

Front-end Weber-Fechner gain control enhances the fidelity of combinatorial odor coding

Nirag Kadakia¹ and Thierry Emonet^{1,2*}

*For correspondence:

thierry.emonet@yale.edu (TE)

¹Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06511; ²Department of Physics, Yale University, New Haven, CT 06511

Abstract We showed previously (Gorur-Shandilya et al 2017) that *Drosophila* olfactory receptor neurons (ORNs) expressing the co-receptor Orco scale their gain inversely with mean odor intensity according to Weber-Fechner's law. Here we show that this front-end 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 this adaptation mechanism 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, according to Weber-Fechner's law. Thus, a common scaling of the gain across Orco-expressing ORNs may be a key feature of ORN adaptation that helps preserve combinatorial odor codes in naturalistic landscapes.

Introduction

Animals identify and discriminate odors using olfactory receptors (Ors) expressed in olfactory receptor neurons (ORNs) (Joseph and Carlson, 2015; Buck and Axel, 1991; Clyne et al., 1999; Vosshall et al., 1999). Individual ORNs, which typically express a single Or, respond to many odorants, while individual odorants activate many distinct ORNs (Friedrich and Korschning, 1997; Hallem and Carlson, 2006; Wang et al., 2010; Nara et al., 2011). Odors are thus encoded by the combinatorial patterns of activity they elicit in the sensing periphery (Malnic et al., 1999; Wang et al., 2010; Hildebrand and Shepherd, 1997; Hallem and Carlson, 2006; de Bruyne et al., 2001; Friedrich and Korschning, 1997), patterns decoded downstream into behavioral response (Wilson, 2013; Davies et al., 2015). Still, ethologically-relevant odors are often mixed with background ones (Saha et al., 2013; Renou et al., 2015) and intensity can vary widely and rapidly as odors are carried by the wind (Murlis, 1992; Weissburg, 2000; Celani et al., 2014; Cardé and Willis, 2008). 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 (Hallem and Carlson, 2006; Si et al., 2019), which together with inhibitory odorants enhance coding capacity (Si et al., 2019; Cao et al., 2017; Hallem and Carlson, 2006; Stevens, 2016). In antennal lobe (AL) glomeruli, mutual lateral inhibition normalizes population response, reducing the dependency of activity patterns on odor concentration (Asahina et al., 2009; Olsen et al., 2010). Further downstream, sparse connectivity to the mushroom body (MB) helps maintain neural representations of odors, and facilitates compressed sensing and associative learning schemes (Caron et al., 2013; Litwin-

41 **Kumar et al., 2017; Krishnamurthy et al., 2017; Dasgupta et al., 2017**). Finally, temporal features
 42 of neural responses contribute to concentration-invariant representations of odor identity (**Brown**
 43 **et al., 2005; Raman et al., 2010; Gupta and Stopfer, 2014; Wilson et al., 2017**).

44 Here we examine how short-time ORN adaptation at the very front-end of the insect olfactory
 45 circuit contributes to the fidelity of odor encoding. Our theoretical study is motivated by the recent
 46 discovery of invariances in the signal transduction and adaptation dynamics of ORNs expressing
 47 the co-receptor Orco. ORN response is initiated upon binding of odorant molecules to olfactory
 48 receptors (ORs), opening the ion channels they form with the co-receptor Orco (**Larsson et al., 2004;**
 49 **Butterwick et al., 2018**). Because of differences in odor-receptor affinities, the responses of ORNs
 50 to diverse odorants of the same concentration differ widely (**Hallez and Carlson, 2006; Montague**
 51 **et al., 2011; Stensmyr et al., 2012**). In contrast, downstream from this input nonlinearity, signal
 52 transduction and adaptation dynamics exhibit a surprising degree of invariance with respect to
 53 odor-receptor identity: reverse-correlation analysis of ORN response to fluctuating stimuli produces
 54 highly stereotyped, concentration-invariant response filters (**Martelli et al., 2013; Si et al., 2019;**
 55 **Gorur-Shandilya et al., 2017**).

56 These properties stem in part from an apparently invariant adaptive scaling law in ORNs:
 57 gain varies inversely with mean odor concentration according to the Weber-Fechner Law of psy-
 58 chophysics (**Weber, 1996; Fechner, 1860**), irrespective of the odor-receptor combination (**Gorur-**
 59 **Shandilya et al., 2017; Cafaro, 2016; Cao et al., 2016**). This invariance can be traced back to
 60 adaptative feedback mechanisms in odor transduction, upstream of ORN firing (**Nagel and Wilson,**
 61 **2011; Cao et al., 2016; Cafaro, 2016; Gorur-Shandilya et al., 2017**), which depend on the activity
 62 of the signaling pathway rather than on the identity of its receptor (**Nagel and Wilson, 2011**). The
 63 generality of the adaptive scaling suggests it could be mediated by the highly conserved Orco
 64 co-receptor (**Butterwick et al., 2018; Getahun et al., 2013, 2016; Guo et al., 2017**), which has been
 65 already implicated in other types of odor adaptation, taking place over longer timescales (**Guo**
 66 **and Smith, 2017; Guo et al., 2017**).

67 While in a simpler system such as *E. coli* chemotaxis (**Waite et al., 2018**), adaptive feedback via the
 68 Weber-Fechner Law robustly maintains sensitivity over concentration changes, the implication for a
 69 multiple-channel system – which combines information from hundreds of cells with overlapping
 70 receptive fields – is less clear. Here we combine a biophysical model of ORN adaptive response and
 71 neural firing with various sparse signal decoding frameworks to explore how ORN adaptation with
 72 Weber-Fechner scaling affects combinatorial coding and decoding of odor signals spanning varying
 73 degrees of intensity, molecular complexity, and temporal structure. We find that this front-end
 74 adaptive mechanism promotes the accurate discrimination of odor signals from backgrounds
 75 of varying molecular complexity, and aids other known mechanisms of neural processing in the
 76 olfactory circuit to maintain representations of odor identity across environmental changes.

77 Results

78 Model of ORN sensing repertoire

79 To model ORN firing rates in response to time-dependent odor signals, we extended a minimal
 80 model (**Gorur-Shandilya et al., 2017**) that reproduces the Weber-Fechner gain adaptation and firing
 81 rate dynamics measured in individual *Drosophila* ORNs in response to Gaussian and naturalistic
 82 signals (code available on GitHub (**Kadakia, 2019**)).

