp(-\beta\epsilon_a) \\ C_a:s_i &= \exp(-\beta(-E_{ai} - \mu_i)) \\ C_a^*:s_i &= \exp(-\beta(-(E_{ai}^* - \epsilon_a) - \mu_i)), \end{aligned} \quad (6)$$

633 where  $\epsilon_a$  (assumed positive) is the free energy differ-  
 634 ence between the active and inactive conformation of  
 635 the unbound receptor, and  $E_{ai}$  and  $E_{ai}^*$  are the free  
 636 energy differences (assumed positive) between the un-  
 637 bound and bound state for the inactive and active re-  
 638 ceptor, respectively.  $\mu_i = \mu_0 + \beta^{-1} \log(s_i/s_0)$  is the  
 639 chemical potential for odorant species  $i$  in terms of  
 640 a reference chemical potential  $\mu_0$  at concentration  $s_0$ ,  
 641  $s_0 \exp(-\beta\mu_0) = s_i \exp(-\beta\mu_i)$ , which can be traded  
 642 for the thermodynamic-relevant disassociation constants  
 643  $K_{ai}^{-1} = K_{D,ai} = s_0 e^{\beta(-E_{ai} - \mu_0)}$ .

Adding up contributions from all  $i$  odorants, the active  
 fraction is:

$$\begin{aligned} A_a &= \frac{C_a^* + \sum_i C_a^*:s_i}{C_a^* + \sum_i C_a^*:s_i + C_a + \sum_i C_a:s_i}^{-1} \\ &= \left( 1 + \frac{C_a + \sum_i C_a:s_i}{C_a^* + \sum_i C_a^*:s_i} \right)^{-1} \\ &= \left( 1 + e^{\epsilon_a} \frac{1 + \mathbf{K}_a \cdot \mathbf{s}(t)}{1 + \mathbf{K}_a^* \cdot \mathbf{s}(t)} \right)^{-1}, \end{aligned} \quad (2)$$

644 where we have expressed free energies in units of  $k_B T =$   
 645  $\beta^{-1}$  for notational convenience.

646 This expression can be generalized for the case of mul-  
 647 tiple, independent binding sites through some simple  
 648 combinatorial factors. Consider first an odorant  $i$  which  
 649 can bind one of two locations on receptor  $a$ . There are  
 650 then 4 possible inactive states: both sites unbound, site  
 651 1 bound, site 2 bound, both sites bound. Combined with  
 652 the active states, there are therefore 8 states for odorant  
 653  $i$  and receptor  $a$ , with energies:

$$\begin{aligned}
\text{active} \quad & \{1, -E_{ai} - \mu_i, \\
& \quad -E_{ai} - \mu_i, -2E_{ai} - 2\mu_i\} \\
\text{inactive} \quad & \{\epsilon_a, -(E_{ai}^* - \epsilon_a) - \mu_i, \\
& \quad -(E_{ai}^* - \epsilon_a) - \mu_i, \\
& \quad -(2E_{ai}^* - \epsilon_a) - 2\mu_i\}
\end{aligned} \tag{7}$$

In the active fraction, Eq. 2, the Boltzmann factors combine through the binomial theorem, giving (for a single odorant environment  $i$ ):

$$\begin{aligned}
A_a(\text{odorant } i, 2 \text{ binding sites}) \\
= \left[ 1 + e^{\epsilon_a} \left( \frac{1 + \mathbf{K}_a \cdot \mathbf{s}(t)}{1 + \mathbf{K}_a^* \cdot \mathbf{s}(t)} \right)^2 \right]^{-1}.
\end{aligned} \tag{8}$$

This expression generalizes for an arbitrary number of odorants and independent binding sites through the appropriate combinatorial factors, giving an active fraction of

$$\begin{aligned}
A_a(N \text{ odorants}, R \text{ binding sites}) \\
= \left[ 1 + e^{\epsilon_a} \left( \frac{1 + \mathbf{K}_a \cdot \mathbf{s}(t)}{1 + \mathbf{K}_a^* \cdot \mathbf{s}(t)} \right)^R \right]^{-1}.
\end{aligned} \tag{9}$$

To generate ORN time traces, equations 2-3 are integrated numerically using the Euler method with a 2 ms time step. For ORN firing (Eq. 4),  $h(t)$  is bi-lobed [37]:  $h(t) = Ap_{\text{Gam}}(t; \alpha_1, \tau_1) - Bp_{\text{Gam}}(t; \alpha_2, \tau_2)$ ,  $A = 190$ ,  $B = 1.33$ ,  $\alpha_1 = 2$ ,  $\alpha_2 = 3$ ,  $\tau_1 = 0.012$ , and  $\tau_2 = 0.016$ , where  $p_{\text{Gam}}$  is the pdf of  $\text{Gamma}(\alpha, 1/\tau)$ . Nonlinearity  $f$  is modeled as a linear rectifier with 5 Hz threshold.

#### Derivation of ORN gain

Weber's Law states that the gain, or differential response, of the receptor activity  $A_a$  scales with the mean odor concentration  $\bar{s}_i$ . To show how this is satisfied in our model, we consider the response, Eq. 2, to a signal  $\mathbf{s} = \bar{\mathbf{s}} + \Delta\mathbf{s}$ , where  $\Delta\mathbf{s}$  consists of only a small fluctuation in the  $i$ th component  $\Delta s_i < |\bar{s}_i|$  about the mean. We derive the change in response to fluctuation  $\Delta s_i$  for general  $\beta$  from 0 (Weber's law) to 1 (no adaptation).

First we write the activity in the form:

$$A_a = (1 + e^{F_a})^{-1}, \tag{10}$$

where

$$F_a = \epsilon_a(\bar{\mathbf{s}}) + \ln \left( \frac{1 + \mathbf{K}_a \cdot \mathbf{s}}{1 + \mathbf{K}_a^* \cdot \mathbf{s}} \right). \tag{11}$$

where  $\epsilon_a(\bar{\mathbf{s}})$  is given by Eq. 5. Then, assuming  $1/\mathbf{K}_a^* \ll s_i \ll 1/\mathbf{K}_a$ , the change in response from the adapted level  $A_a(\bar{\mathbf{s}})$  is

$$\begin{aligned}
A_a(\mathbf{s}) - A_a(\bar{\mathbf{s}}) &= \Delta A_a = \frac{dA_a}{dF_a} \frac{dF_a}{ds} \Big|_{\bar{\mathbf{s}}} \Delta s_i \\
&= -\frac{e^{F_a}}{(1 + e^{F_a})^2} \Big|_{\bar{\mathbf{s}}} \left( \frac{-K_{ai}^*}{\mathbf{K}_a^* \cdot \bar{\mathbf{s}}} \right) \Delta s_i.
\end{aligned} \tag{12}$$

We use Eq. 5 to evaluate  $e^{F_a}$  at  $\bar{\mathbf{s}}$ , obtaining:

$$e^{F_a} \approx \frac{1 - A_{0a}}{A_{0a}} (\mathbf{K}_a^* \cdot \bar{\mathbf{s}})^{-\beta}, \tag{13}$$

whereby

$$\begin{aligned}
\frac{\Delta A_a}{\Delta s_i} &= \frac{\frac{1-A_{0a}}{A_{0a}} (\mathbf{K}_a^* \cdot \bar{\mathbf{s}})^{-\beta}}{(1 + \frac{1-A_{0a}}{A_{0a}} (\mathbf{K}_a^* \cdot \bar{\mathbf{s}})^{-\beta})^2} \left( \frac{K_{ai}^*}{\mathbf{K}_a^* \cdot \bar{\mathbf{s}}} \right) \\
&= \frac{(1 - A_{0a}) A_{0a} K_{ai}^*}{[A_{0a} (\mathbf{K}_a^* \cdot \bar{\mathbf{s}})^{\frac{1+\beta}{2}} + (1 - A_{0a}) (\mathbf{K}_a^* \cdot \bar{\mathbf{s}})^{\frac{1-\beta}{2}}]^2}.
\end{aligned} \tag{14}$$