83 We consider a repertoire of $M = 50$ ORN types that each express one type of Or together with
 84 the co-receptor Orco (**Larsson et al., 2004**). Within ORNs of type $a = 1, \dots, M$, Or-Orco complexes
 85 form non-selective cation channels (**Butterwick et al., 2018**) (**Figure 1A**) that switch between active
 86 and inactive conformations, while simultaneously binding to odorants i with affinity constants, K_{ai}^*
 87 and K_{ai} , respectively (**Nagel and Wilson, 2011; Gorur-Shandilya et al., 2017**). For simplicity we only
 88 consider agonists, i.e. $K_{ai}^* > K_{ai}$, and assume receptors can only bind one odorant at a time. The
 89 analysis can easily be extended to include inhibitory odorants, which increases coding capacity (**Cao**

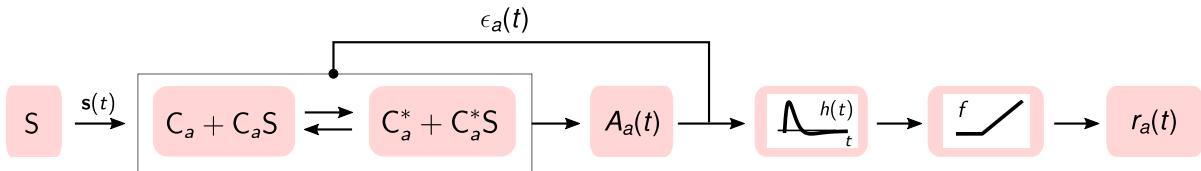
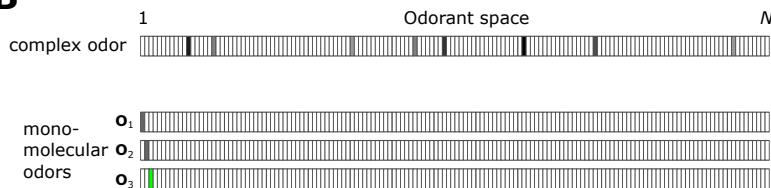
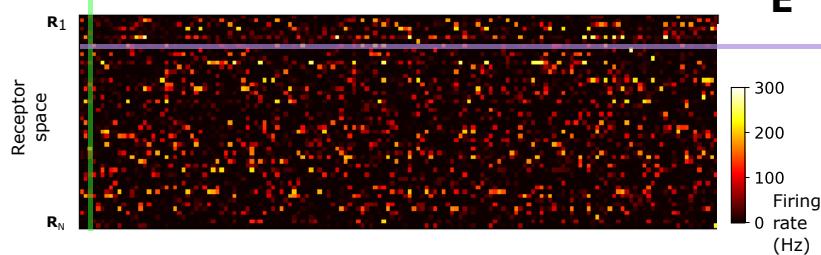
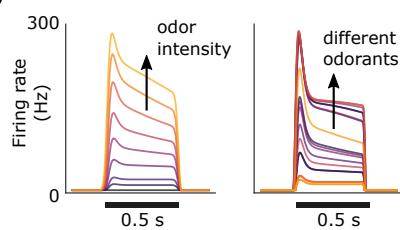
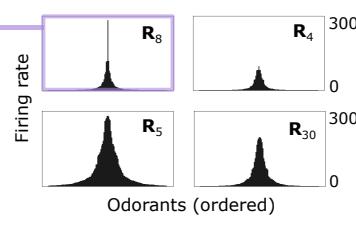
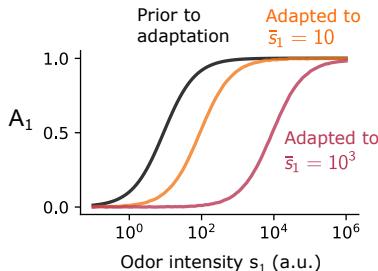
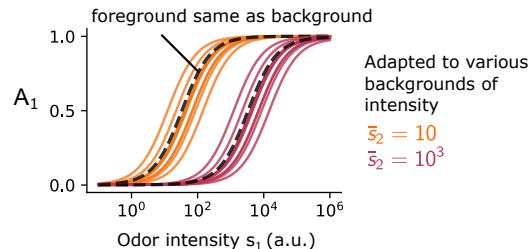
A**B****C****D****E****F****G**

Figure 1. Simple ORN model. (Gorur-Shandilya et al., 2017) **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 $K_a^* = (K_{a1}^*, \dots, K_{ai}^*, \dots, K_{aN}^*)$ and K_a for the active and inactive conformations, respectively (Eqs. 1-2). Adaptation is mediated by a negative feedback (Nagel and Wilson, 2011) from the activity of the channel onto the free energy difference $\epsilon_a(t)$ with timescale τ . 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 $s = [s_1, \dots, s_i, \dots, s_N]$, whose components s_i are the concentrations of the individual molecular constituents of s . **C** Step-stimulus firing rate of 50 ORNs to the $N=150$ possible monomolecular odorants $s = s_i$, given power-law distributed affinity constants (Si et al., 2019). **D** Temporal responses of a representative ORN to a pulse stimulus, for a single odorant at several intensities (left), or to many odors 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 (Hallem and Carlson, 2006). **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).

90 *et al., 2017*). Dissociation (inverse affinity) constants are chosen from a power law distribution
 91 ($\alpha = 0.35$) recently found across ORN-odor pairs in *Drosophila* larvae (*Si et al., 2019*). For a handful of
 92 ORNs, we choose a very large value for one of the K_{ai}^* to mimic high responders to private odors
 93 relevant to innate responses (*Stensmyr et al., 2012*). These private odors do not affect the general
 94 findings.

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

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

97 C_a and C_a^* represent unbound channels in the inactive and active conformation. Here, $\mathbf{K}_a \cdot \mathbf{s}(t) =$
 98 $\sum_i^K K_{ai} s_i(t)$, where $s_i(t)$ is the time-dependent concentration of the i -th monomolecular component
 99 of the odor signal $\mathbf{s}(t)$ at time t (*Figure 1B*). $N = 150$ is the size of the molecular odorant space
 100 (*Figure 1B*). Eq. 1 can be rearranged as (derivation in Methods):

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

101 The two terms in the exponential represent the change in the channel's free energy due to the
 102 binding of odorant i , and the free energy difference ϵ_a between the unbound states C_a and C_a^* , in
 103 units of $k_B T$. Because $K_{ai}^* > K_{ai}$, a sudden increase in the concentration of excitatory odor results in
 104 an increase in activity A_a .

105 Upon prolonged stimulation, ORNs adapt. At least one form of adaptation, which takes place
 106 over short time scale, $\tau \approx 250$ ms (*Gorur-Shandilya et al., 2017*), involves a negative feedback of
 107 the Or-Orco channel activity onto the channel sensitivity (*Nagel and Wilson, 2011; Gorur-Shandilya*
 108 *et al., 2017*). To model this adaptation process, we assume that inward currents elicited by activating
 109 Or-Orco channels eventually result in an increase of the free energy difference $\epsilon_a(t)$, possibly via a
 110 feedback onto Orco (*Butterwick et al., 2018*):

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

111 where $\epsilon_{L,a} < \epsilon_a(t) < \epsilon_{H,a}$. The lower bound $\epsilon_{L,a}$ determines the spontaneous activity of the channel.
 112 The higher bound $\epsilon_{H,a}$ determines the concentrations of odors at which adaptation is unable to
 113 keep up and saturation occurs (*Gorur-Shandilya et al., 2017*). Through these dynamics, $\epsilon_a(t)$ can
 114 compensate for changes in free energy due to ligand binding (see Eq. 2), returning the activity
 115 A_a towards an adapted level A_{0a} above the spontaneous activity. Since ϵ_a is bounded below, a
 116 minimum amount of signal intensity is needed for adaptation to kick in. Finally, the firing rate is
 117 modeled by passing the activity $A_a(t)$ through the derivative-taking bi-lobed filter $h(t)$ and a rectifying
 118 nonlinearity f (*Gorur-Shandilya et al., 2017*):

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

119 where \otimes is convolution. When deconvolved from stimulus dynamics, the shapes of the temporal
 120 kernels of *Drosophila* ORNs that express Orco tend to be stereotyped for many odor-receptor
 121 combination (*Martelli et al., 2013; Gorur-Shandilya et al., 2017; Si et al., 2019*) (although there
 122 are known exceptions such as super-sustained responses (*Montague et al., 2011*)). Moreover,
 123 adaptation is not intrinsic to the receptor (*Nagel and Wilson, 2011*). Accordingly, for simplicity τ ,
 124 $h(t)$, and f are assumed independent of receptor and odorant identities.

125 This minimal model reproduces the essential features of ORN response to odorant pulses (*Nagel*
 126 *and Wilson, 2011; Martelli et al., 2013; Cao et al., 2016*). In the absence of stimulus, ORNs fire

127 spontaneously at rates (1-10 Hz) (*Hallem and Carlson, 2006*) set by the lower free energy bound $\epsilon_{L,a}$,
 128 which we choose from a normal distribution (*Figure 1D*). For sufficiently strong stimuli, adaptation
 129 causes ϵ_a to increase, compensating for the drop in free energy difference due to ligand binding.
 130 This gradually reduces the firing rate to a steady state level $r(A_{0a}) \simeq 30\text{-}40$ Hz (*Gorur-Shandilya*
 131 *et al., 2017*) (*Figure 1D*). The diversity of temporal firing responses and tuning curves measured
 132 experimentally (*Hallem and Carlson, 2006; Montague et al., 2011; Brown et al., 2005; Gupta and*
 133 *Stopfer, 2014; Raman et al., 2010*) arise naturally in the model due to the distribution of chemical
 134 affinity constants and the nonlinearity of Eq. 2 (*Figure 1B-Figure 1E*).

135 The model also reproduces Weber-Fechner scaling of the gain with the inverse of the mean
 136 odorant intensity \bar{s}_i (*Gorur-Shandilya et al., 2017; Cao et al., 2016*). For small fluctuations Δs_i
 137 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
 138 satisfied provided $A_a(\bar{s}_i)$ is approximately constant (derivation in Methods). In our model, since
 139 the rate of adaptation depends only on the activity of the ion channel (right hand-side of Eq. 3),
 140 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
 141 process adjusts the sensitivity of the ORN by matching the dose responses to the mean signal
 142 concentration, while maintaining their log-slopes (*Figure 1F*). However, for foreground odors mixed
 143 with background odors to which the system has adapted, the dose response curves now exhibit
 144 background-dependent shifts (*Figure 1G*).

145 While this phenomenological model could be extended to include further details – e.g. we
 146 could relax the quasi-steady-state assumption in Eq. 2, use a more complex model for channel
 147 adaptation and neural firing (*Gorur-Shandilya et al., 2017*), or consider feedforward mechanisms
 148 in addition to negative integral feedback (*Schulze et al., 2015*) – this minimally-parameterized form
 149 captures the key dynamical properties of Orco-expressing ORNs relevant to our study: receptor-
 150 independent adaptation (*Nagel and Wilson, 2011*) with Weber-Fechner scaling (*Gorur-Shandilya*
 151 *et al., 2017; Cafaro, 2016; Cao et al., 2016*) that maintains response time independent of mean
 152 stimulus intensity (*Martelli et al., 2013; Gorur-Shandilya et al., 2017*), along with a diversity of
 153 temporal firing patterns in response to a panel of monomolecular odorants (*Hallem and Carlson,*
 154 *2006; Montague et al., 2011; Brown et al., 2005; Gupta and Stopfer, 2014; Raman et al., 2010*
 155 (*Figure 1D-1E*)).

156 **Front-end Weber-Fechner adaptation preserves odor coding among background 157 and intensity confounds**

158 The identity of an odor is encoded by the pattern of ORN firing responses. However, when a novel
 159 foreground odor is presented atop an existing background odor, this pattern may depend also on
 160 the background odor, rendering ORN responses less informative about foreground odor identity. To
 161 understand how front-end Weber-Fechner adaptation might help encode novel foreground odors in
 162 the presence of background odors, we considered environments containing various combinations
 163 of foreground odors s and background odors \bar{s} , and asked how similar are the ORN responses r to
 164 a given s but different \bar{s} .

165 Since it is not possible to visualize the 50-dimensional space of ORN responses, we projected
 166 ORN responses onto a two-dimensional space using t-distributed stochastic neighbor embedding
 167 (t-SNE) (*van der Maaten and Hinton, 2008*). Like principle component analysis (PCA), t-SNE allows
 168 a visualization of high-dimensional objects in such a way that desirable features of the original
 169 dataset are preserved (*Figure 2A*). PCA, for example, retains much of the data variance. t-SNE
 170 retains the proximity of an object to its nearest neighbors. Specifically, it constructs a probability
 171 distribution Q_H based on pairwise distances between nearby objects, assigning higher probability
 172 to closer objects. It then determines where the objects would live in a lower-dimensional space,
 173 such that the analogous distribution Q_L in this space is most similar to Q_H . t-SNE is widely used
 174 to cluster objects (in our case, ORN responses r to different foreground odors on top of diverse
 175 background odors) by similarity (here, foreground odor identity). However, because t-SNE uses
 176 local information from only nearest neighbors, global distances and scales are not preserved (*Zhou*