For  $\beta = 0$  (the fully adaptive case) and a single odorant, this expression for the gain reduces to  $(1 - A_{0a}) A_{0a} / s_i$ . For small  $\beta$ , and given  $A_{0a} \simeq 0.1$  (corresponding to 30 Hz on a 300 Hz firing rate scale), the denominator is dominated by the  $1 - A_{0a}$  term, giving:

$$\frac{\Delta A_a}{\Delta s_i} \Big|_{(\beta \ll 1)} = \frac{A_{0a} K_{ai}^*}{(1 - A_{0a}) (\mathbf{K}_a^* \cdot \bar{\mathbf{s}})^{1-\beta}}. \tag{15}$$

The implication of this is that the gain scaling of the inverse mean intensity, which is 1 for perfect adaptation ( $\text{gain} \sim (1/s_i)^1$ ), is now sublinear. Thus, when Weber's Law is weakly broken, the gain still reduces with mean odor intensity, but not as quickly.

#### t-SNE dimensionality reduction

For t-SNE dimensionality reduction [51], ORN responses were generated for odor signal combinations consisting of 1 (among 10) distinct sparse foreground odors A atop 1 (among 50) distinct sparse background odors B, for Fig. 2b. Fig. 2c plots responses for 10 odors at 40 concentrations spanning 4 decades, atop a random sparse background odor of similar magnitude. For adaptive systems,  $\epsilon_a$  were set to their fully adapted values to the background odor, given by Eq. 5, with  $\beta = 0$ . For Figs. 2d, for each  $\beta$ , we averaged the silhouette score  $H$  (defined precisely below) over 5 different trials, each trial with a different seed for randomly choosing the odor identities.

To calculate performance of clustering in the t-SNE plots, we use the silhouette score  $H$  [50], which is a function of the 2D Euclidean distance  $d(i, j|m, n)$  between the point  $p_{ij}$  representing odor  $s_i$  on background  $b_j$  and the point  $p_{mn}$  representing odor  $s_m$  on background  $b_n$ . For each  $p_{ij}$ , we define  $a(ij)$  as:

$$a(ij) = \frac{1}{1 - N_i} \sum_{n \in C_i} d(i, j|i, n) \quad (16)$$

where  $C_i$  is the set of all points in cluster  $i$ , i.e. with foreground  $i$ . This quantifies the average distance of  $p_{ij}$  to other points in the same cluster, i.e. with the same foreground but different backgrounds. We next define  $b(ij)$  as:

$$b(ij) = \min_{m \neq i} \frac{1}{N_m} \sum_{n \in C_m} d(i, j|m, n), \quad (17)$$

where the sum is taken over all points  $p_{mn}$  in cluster  $m$  – i.e. with foreground  $m$  not equal to  $i$  – and the minimum is taken over all clusters  $m$ . This quantifies the average distance from  $p_{ij}$  to the points in its “nearest neighbor” cluster. Then,  $H$  is defined as:

$$H = \langle \frac{a(ij) - b(ij)}{\max\{a(ij), b(ij)\}} \rangle_{ij}, \quad (18)$$

where the average is taken over all points  $ij$  in the t-SNE plot.  $H$  is nearer to 1 when points cluster by foreground, and is near zero or slightly negative when the clusters cannot be disambiguated.

### Compressed sensing decoding of ORN responses

Compressed sensing (CS) addresses the problem of determining a sparse signal from a set of linear measurements, when the number of measurements is less than the signal dimension. Specifically, it is a solution to

$$\mathbf{y} = \mathbf{D}\mathbf{x}, \quad (19)$$

where  $\mathbf{x} \in \mathbb{R}^N$  and  $\mathbf{y} \in \mathbb{R}^M$  are vectors of signals and responses, respectively, and  $\mathbf{D}$  is the measurement matrix. Since measurements are fewer than signal components, then  $M < N$ , whereby  $\mathbf{D}$  is wide rectangular and so Eq. 19 cannot be simply inverted to produce  $\mathbf{x}$ . The idea of CS is to utilize the knowledge that  $\mathbf{x}$  is sparse, i.e. only  $K$  of its components,  $K \ll N$  are nonzero. Both the measurements and sparsity are thus combined into a single constrained optimization routine:

$$\hat{x}_i = \operatorname{argmin}_i \sum_i^N |x_i| \quad \text{such that } \mathbf{y} = \mathbf{D}\mathbf{s} \quad (20)$$

where  $\hat{x}_i$  are the optimal estimates of the signal components and the sum, which is known as the  $L_1$  norm of  $\mathbf{x}$ , is a natural metric of sparsity [53].

The  $L_1$  norm is a convex operation and the constraints are linear, so the optimization has a unique global minimum. To incorporate the nonlinear response of our encoding model into this linear framework, we assume that the responses are generated through the full nonlinear steady state response, Eq. 2- 4, but that the measurement matrix  $\mathbf{D}$  needed for decoding uses a linear approximation of this transformation. Expanding Eq. 4 around  $\bar{\mathbf{s}} = \mathbf{s} - \Delta\mathbf{s}$  gives

$$\begin{aligned} \Delta r_a(t) &= r_a(\mathbf{s}(t)) - r_a(\bar{\mathbf{s}}(t)) \\ \Delta r_a(t) &= \int^t d\tau h(t - \tau) \sum_i^N \frac{dA_{ai}}{ds_i} \Big|_{\bar{\mathbf{s}}} \Delta s_i \end{aligned} \quad (21)$$

where

$$r_a(\mathbf{s}_0) = \int^t d\tau h(t - \tau) \sum_i^N A_{0a} \quad (22)$$

and where  $\frac{dA_{ai}}{ds_i} \Big|_{\bar{\mathbf{s}}}$  is given by the right-hand side of Eq. 14 with  $\beta = 0$ . Eqs. 21 and 22 hold only for integrands above 5 Hz (and are zero below), as per the linear rectifier  $f$ . We assume that the neural decoder has access to background  $\bar{\mathbf{s}}$ , presumed learned (this assumption can be relaxed; see below), and to the linearized response matrix, Eq. 14, but must infer the excess signals  $\Delta s_i$  from excess ORN firing rates  $\Delta r_a(t)$ . Thus, this corresponds to the CS framework (Eq. 20) via  $\Delta\mathbf{r} \rightarrow \mathbf{y}$ ,  $\Delta\mathbf{s} \rightarrow \mathbf{x}$ , and  $dA_{ai}/ds_i \Big|_{\bar{\mathbf{s}}} \rightarrow \mathbf{D}$ . We optimize the cost function in Eq. 20 using sequential least squares programming, implemented in Python through using the scientific package SciPy.