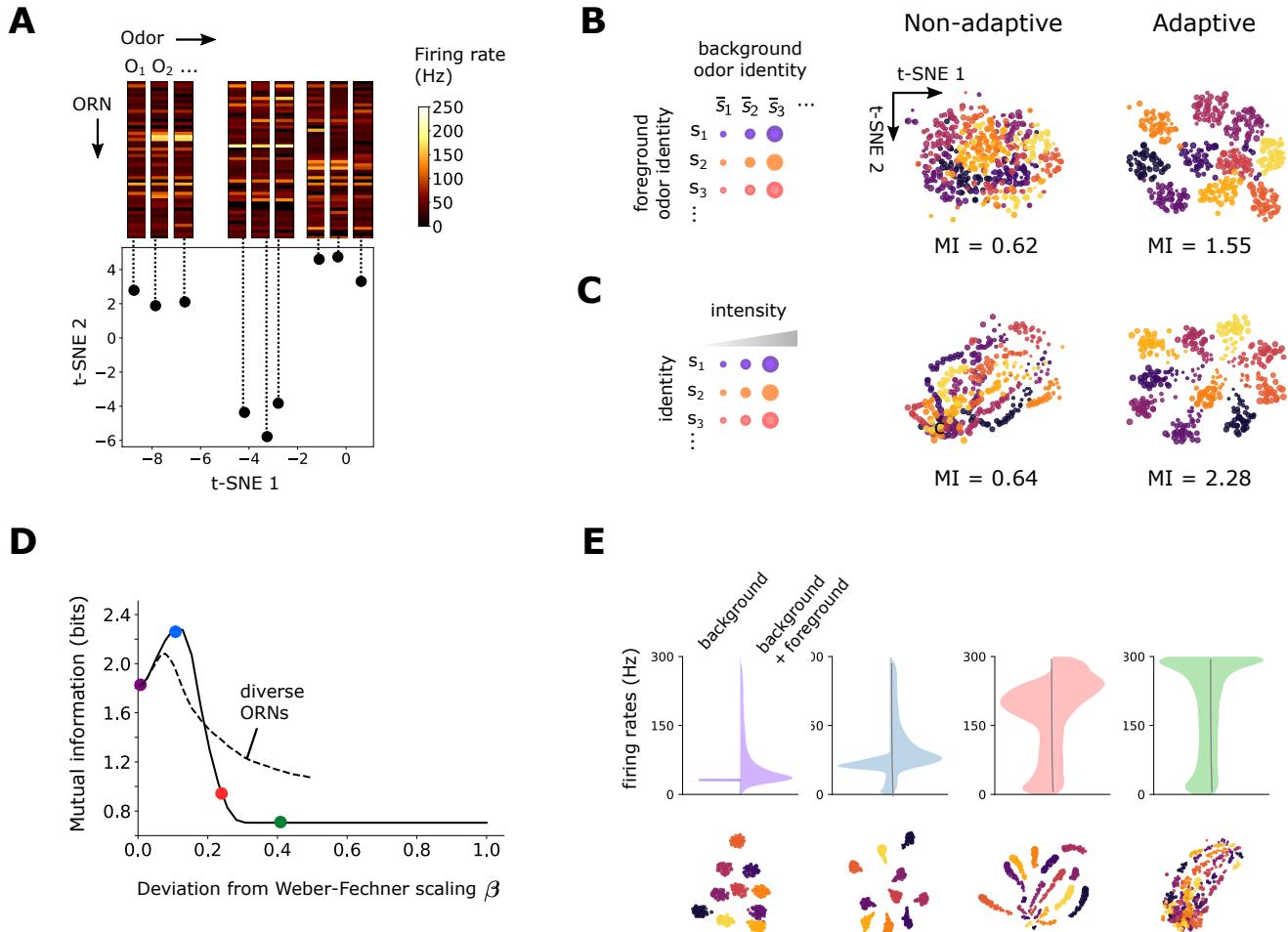


Figure 2. Front-end adaptation maintains representations of odor identity across background and intensity confounds. **A** Example t-SNE projection of the 50-dimensional vector of ORN firing rates to 2 dimensions. Each point represents the firing response to a distinct odor. Nearby points exhibit similarities 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, ϵ_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; **Figure 3–Figure Supplement 1**). Clustering by color implies that responses cluster by foreground odor identity. Since global distances are not preserved by t-SNE, distances between plots cannot be meaningfully compared, and so we do not label the axes with units. **Mutual information, in bits, is indicated below the plots.** **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 β , 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 mutual information between foreground odor and ORN responses in bits (Methods).** Line: same scaling $\beta_a = \beta$ for all ORNs. Dashed: β_a is uniformly distributed between 0 and $2\beta < 1$ (i.e. has mean β). **E** Distribution of ORN responses and t-SNE projections for $\beta = 0, 0.10, 0.22, 0.40$ in (D).

Figure 2–Figure supplement 1. t-SNE projections when background adapted activity level A_{0a} depends on ORN.

Figure 2–Figure supplement 2. Front-end adaptive feedback preserves information capacity of the ORN sensing repertoire.

177 *and Sharpee, 2018*). Thus, we use t-SNE only for visualization. To more rigorously quantify how
 178 foreground identity is preserved in ORN activity, we calculate the mutual information (MI) between
 179 foreground odor s and ORN firing rates r in the 50-dimensional space (Methods). The MI quantifies
 180 how much information a response contains about the stimulus. High MI means that responses
 181 exhibit larger variability for different stimuli than for repeated presentations of the same stimulus.
 182 In our case, this would be true if r were uniquely defined for different foregrounds s , irrespective of
 183 the background \bar{s} . Conversely, the MI would be low if responses varied more by background \bar{s} than
 184 by foreground s .

185 We first examined how an adaptive or non-adaptive ORN repertoire encodes odor identity in an
 186 odor environment that contains a foreground odor s atop a background odor \bar{s} (*Figure 2B*). Both
 187 odors are sparse mixtures, with $K \ll N$ odorants of similar concentrations, odor “identity” being
 188 the particular set of odorants in the mixture. In the adaptive case, we assume that the system
 189 has fully adapted to the background \bar{s} before the foreground s is presented. This is enacted by
 190 calculating the firing response to the foreground odor $r(s)$ only after having set the ϵ_a in Eq. 2 to
 191 their steady state values in response to the background odor \bar{s} :

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

192 where we have introduced the new parameter β_a to allow us to control the scaling of gain
 193 adaptation: for $\beta_a = 0$ the system exactly follows Weber-Fechner’s law, while for $\beta_a = 1$ there is
 194 no adaptation. For small but nonzero β_a , the inverse gain scales sub-linearly (see Methods), and
 195 the adapted activity $A_a(\bar{s})$ increases weakly with background \bar{s} . In experiments, small deviations
 196 from the strict Weber-Fechner scaling on the order of $\beta \approx 0.1$ are observed (see extended figures
 197 in (*Gorur-Shandilya et al., 2017*)).