For our simulations, we let sparse components  $s_i$  be chosen as  $s_i = \bar{s}_i + \Delta s_i$ , where  $\bar{s}_i = s_0$  and  $\Delta s_i \sim \mathcal{N}(s_0/3, s_0/9)$ . The measurement matrix  $\mathbf{D}$  depends on the free energy differences  $\epsilon_a$ . For static stimuli,  $\epsilon_a$  equals the fixed point of Eq. 3 in response to the background stimulus with  $\beta = 0$ . For fluctuating stimuli,  $\epsilon_a$  is updated in time by continuously integrating  $r_a(t)$ , via Eqs. 3 and 4; thus, only knowledge of the response  $r_a(t)$  are needed by the decoder. To quantify decoding accuracy, we treat the zero and nonzero components of the sparse odor vector separately. We demand that the  $K$  nonzero components  $\hat{s}_i$  of the estimated sparse vector are within 25% of their true values  $s_i$ , and that the  $N - K$  zero components are estimated less than 10% of  $s_0$ . Together, this ensures that the odorants comprising the odor mixture are estimated sufficiently close to their concentrations, and that the remaining components are sufficiently small. Odor signals  $\mathbf{s}$  are considered correctly decoded if both of these conditions are satisfied for all

790 components  $s_i$ . The relatively lax accuracy demanded  
 791 on the nonzero components is to prevent oversensitivity  
 792 on the unavoidable errors introduced by linearization.  
 793 Qualitatively, our findings are robust to these choices.

794 The naturalistic odor signal (Fig. 3d) was generated  
 795 by randomly varying flow rates of ethyl acetate and  
 796 measuring the concentration with a photo-ionization  
 797 detector [38]. Statistics mirroring a turbulent flow [18]  
 798 were verified (Fig. 3–figure supplement 6).

799  
 800 *Iterative hard thresholding (IHT) and the restricted  
 801 isometry property in compressed sensing*

802  
 803 The purpose of response linearization (Eq. 21) is sim-  
 804 ply to apply compressed sensing reconstruction directly  
 805 using linear programming, without worrying about issues  
 806 of local minima in Eq. 20. This allows us to isolate the  
 807 impact of Weber Law adaptation from the particulari-  
 808 ties of the numerics. An alternate technique for com-  
 809 pressed signal reconstruction, *iterative hard thresholding*  
 810 (IHT), does not minimize the constrained  $L_1$  norm di-  
 811 rectly, rather applying a hard threshold to an iteratively  
 812 updated signal estimate [66]. IHT can be generalized  
 813 straightforwardly to nonlinear constraints, and would ac-  
 814 tually dispense with the need for a learned background  
 815  $\bar{s}$ , simply initializing the iterations from  $\bar{s} = \mathbf{0}$ . Remark-  
 816 ably, this technique works quite well even for non-linear  
 817 measurements [67]. We demonstrate the applicability of  
 818 the IHT algorithm to our odor decoding system in Fig. 3–  
 819 figure supplement 5, which reproduces qualitatively the  
 820 findings in the main text. For these calculations, no back-  
 821 ground odor was assumed, each iterative decoding being  
 822 initialized  $\bar{s} = \mathbf{0}$ .

823 IHT provides an alternate computational technique of  
 824 nonlinear CS, which could be used to both extend and  
 825 verify our results. Further, it allows us to illustrate why  
 826 Weber Law adaptation maintains signal reconstruction  
 827 fidelity in our olfactory sensing model. Like CS using  
 828  $L_1$ -norm minimization, IHT exhibits amenable recon-  
 829 struction and convergence properties under the guarantee  
 830 of the so-called restricted isometry property (RIP) [68].  
 831 Loosely, RIP measures how closely a matrix operator  
 832 resembles an orthogonal transformation when acting on  
 833 sparse vectors. The degree to which RIP is satisfied can  
 834 be understood in terms of the spectrum of a measure-  
 835 ment matrix  $\mathbf{D}$ . In particular, if  $\lambda_i$  are the eigenvalues  
 836 of  $\mathbf{D}_k^T \mathbf{D}_k$ , where  $\mathbf{D}_k$  is any  $k \times m$  submatrix of  $\mathbf{D}$ , and

$$1 - \delta_k \leq \lambda_{\min} \leq \lambda_{\max} \leq 1 + \delta_k \quad (23)$$

837 is satisfied for some  $\delta_k$ , then  $\mathbf{D}$  satisfies the RIP with  
 838 constant  $\delta_k$ . Plainly, the RIP states that the eigenval-  
 839 ues of  $\mathbf{D}_k^T \mathbf{D}_k$ , when acting on  $k$ -sparse vectors, are cen-  
 840 tered around 1. Thus, to intuit why signal reconstruction  
 841 breaks down in the non-adaptive sensing system, we can  
 842 investigate the eigendecomposition of various lineariza-

843 tions of the measurement matrix. We do this now, start-  
 844 ing with a brief description of the IHT.

In the linear setting, IHT seeks sparse signals via the  
 following iterative procedure [66]:

$$\mathbf{x}_{k+1} = H_K(\mathbf{x}_k + \mu \mathbf{D}^T (\mathbf{x}_k + (\mathbf{y} - \mathbf{D}\mathbf{x}_k))) \quad (24)$$

845 where  $\mathbf{x}_k$  is the  $k$ th estimate of the sparse signal  $\mathbf{x}$ ,  $\mu$   
 846 is a step size for the iterations, and  $\mathbf{y}$ ,  $\mathbf{D}$  are as defined  
 847 above.  $H_k(\cdot)$  is a thresholding function which sets all  
 848 but the largest  $K$  values of its argument to zero. The  
 849 nonlinear extension to IHT is [67]:

$$\mathbf{x}_{k+1} = H_K(\mathbf{x}_k + \mu \mathbf{D}_{\mathbf{x}_n}^T (\mathbf{x}_k + (\mathbf{y} - D(\mathbf{x}_k)))), \quad (25)$$

850 where  $D$  is a nonlinear sensing function and  $\mathbf{D}_{\mathbf{x}_n}$  is a  
 851 linearization of  $D$  about the point  $\mathbf{x}_n$ . Reconstructibility  
 852 for  $k$ -sparse signals is guaranteed if  $\mathbf{D}_{\mathbf{x}_n}$  satisfies RIP  
 853 for all  $\mathbf{x}_n$  and all  $k$ -sparse vectors [66]. To get a sense  
 854 of how this is preserved in the adaptive system, we  
 855 calculate the eigenvalues for 1000 choices of  $\mathbf{x}_n$ , acting  
 856 on random signals of given sparsity  $K$  (Fig. 3–figure  
 857 supplement 4). Since the RIP is sensitive to constant  
 858 scalings of the measurement matrix (while the actual  
 859 estimation problem is not), we scaled all columns of  
 860  $\mathbf{D}_{\mathbf{x}_n}$  to norm unity [69]. This normalizes the eigenvalues  
 861 of  $\mathbf{D}_{\mathbf{x}_n}^T \mathbf{D}_{\mathbf{x}_n}$  to center near unity before calculating the  
 862 eigendecomposition, allowing us to assess the degree  
 863 to which the RIP is satisfied. This scaled matrix can  
 864 be used directly in Eq. 25 [67, 69]. The spectra of  
 865 these matrices indicates that the RIP becomes far more  
 866 weakly satisfied in the non-adaptive system than in the  
 867 adaptive one, for sufficient odor complexity and intensity.