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

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

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

224 (**Figure 2E**). Note that this effect is not specific to our model. A similar enhancement would be
 225 observed if Weber's Law were maintained, but kicked in only above a minimum signal intensity.
 226 Thus, while Weber-Fechner scaling largely preserves the representation of foreground odor identity
 227 amid backgrounds, in some cases it may benefit from a slight relaxation so that the full dynamical
 228 range of the ORNs can be exploited.

229 **Front-end adaptation enhances odor decoding in complex environments**

230 Given that front-end adaptation helps maintain combinatorial odor codes in the presence of back-
 231 grounds, we wondered how it affects the capability to decode odor signals from ORN response.
 232 One potentially complicating factor is the disparity between sensor dimension and stimulus di-
 233 mension: while *Drosophila* only express ~ 60 Or genes (Vosshall *et al.*, 2000), the space of odorants
 234 is far greater (Krishnamurthy *et al.*, 2017). An N -dimensional odor signal would naively need N
 235 sensory neurons to decode it – one for each odorant. However, naturally-occurring odors are
 236 sparse, typically comprised of only a few odorants. Enforcing sparsity of the signal during decoding
 237 greatly restricts the number of possible odors consistent with a given ORN response, suggesting
 238 that such high-dimensional signals might be inferred from less than N ORNs. Indeed, the decoding
 239 of sufficiently sparse signals from lower-dimensional responses is rigorously guaranteed by the
 240 theory of compressed sensing (CS) (Donoho, 2006; Candes *et al.*, 2006a). It is unknown whether
 241 CS is implemented in the *Drosophila* olfactory circuit (Pehlevan *et al.*, 2017). Here we use this
 242 framework mainly as a tool to quantify how front-end adaptation potentially affects odor decoding,
 243 later verifying our conclusions with other classification techniques that incorporate the known
 244 architecture of the olfactory system.

245 CS is performed as a constrained linear optimization. The constraints in the optimization are
 246 $\mathbf{r} = \mathbf{Ds}$, where \mathbf{s} is the stimulus to be estimated, \mathbf{D} is the response matrix, and \mathbf{r} is the vector of ORN
 247 responses. The cost function to be minimized, $C = \sum_i |s_i|$, enforces sparsity by driving the estimate
 248 of each odorant component to zero; the constraints balance this tendency by simultaneously
 249 enforcing information from the ORN firing rates. The result is a reconstructed odor signal $\hat{\mathbf{s}}$ that is as
 250 sparse as possible, consistent with the ORN responses. In practice, one uses a linear optimization
 251 routine to numerically minimize $\sum_i |s_i|$ over s_i , subject to $\mathbf{r} = \mathbf{Ds}$. The result is an estimate of the
 252 magnitude of each signal component s_i . Thus, both the identity and the intensity of the odor signal
 253 are estimated.

254 To incorporate this linear framework of CS into our nonlinear odor encoding model, we treat
 255 the nonlinear odor encoding exactly, but approximate the decoding to first order around the
 256 background concentration (**Figure 3A**). Specifically, we use Eqs. 2-4 to generate ORN responses \mathbf{r}
 257 for sparse odors \mathbf{s} having $K \ll N$ nonzero components $s_i = \bar{s}_i + \Delta s_i$, where the mean concentration
 258 is \bar{s}_i . To estimate signals using CS, we minimize $\sum_i |\Delta s_i|$ while enforcing the constraints $\mathbf{r} = \mathbf{D}\Delta\mathbf{s}$,
 259 where \mathbf{D} is the linearization of Eq. 2 around \bar{s}_i (details in Methods). The perturbations are chosen
 260 as $\Delta s_i \sim \mathcal{N}(s_0/3, s_0/9)$, where $\bar{s}_i = s_0$. This linearization simplifies the CS decoding – namely it
 261 enforces a single, global minimum – but it is not critical for our general results; see Methods and
 262 **Figure 3–Figure Supplement 5**. We perform the minimization using the sequential least squares
 263 algorithm, producing an estimate of the concentration Δs_i of each individual odorant. The matrix
 264 \mathbf{D} depends on ϵ_a , and as above, we assume precise adaptation by setting ϵ_a to their steady state
 265 values in response to the background odor alone (via Eq. 5 with $\beta = 0$). In the nonadaptive case, ϵ_a
 266 are held at their minimum values $\epsilon_{L,a}$.

267 We first examine how foreground odors are recognized when mixed with background odors of
 268 a distinct identity but similar intensities, quantifying decoding accuracy as the percentage of odors
 269 correctly decoded within some tolerance (see Methods). Without adaptation, accuracy is maintained
 270 within the range of receptor sensitivity for monomolecular backgrounds, but is virtually eliminated
 271 as background complexity rises (**Figure 3B**). The range of sensitivity is broader in the adaptive
 272 system, and is substantially more robust across odor concentration and complexity (**Figure 3C**).

273 In realistic odor environments, the concentration and duration of individual odor whiffs vary

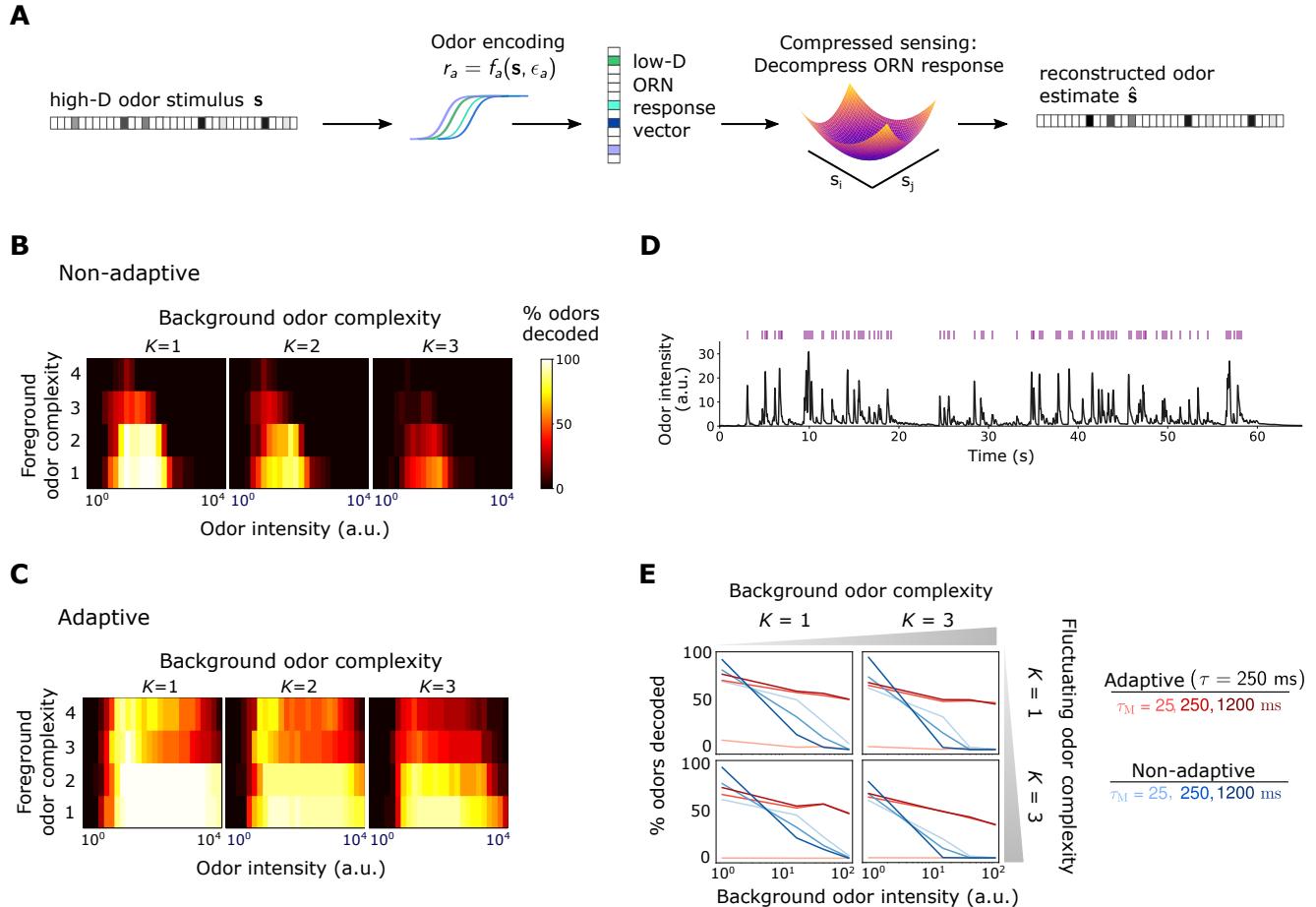


Figure 3. Front-end adaptation promotes accurate odor decoding in static and naturalistic odor environments. **A** Odor stimuli produce ORN responses via odor-binding and activation and firing machinery, as described by Eqs. 2-4. Odors are then decoded using compressed sensing optimization. Odors are 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 τ_M (line shades).

Figure 3-Figure supplement 1. Decoding accuracy for system with ORN-dependent adaptive timescales τ .

Figure 3-Figure supplement 2. Decoding accuracy for system with ORN-dependent adapted firing rates.

Figure 3-Figure supplement 3. Decoding accuracy for receptors with multiple binding sites.

Figure 3-Figure supplement 4. Preservation of restricted isometry property in CS shows how decoding accuracy is maintained by adaptation.

Figure 3-Figure supplement 5. Odor decoding accuracy using the iterative hard thresholding algorithm for nonlinear compressed sensing.

Figure 3-Figure supplement 6. Whiff duration distribution in naturalistic stimulus.

widely (*Celani et al., 2014*). We wondered how well a front-end adaptation mechanism with a single timescale τ could promote odor identity detection in such environments. As inputs to our coding/decoding framework, we apply a naturalistic stimulus intensity recorded from a photo-ionization detector (*Gorur-Shandilya et al., 2017*) (*Figure 3D*), to which we randomly assign sparse identities from the N -dimensional odorant space (odor concentration fluctuates in time, but identity is fixed). To mimic background confounds, we combine these signals with a static odor background of a different identity. We decode the odor at each point in time using CS optimization. To assess performance, we consider decoding accuracy only during odor whiffs, rather than blanks, where the concentration is too low to be perceived. We assess performance by the percentage of correctly decoded whiffs (signal must be fully decoded at some point during the whiff), and average our results over distinct choices of foreground and background identity. Finally, we assume the decoder has short-term memory: detected odor signals are only retained for τ_M seconds in the immediate past, bounding the amount of past information utilized in signal reconstruction.

Without ORN adaptation, sufficiently strong backgrounds eliminate the ability to reconstruct the identity of individual odor whiffs, irrespective of the complexity of either the foreground or background odor (*Figure 3E*, blue lines). In the adaptive system, this is substantially mitigated (red lines in *Figure 3E*), provided the memory duration τ_M is at least as long as the adaptation timescale τ (darker red lines). The memory τ_M must be long enough so that information about the background concentration \bar{s}_i , which is needed for decoding, can be acquired over a window at least as long as the adaptation timescale. Because short-term adaptation depends on the activity of the Or-Orco channel rather than on the identity of the receptor (*Nagel and Wilson, 2011; Martelli et al., 2013; Gorur-Shandilya et al., 2017*), the values of τ and A_{0a} were assumed the same for all ORNs; still, our results hold if these invariances are relaxed (*Figure 3-Figure Supplement 1* and *Figure 3-Figure Supplement 2*).

Front-end adaptation enhances primacy coding

The primacy coding hypothesis has recently emerged as an intriguing framework for combinatorial odor coding. Here, odor identity is encoded by the set (but not temporal order) of the p earliest responding glomeruli/ORN types, known as primacy set of order p (*Wilson et al., 2017*). If the activation order of ORNs were invariant to the strength of an odor step or pulse, primacy sets would in principle form concentration-invariant representation of odor identity. Though our coding framework 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 (*Figure 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 (*Figure 4-Figure Supplement 1*).

Beyond mere consistency, however, front-end adaptation 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, with adaptation, primacy sets are highly consistent across background confounds for all but the smallest primacy orders (*Figure 4B-Figure 4C*). This also holds true for backgrounds of different concentrations (*Figure 4-Figure Supplement 1*), suggesting a central role for front-end adaptation in reinforcing primacy codes across differing environmental conditions.