#### 868 Network model and classification

869  
 870 For the network model, the AL-to-MB connectivity  
 871 matrix  $\mathbf{J}_1$ , is chosen such that each KC connects pre-  
 872 synaptically to 7 randomly chosen AL glomeruli [25, 26].  
 873 The results shown in Fig. 5 are an average of 10 distinct  
 874 instantiations of this random topology. The  $Z = 2500$   
 875 KCs are then connected by a matrix  $\mathbf{J}_2$  to a readout layer  
 876 of dimension  $Q$ , where  $Q = 2$  for binary and  $Q = N_{ID}$  for  
 877 multi-class classification. Both AL-to-MB and MB-to-  
 878 readout connections are perceptron-type with rectified-  
 879 linear thresholds. The weights of  $\mathbf{J}_1$  and  $\mathbf{J}_2$  are chosen  
 880 randomly from  $\sim \mathcal{N}(0, 1/\sqrt{7})$  and  $\sim \mathcal{N}(0, 1/\sqrt{Z})$ , re-  
 881 spectively. Only the  $\mathbf{J}_2$  and the MB-to-output thresh-  
 882 olds are updated during supervised network training, via  
 883 logistic regression (for binary classification) or its higher-  
 884 dimensional generalization, the softmax cross entropy  
 885 (for multi-class classification).

---

886 [1] R. M. Joseph and J. R. Carlson, Trends in Genetics 31,  
 887 683 (2015).

- [2] L. Buck and R. Axel, *Cell* **65**, 175 (1991).
- [3] P. J. Clyne, C. G. Warr, M. R. Freeman, D. Lessing, J. Kim, and J. R. Carlson, *Neuron* **22**, 327 (1999).
- [4] L. B. Vosshall, H. Amrein, P. S. Morozov, A. Rzhetsky, and R. Axel, *Cell* **96**, 725 (1999).
- [5] R. W. Friedrich and S. I. Korschning, *Neuron* **18**, 737 (1997).
- [6] E. Hallem and J. Carlson, *Cell* **125**, 143 (2006).
- [7] G. Wang, A. F. Carey, J. R. Carlson, and L. J. Zwiebel, *Proceedings of the National Academy of Sciences* **107**, 4418 (2010).
- [8] K. Nara, L. R. Saraiva, X. Ye, and L. B. Buck, *Journal of Neuroscience* **31**, 9179 (2011).
- [9] B. Malnic, J. Hirono, T. Sato, and L. B. Buck, *Cell* **96**, 713 (1999).
- [10] J. G. Hildebrand and G. M. Shepherd, *Annual review of neuroscience* **20**, 595 (1997).
- [11] M. de Bruyne, K. Foster, and J. R. Carlson, *Neuron* **30**, 537 (2001).
- [12] R. I. Wilson, *Annual Review of Neuroscience* **36**, 217 (2013).
- [13] A. Davies, M. Louis, and B. Webb, *PLOS Computational Biology* **11**, 1 (2015).
- [14] D. Saha, K. Leong, S. Peterson, G. Siegel, and B. Raman, *Nature Neuroscience* **16**, 1830 (2013).
- [15] M. Renou, V. Party, A. Rouyar, and S. Anton, *Biosystems* **136**, 35 (2015).
- [16] J. Murlis, *Annual Review of Entomology* **37**, 505 (1992).
- [17] M. Weissburg, *The Biological Bulletin* **198**, 188 (2000), pMID: 10786940.
- [18] A. Celani, E. Villermaux, and M. Vergassola, *Phys. Rev. X* **4**, 041015 (2014).
- [19] R. T. Cardé and M. A. Willis, *Journal of Chemical Ecology* **34**, 854 (2008).
- [20] G. Si, J. Kanwal, Y. Hu, C. J. Tabone, J. Baron, M. E. Berck, G. Vignoud, and A. D. Samuel, *Neuron* **101**, 950962 (2019).
- [21] L.-H. Cao, D. Yang, W. Wu, X. Zeng, B.-Y. Jing, M.-T. Li, S. Qin, C. Tang, Y. Tu, and D.-G. Luo, *Nature Communications* **8**, 1357 (2017).
- [22] C. F. Stevens, *Proceedings of the National Academy of Sciences* **113** (2016).
- [23] K. Asahina, M. Louis, S. Piccinotti, and L. Vosshall, *Journal of Biology* **8**, 9 (2009).
- [24] S. R. Olsen, B. Vikas, and R. I. Wilson, *Neuron* **66**, 287 (2010).
- [25] S. Caron, V. Ruta, L. Abbott, and R. Axel, *Nature* **497**, 113 (2013).
- [26] A. Litwin-Kumar, K. D. Harris, R. Axel, H. Sompolinsky, and L. Abbott, *Neuron* **93**, 1153 (2017).
- [27] K. Krishnamurthy, A. M. Hermundstad, T. Mora, A. M. Walczak, and V. Balasubramanian, *bioRxiv* doi:[10.1101/160382](https://doi.org/10.1101/160382) (2017), 10.1101/160382.
- [28] S. Dasgupta, C. F. Stevens, and S. Navlakha, *Science* **358**, 793 (2017).
- [29] S. L. Brown, J. Joseph, and M. Stopfer, *Nature Neuroscience* **8**, 1568 (2005).
- [30] B. Raman, J. Joseph, J. Tang, and M. Stopfer, *Journal of Neuroscience* **30**, 1994 (2010).
- [31] N. Gupta and M. Stopfer, *Neuron* , 2247 (2014).
- [32] C. D. Wilson, G. O. Serrano, A. A. Koulakov, and D. Rinberg, *Nature Communications* **8**, 1477 (2017).
- [33] M. C. Larsson, A. I. Domingos, W. D. Jones, M. Chiappe, H. Amrein, and L. B. Vosshall, *Neuron* **43**, 703 (2004).
- [34] J. A. Butterwick, J. del Már Mol, K. H. Kim, M. A. Kahlson, J. A. Rogow, T. walz, and V. Ruta, *Nature* **560**, 447 (2018).
- [35] S. A. Montague, D. Mathew, and J. R. Carlson, *Journal of Neuroscience* **31**, 7891 (2011).
- [36] M. C. Stensmyr, H. K. Dweck, A. Farhan, I. Ibba, A. Strutz, L. Mukunda, J. Linz, V. Grabe, K. Steck, S. Lavista-Llanos, D. Wicher, S. Sachse, M. Knaden, P. G. Becher, Y. Seki, and B. S. Hansson, *Cell* **151**, 1345 (2012).
- [37] C. Martelli, J. R. Carlson, and T. Emonet, *Journal of Neuroscience* **33**, 6285 (2013).
- [38] S. Gorur-Shandilya, M. Demir, J. Long, D. A. Clark, and T. Emonet, *eLife* **6**, e27670 (2017).
- [39] E. H. Weber, *EH Weber on the tactile senses* (Psychology Press, 1996).
- [40] G. T. Fechner, *Elemente der psychophysik* (Breitkopf und Härtel, 1860).
- [41] J. Cafaro, *Physiological Reports* **4**, e12762 (2016).
- [42] L.-H. Cao, B.-Y. Jing, D. Yang, X. Zeng, Y. Shen, Y. Tu, and D.-G. Luo, *Proceedings of the National Academy of Sciences* **113**, E902 (2016).
- [43] K. Nagel and R. Wilson, *Nature Neuroscience* **14**, 208216 (2011).
- [44] M. N. Getahun, S. B. Olsson, S. Lavista-Llanos, B. S. Hansson, and D. Wicher, *PLoS One* **8**, e58889 (2013).
- [45] M. N. Getahun, M. Thoma, S. Lavista-Llanos, I. Keesey, R. A. Fandino, M. Knaden, D. Wicher, S. B. Olsson, and B. S. Hansson, *Journal of Experimental Biology* **219**, 3428 (2016).
- [46] H. Guo, K. Kunwar, and D. Smith, *The Journal of Neuroscience* **37**, 9465 (2017).
- [47] H. Guo and D. P. Smith, *Journal of Experimental Neuroscience* **11**, 1 (2017).
- [48] A. J. Waite, N. W. Frankel, and T. Emonet, *Annual Review of Biophysics* **47**, 595 (2018), pMID: 29618219, <https://doi.org/10.1146/annurev-biophys-062215-010954>.
- [49] Kadakia, N. 2019. ORN-WL-gain-control. GitHub. <https://github.com/emonetlab/ORN-WL-gain-control.git>. 926cfa7.
- [50] P. J. Rousseeuw, *Journal of Computational and Applied Mathematics* **20**, 53 (1987).
- [51] L. van der Maaten and G. Hinton, *Journal of Machine Learning Research* **9**, 2579 (2008).
- [52] L. B. Vosshall, A. M. Wong, and R. Axel, *Cell* **102**, 147 (2000).
- [53] D. Donoho, *IEEE Transactions on Information Theory* **52**, 1289 (2006).
- [54] E. Candes, J. Romberg, and T. Tao, *Communications on Pure and Applied Mathematics* **59**, 1207 (2006).
- [55] C. Pehlevan, A. Genkin, and D. B. Chklovskii, in *2017 51st Asilomar Conference on Signals, Systems, and Computers* (2017) pp. 593–600.
- [56] S. R. Olsen and R. I. Wilson, *Nature* **452**, 952 (2008).
- [57] A. C. Keene and S. Waddell, *Nature Reviews Neuroscience* **8**, 341 (2007).
- [58] R. G. Vogt and L. M. Riddiford, *Nature* **293**, 161 (1981).
- [59] K. Menuz, N. K. Larter, J. Park, and J. R. Carlson, *PLoS Genetics* **10**, e1004810 (2014).
- [60] G. Reddy, J. Zak, M. Vergassola, and V. N. Murthy, *eLife* **7**, 344958 (2018).
- [61] C.-Y. Su, K. Menuz, J. Reisert, and J. Carlson, *Nature* **492**, 76 (2012).

- 1017 [62] M. Papadopoulou, S. Cassenaer, T. Nowotny, and 1027  
 1018 G. Laurent, *Science* **332**, 721 (2011).  
 1019 [63] N. Gupta and M. Stopfer, *Current Opinion in Neurobi- 1028*  
 1020 ology **21**, 768 (2011).  
 1021 [64] B. Wark, B. N. Lundstrom, and A. Fairhall, *Current 1029*  
 1022 Opinion in Neurobiology **17**, 423 (2007).  
 1023 [65] I. Nemenman, in *Quantitative Biology: From Molecular 1030*  
 1024 to Cellular Systems, edited by M. E. Wall (CRC Press, 1031  
 1025 USA, 2012) Chap. 4, pp. 73–91.  
 1026 [66] T. Blumensath and M. E. Davies, *Applied and Compu- 1032*  
 1027 tational Harmonic Analysis **27**, 265 (2009).  
 1028 [67] T. Blumensath, *IEEE Transactions on Information The- 1029*  
 1030 ory **69** (2013).  
 1031 [68] E. Candes, J. Romberg, and T. Tao, *Communications on 1032*  
 1033 Pure and Applied Mathematics **LIX**, 12071223 (2006).  
 1034 [69] T. Blumensath and M. E. Davies, in *Proceedings of 1035*  
 1035 *SPARS'09 - Signal Processing with Adaptive Sparse Structured Representations* (Saint Malo, France, April 2009) (2009).

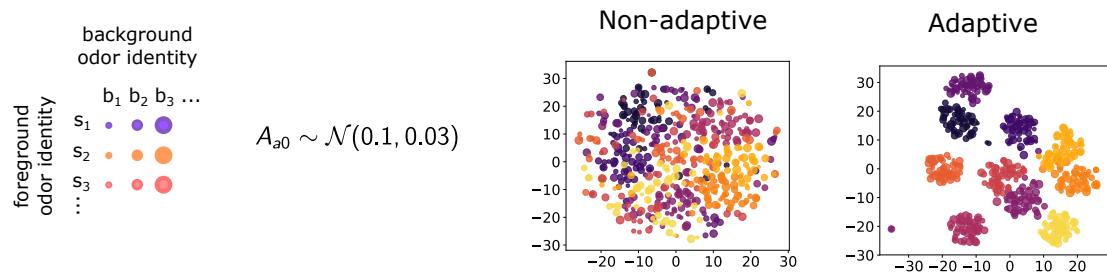


FIG. 2–figure supplement 1: t-SNE projections for non-adaptive ( $\beta = 1$ ) and adaptive ( $\beta = 0$ ) systems, when background firing rates depend on ORN identity. Background active fractions  $A_{0a}$  are chosen normally with mean 0.1 and deviation 0.03, corresponding to background firing rates around of  $\simeq 20\text{-}40$  Hz.