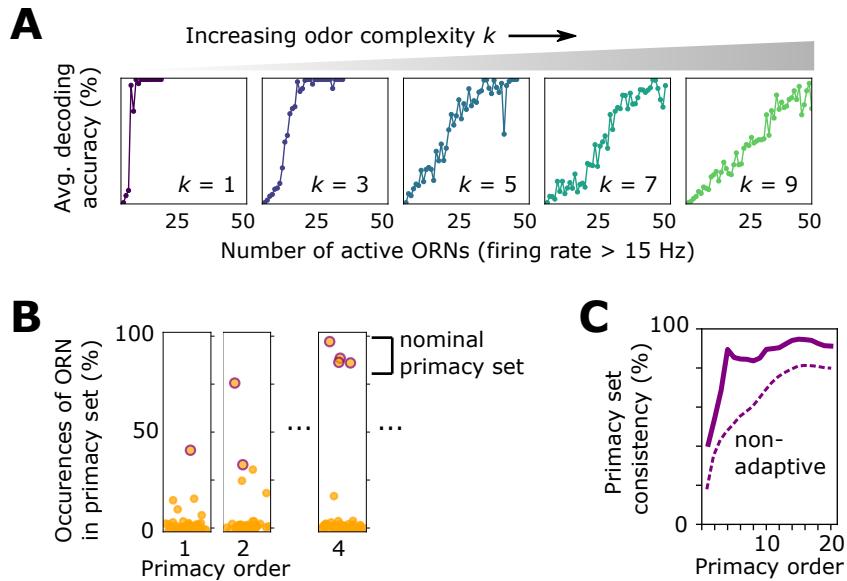


Figure 4. Effect of front-end adaptation on primacy coding. **A** Decoding accuracy as a function of the number of active ORNs, for different odor complexities. The primacy set consists of those ORNs required to be active for accurate decoding. **B** Frequency of particular ORNs in primacy sets of an 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 , for the adaptive (solid) and non-adaptive (dashed) system. 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.

Figure 4–Figure supplement 1. Additional results for primacy coding in the adaptive ORN model.

322 **Contribution of front-end adaptation for odor recognition within the *Drosophila***
 323 **olfactory circuit**

324 Signal transformations in the sensing periphery are propagated through the remainder of the
 325 olfactory circuit. How does front-end adaptation interact with these subsequent neural trans-
 326 formations? ORNs expressing the same OR converge to a unique AL glomerulus, where they receive
 327 lateral inhibition from other glomeruli (*Olsen and Wilson, 2008; Asahina et al., 2009*). This inhi-
 328 bition implements a type of divisive gain control (*Olsen et al., 2010*), normalizing the activity of
 329 output projections neurons, which then synapse onto a large number of Kenyon cells (KCs) in the
 330 mushroom body. To investigate how odor representations are affected by interactions between
 331 front-end ORN adaptation and this lateral inhibition and synaptic divergence, we extended our ORN
 332 encoding model by adding uniglomerular connections from ORNs to the antennal lobe, followed by
 333 sparse, divergent connections to 2500 KCs (*Keene and Waddell, 2007; Litwin-Kumar et al., 2017;*
 334 *Caron et al., 2013*). Inhibition was modeled via divisive normalization, with parameters chosen
 335 according to experiment (*Olsen et al., 2010*). We quantified decoding accuracy by training and
 336 testing a linear classifier on the KC activity output of sparse odors of distinct intensity and identity.
 337 We trained the classifier on N_{ID} sparse odor identities at intensities chosen randomly over 4 orders
 338 of magnitude, then tested the classifier accuracy on the same set of odor identities but of differing
 339 concentrations.

340 With both ORN adaptation and divisive normalization, the accuracy of the classification by
 341 odor identity remains above 85% for more than 1000 odor identities ($N_{ID} > 1000$). Removing ORN
 342 adaption while maintaining divisive normalization substantially reduces accuracy (down to 65%
 343 for 1000 odor identities). Further removing divisive normalization gives similar results, apart for
 344 very large numbers of odors identities ($N_{ID} > 1000$), where divisive normalization provides benefits
 345 (*Figure 5A*). These results strongly implicate front-end adaptation as a key player in maintaining

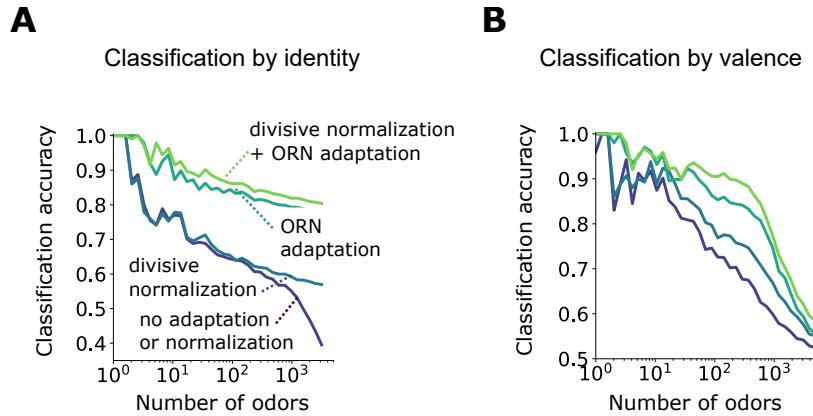


Figure 5. Front-end adaptation enhances odor recognition by the *Drosophila* olfactory circuit. **A** Accuracy of linear classification by odor identity, as a function of the number of distinct odor identities classified by the trained network (concentrations span 4 orders of magnitude), in systems with only ORN adaptation, only divisive normalization, both or neither. **B** Same as (A) but now classifying odors by valence. Odors were randomly assigned valence. For a given odor identity the valence is the same for all concentrations.

Figure 5–Figure supplement 1. Binary classification for odors whose concentrations span a narrow range of concentration.

346 odor identity representations, before signals are further processed downstream.

347 As a simpler task, we also considered binary classification, categorizing odors as appetitive or
 348 aversive. For simplicity, odor signals of the same identity but differing intensity were assigned
 349 the same valence. Classification accuracy degrades to chance level as N_{ID} becomes very large
 350 (**Figure 5B**). When acting alone, either divisive normalization or ORN adaptation can help, although
 351 the effect of ORN adaptation is slightly stronger. When both are active, accuracy improves further,
 352 suggesting that these distinct adaptive transformations may act jointly at different stages of neural
 353 processing in preserving representations of odor identity. As expected, these gains mostly vanish for
 354 the same odors chosen from a narrower range of concentrations (**Figure 5–Figure Supplement 1**).