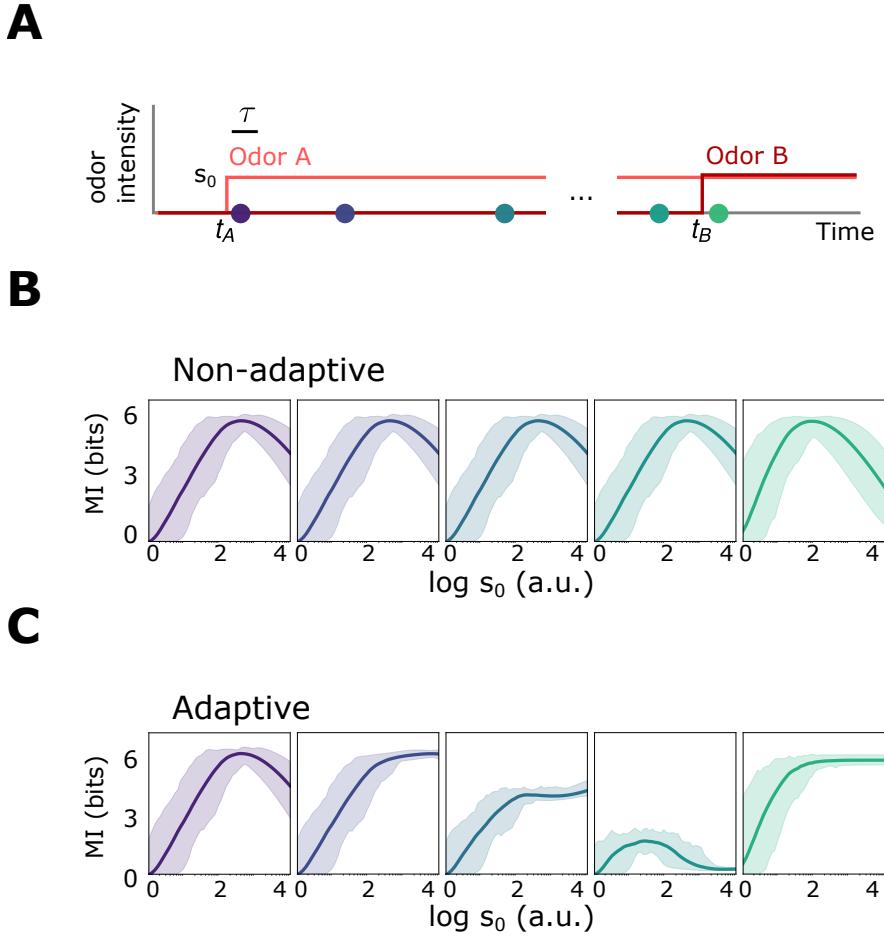


FIG. 2—figure supplement 2: Front-end adaptive feedback preserves information capacity of the ORN sensing repertoire. **A** Mutual information between signal  $\mathbf{s}(t) = \mathbf{s}_A(t) + \mathbf{s}_B(t)$  and response  $\mathbf{r}(t)$  is calculated at various points in time  $t$  for an odor environment consisting of two step odors, A and B. Odor A, with concentration  $\mathbf{s}_A(t)$ , turns on at time  $t_A$  and a odor B, with concentration  $\mathbf{s}_B(t)$ , turns on at some later time  $t_B$ . Both odors have similar intensities  $\sim s_0$  and similar molecular complexity ( $k = 4$ ). **B** Mutual information as a function of  $s_0$  for the non-adaptive system, respectively, at different time points after  $t_A$ , corresponding to the dots in A. The mutual information carried by distinct ORNs is represented by the shaded region; their average is plotted by the heavy line. In the non-adaptive system, the mutual information peaks in the regime of high sensitivity after the arrival of odor A (purple, blue), and shifts leftward with the onset of odor B (teal, green). The leftward shifts occurs since stronger signals are more prone to response saturation (compromising information transfer) as odor B arrives. **C** Same as B, now for the adaptive system. The MI mimics the non-adaptive case at the onset of odor A, before adaptation has kicked in (purple). As the system adapts and responses decrease toward baseline, previously saturating signal intensities now cross the regime of maximal sensitivity, which therefore shifts rightward to higher  $s_0$  (dark blue). Much later, but before the arrival of odor B, the ORNs that responded now fire at a similar adapted firing rate  $\sim 30$  Hz, irrespective of odor identity, so the mutual information drops to zero. However, having now adjusted its sensitivity to the presence of odor A, the system can respond appropriately to odor B: the MI at  $t_B$  is nearly 6 bits across decades of concentration immediately following  $t_B$  (green).

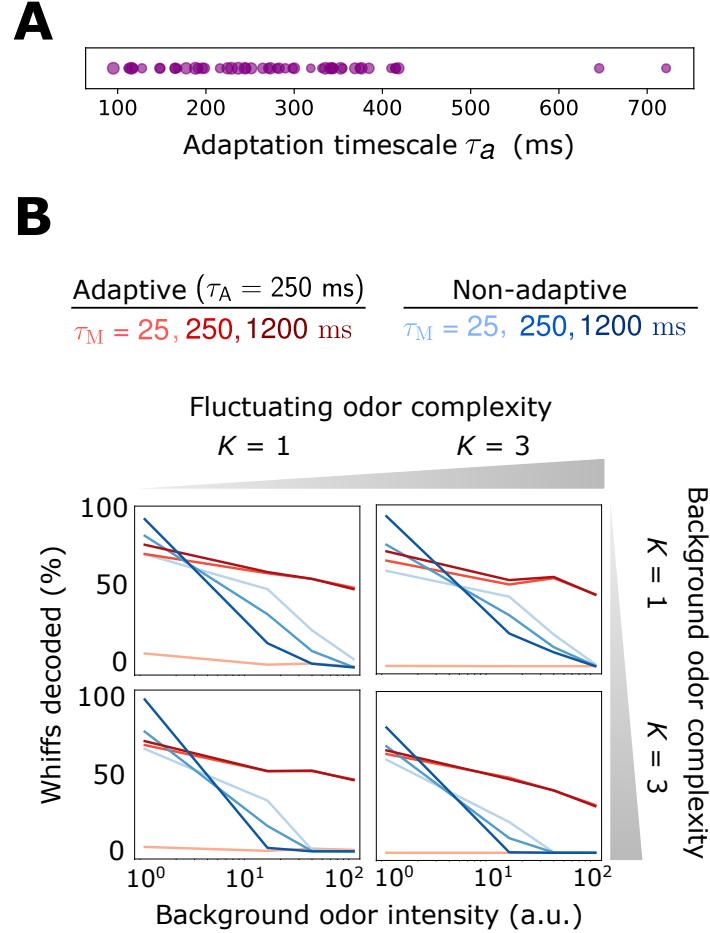


FIG. 3—figure supplement 1: Decoding accuracy for system with broader distribution of adaptive timescales  $\tau$ . **A** Distribution of timescales for all ORNs  $a$  (purple dots). Here,  $\tau_a \sim= 10^X$  where  $\tau = 250$  ms as in the main text and  $X \sim \mathcal{N}(0, 0.2)$ . **B** Individual plots show the percent of accurately decoded odor whiffs (same fluctuating odor signal used in the main text) as a function of background odor intensity, for the non-adaptive (blue) and adaptive (red) systems, for different  $\tau_M$  (line shades). Plots are arrayed by the complexity of the naturalistic signal (column-wise) and the complexity of the background odor (row-wise).

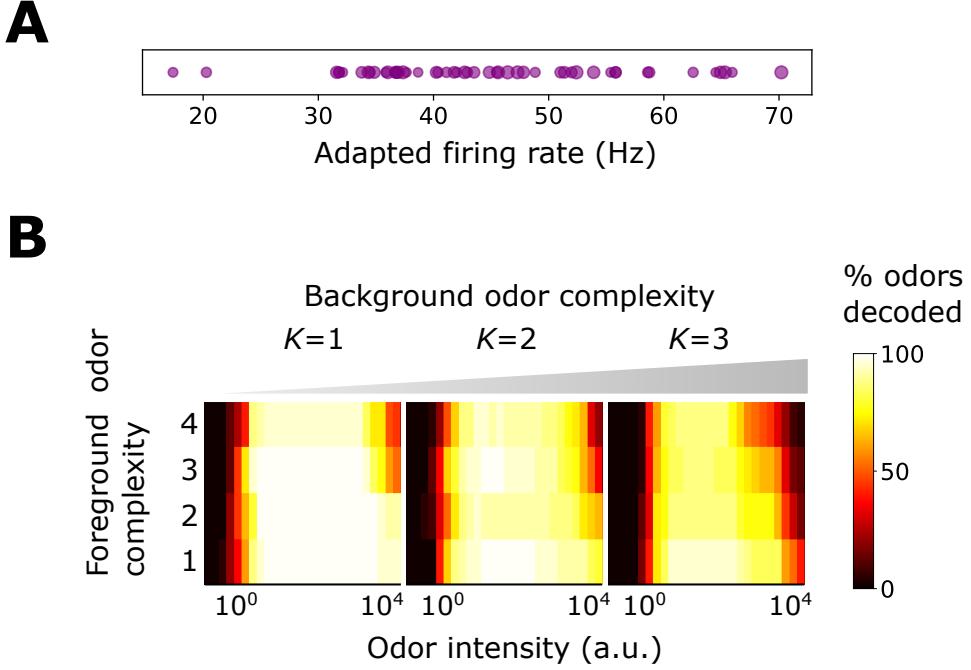


FIG. 3—figure supplement 2: Benefits conferred by Weber-Fechner adaptation remain for a broader distribution of baseline firing rates  $A_{0a}$ , now assumed to be ORN-dependent and chosen from a normal distribution. **A** Distribution of  $A_{0a}$ . **B** Decoding accuracy of foreground odors in the presence of background odors.

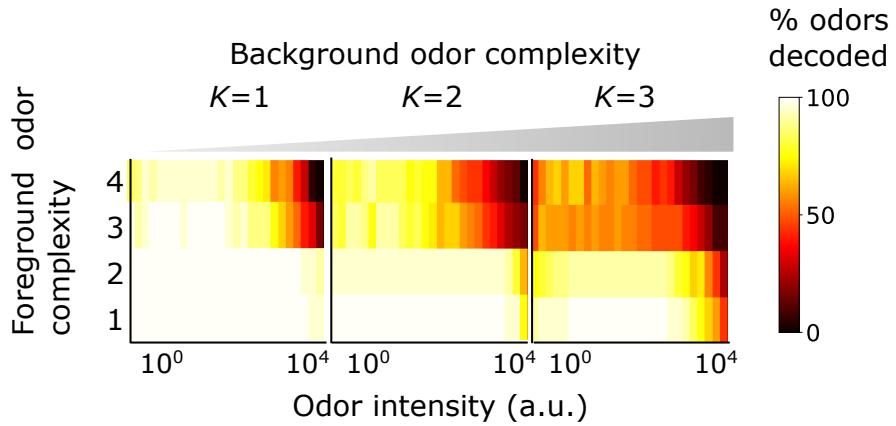


FIG. 3—figure supplement 3: Benefits conferred by Weber-Fechner adaptation remain for 2 binding sites per receptor. This might conceivably occur in insect olfactory receptors, heterotetramers consisting of 4 Orco/Or subunits that gate a central ion channel pathway [34]. Plotted is the decoding accuracy of foreground odors in the presence of background odors.

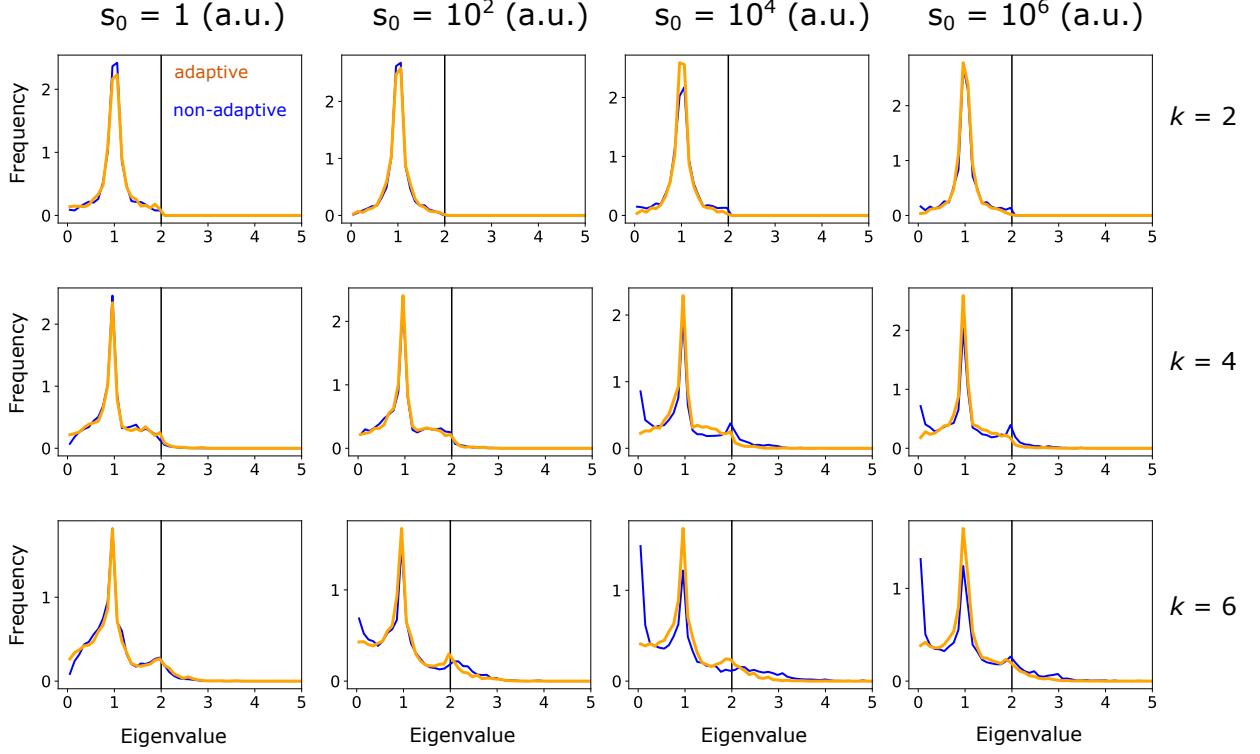


FIG. 3—figure supplement 4: Eigenvalue distribution of  $\mathbf{A}_{\mathbf{x}_n}^T \mathbf{A}_{\mathbf{x}_n}$ , where  $\mathbf{A}_{\mathbf{x}_n}$  is a  $m \times k$  submatrix of the column-normalized linearized ORN response matrix  $\mathbf{A}$ , evaluated at the linearization point  $\mathbf{x}_n$ . Note that  $\mathbf{x}_n$  is  $k$ -sparse, but its components do not necessarily align with the  $k$  columns chosen for the sub-matrix. Eigenvalues are calculated for the adaptive (orange) and non-adaptive (blue) systems, for 1000 randomly chosen linearization points  $\mathbf{x}_n$  and submatrices. Plots are arranged for various odor sparsities (by row) and odor intensities (by column). The restricted isometry property is satisfied when the eigenvalues lie between 0 and 2 (black vertical line), and is more strongly satisfied the more centered the distribution is around unity. The increase in near-zero eigenvalues for the non-adaptive system at higher odor complexities and intensities (lower right plots) indicates the weaker fulfillment of the restricted isometry property for these signals, and leads to higher probability of failure in compressed sensing signal reconstruction.

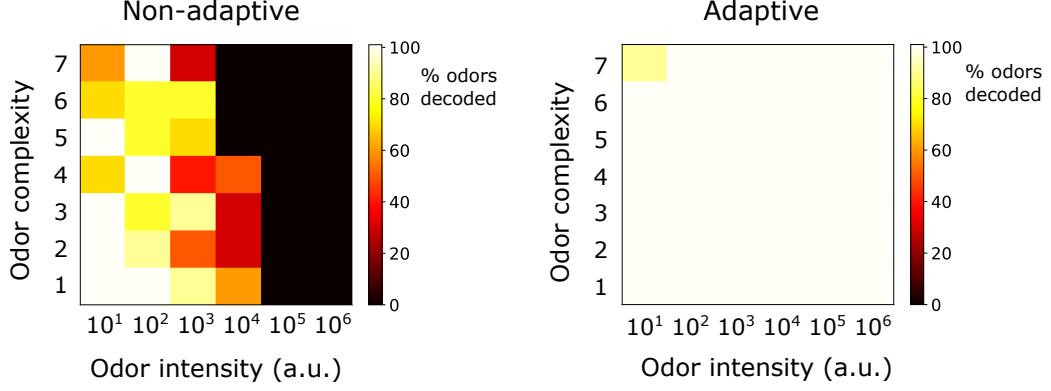


FIG. 3—figure supplement 5: Decoding of odor signals (no background odors) using the IHT algorithm [66, 67] qualitatively reproduces the results from the main text, which used traditional CS with background linearization. In the adaptive case, IHT actually exhibits superior accuracy to traditional CS, though IHT demands more compute time. The results here show odor decoding accuracy for sparse odor signals of given complexity and intensity, averaged over 10 distinct identities. Odors are considered accurately decoded if the  $K$  sparse components are estimated within 25% and the components not in the mixture are estimated below 10% of  $s_0$ . The iterative algorithm was initialized at  $\hat{\mathbf{x}} = \mathbf{0}$  and run forward until  $\hat{\mathbf{x}}$  was stationary, or 10000 iterations were reached. Step size  $\mu$  in Eq. 25 was set to  $s_0/20$ . At each step, the linearized response ( $\mathbf{A}_{\mathbf{x}_n}$  in Eq. 25) was evaluated at the result of the previous iteration. IHT also requires an assumption on the number of components in the mixture (which defines  $H_K(\cdot)$  in Eq. 25); here, that was set to twice the actual sparsity of true signal.

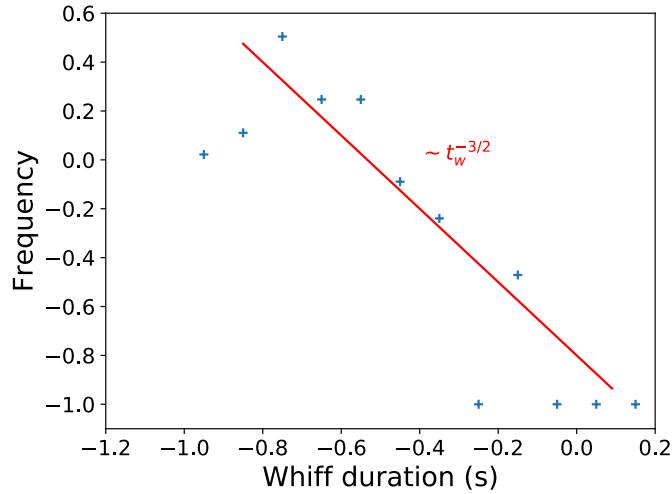


FIG. 3—figure supplement 6: Distribution of whiff durations in naturalistic stimulus, compared to the theoretical prediction [18].

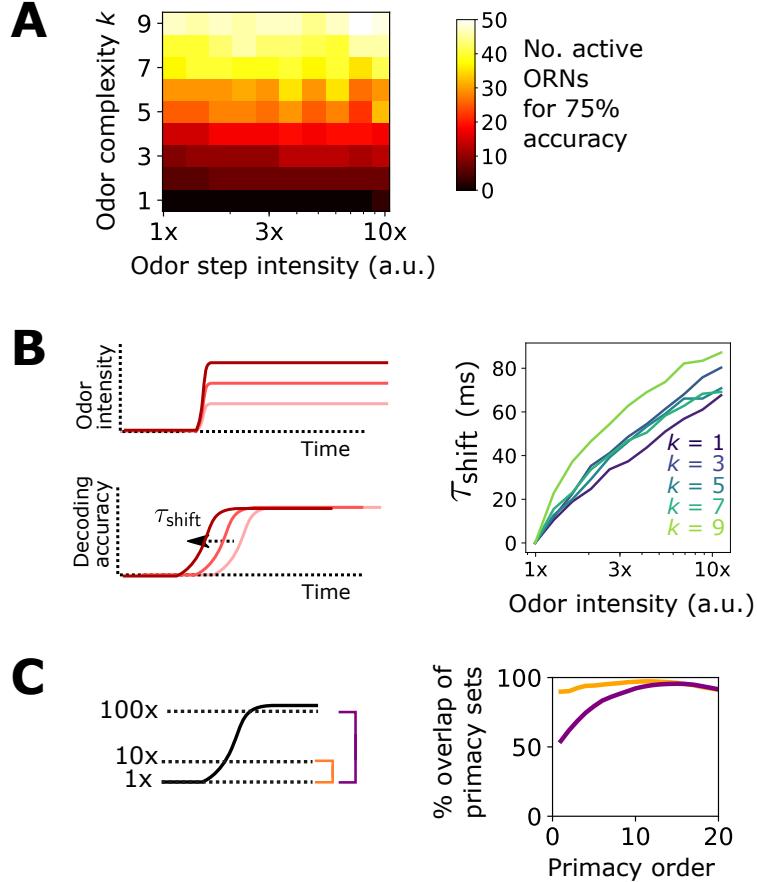


FIG. 4—figure supplement 1: Additional results pertaining to the primacy coding hypothesis. **A** Percent of active ORNs required for 75% accuracy of a steep sigmoidal odor step, as a function of odor step intensity and odor complexity. For low complexities, a primacy set of fewer ORNs may be sufficient to decode the full odor signal; for higher complexities, the entire ORN repertoire is required. **B** In the primacy coding hypothesis, the primacy set is realized sooner for stronger odor signals, so odors are decoded earlier in time, resulting in a perceptual time shift with increasing odor concentration [32]. We also find this shift in our compressed sensing decoding framework (right plot), which rises monotonically with step height for various odor complexities, in agreement with primacy coding. **C** The consistency of a primacy code across changes in background odor concentration, in a system with Weber Law adaptation. We calculate the primacy set for odor A (step odor; black) in the presence of either a weak, medium, or strong background (dotted lines; 1x, 10x, 100x a.u.), assuming the system has adapted its response to the background as described in the main text. Averaged across odor A identities, primacy sets for odor A when in the 1x background are nearly identical to those when odor A is in the 10x background (right plot; yellow). The same holds true when comparing the 1x and 100x backgrounds, for sufficiently large primacy order, above 8 or so (right plot; purple). This indicates that Weber Law adaptation preserves primacy codes across disparate environmental conditions.

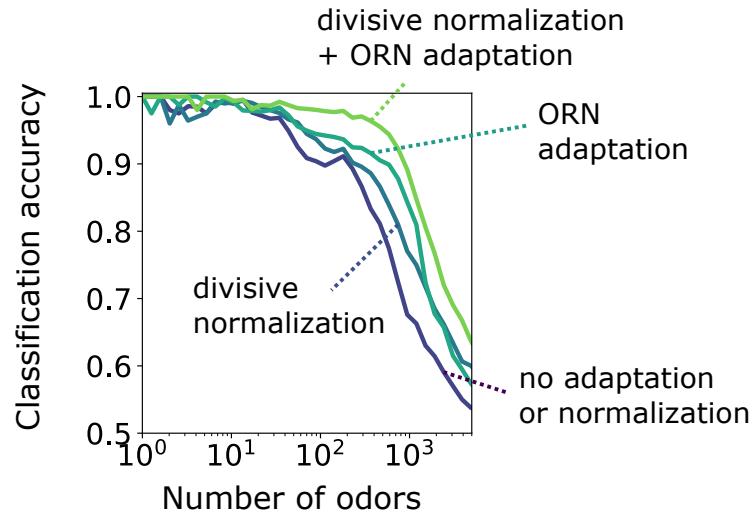


FIG. 5—figure supplement 1: Accuracy of binary classification by odor valence, for odors whose concentrations span a narrow range of concentrations (1 order of magnitude). Accuracy is plotted as a function of the number of distinct odor identities classified by the trained network, in systems with only ORN adaptation, only divisive normalization, both or neither. Decoding gains conferred by divisive normalization and/or ORN adaptation are much smaller than when odors span a much larger range of concentrations, as shown in the main text.