355 Previous simulation results have shown that divisive normalization aids identity decoding
 356 from PN response to a stronger degree than we find here (*Olsen et al., 2010*). There, 19 distinct
 357 odor identities at three concentrations were classified more accurately with divisive normalization
 358 (80%) than without (68%). In our case, we find about ~ 75% accuracy, with and without divisive
 359 normalization. This discrepancy is not necessarily inconsistent. First, we decode mixtures, not
 360 single odorants, and the combinatorics may reduce the benefit of divisive normalization. Second,
 361 we classify the responses of 2500 KCs, rather than 50 PNs (or 24 PNs as in *Olsen et al. (2010)*).
 362 Kenyon cell responses follow a high degree of postsynaptic divergence from PNs, which could
 363 decorrelate neural responses (*Caron et al., 2013; Litwin-Kumar et al., 2017; Krishnamurthy et al., 2017*) similarly to divisive normalization, reducing the gains from the latter. Finally, the divisive
 364 normalization model is a simple one in which glomeruli are all mutually inhibiting. A more complex
 365 model in which each glomerulus inhibits only a subset of other glomeruli through local neurons
 366 might produce a larger contribution.

368 In sum, these results indicate that ORN adaptation might contribute significantly to odor recogni-
 369 tion by identity and valence, and that divisive normalization also contributes, though possibly
 370 more to classification by valence than by identity. An intriguing possibility is that these two forms of
 371 gain control play different roles in coding discrete odor categories versus odor identities.

372 Discussion

373 Weber-Law adaptation at the very front-end of the insect olfactory circuit (*Gorur-Shandilya et al., 2017; Cafaro, 2016; Cao et al., 2016*) may contribute significantly to the preservation of neural
 374 representations of odor identity amid confounding odors and intensity fluctuations. Drawing on
 375 experimental evidence for a number of ORN-invariant response features (*Nagel and Wilson, 2011; Martelli et al., 2013; Stevens, 2016; Gorur-Shandilya et al., 2017; Si et al., 2019*), we have found
 376 that this mechanism of dynamic adaptation confers significant benefits in coding fidelity, without
 377 the need for ORN-specific parameterizations. Still, our results hold when these invariances such as
 378 adaptation timescale or baseline activity are relaxed (*Figure 3-Figure Supplement 1 and Figure 3-Figure Supplement 2*). In the olfactory periphery, front-end Weber Law adaptation therefore
 379 appears fairly robust, a consequence of controlling gain via feedback from channel activity (*Waite et al., 2018; Nagel and Wilson, 2011; Gorur-Shandilya et al., 2017*), rather than through intrinsic,
 380 receptor-dependent mechanisms.

381 Our results also suggest that a slight breaking of Weber scaling may aid combinatorial coding,
 382 by spreading firing rates more fully over the ORN dynamic range, while still preventing saturation.
 383 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
 384 small deviations from the strict Weber-Fechner scaling have been observed (see extended figures
 385 in (*Gorur-Shandilya et al., 2017*)).

386 While our framework incorporates many observed features of the *Drosophila* olfactory system –
 387 Weber-Law adaptation, power-law distributed receptor affinities, temporal filter invariance, connec-
 388 tivity topologies – it is minimal. We considered only one of the chemoreceptor families expressed
 389 in the fly antenna (*Joseph and Carlson, 2015*) and ignored possible contributions of odor bind-
 390 ing proteins (*Vogt and Riddiford, 1981; Menuz et al., 2014*), inhibitory odorants (*Cao et al., 2017*),
 391 and odorant-odorant antagonism (*Reddy et al., 2018*), which could further boost coding capacity
 392 and preserve representation sparsity. Useful extensions to our nonlinear-linear-nonlinear model
 393 might incorporate ephaptic coupling between ORNs housed in the same sensillum (*Su et al., 2012*),
 394 global inhibition in the mushroom body (*Papadopoulou et al., 2011*), and the effects of long-term
 395 adaptation (*Guo et al., 2017*).

396 Previous studies have characterized various neural mechanisms that help preserve combi-
 397 natorial codes. Lateral inhibition between glomeruli helps tame saturation and boost weak sig-
 398 nals (*Olsen et al., 2010*). The sparse degree of connectivity to either the olfactory bulb (vertebrates)
 399 or mushroom body (insects) may also be precisely tuned to optimize the capacity to learn associa-
 400 tions (*Litwin-Kumar et al., 2017*). In this work, we find that some of these downstream features act
 401 in concert with front-end dynamic adaptation in maintaining representations of odor identity.

402 Other studies have implicated the unique temporal patterns of neural response as signatures
 403 of odor identity (*Raman et al., 2010; Gupta and Stopfer, 2011; Brown et al., 2005; Gupta and*
404 Stopfer, 2014). ORN and projection neuron time traces form distinct trajectories in low-dimensional
 405 projections, and cluster by odor identity, much as we have found here for static responses at
 406 different concentrations (*Figure 2*). In locusts PNs, the trajectories elicited by foreground odors
 407 when presented in distinct backgrounds exhibit some degree of overlap; though partial, these
 408 overlaps were nonetheless sufficient to maintain background-invariant decoding from Kenyon
 409 cell responses (*Saha et al., 2013*). It was therefore suggested that background filtering likely
 410 occurs at the level of ORNs themselves (*Saha et al., 2013*). Likewise, in our framework, temporal
 411 coding is implicit: because the input nonlinearity depends on the diversity of binding affinities,
 412 odor signals are naturally formatted into temporal patterns that are both odor- and ORN-specific
 413 (*Figure 1D-Figure 1E*). Further, the short required memory timescales ($\tau_M \sim \tau \sim 250$ ms) suggest
 414 that only brief time windows are needed for accurate odor identification, consistent with previous
 415 findings (*Brown et al., 2005; Saha et al., 2013*). Moreover, we find that front-end adaptation
 416 enhances the robustness of other combinatorial coding schemes, such as primacy coding (*Wilson*
 417

422 *et al., 2017*), which relies on the temporal order of ORN activation but not absolute firing rate
 423 (**Figure 4**).

424 In the well-characterized chemosensory system of bacterial chemotaxis, Weber Law adaptation
 425 is enacted through a feedback loop from the output activity of the receptor-kinase complexes onto
 426 the enzymes modifying receptor sensitivity (*Waite et al., 2018*). It is interesting that some aspects
 427 of this logic are also present in ORNs: although the molecular players are different (and still largely
 428 unknown, though likely involving calcium channel signaling (*Cao et al., 2016*)), it has been shown
 429 that transduction activity feeds back onto the sensitivity of Or-Orco ligand-gated cation channels,
 430 enabling the Weber-Fechner relation (*Nagel and Wilson, 2011; Gorur-Shandilya et al., 2017; Cao*
et al., 2016). That this adaptation mechanism appears to act similarly across ORNs (*Gorur-Shandilya*
et al., 2017; Martelli et al., 2013; Cao et al., 2016) suggests the possible involvement of the universal
 431 co-receptor Orco, whose role in long-term adaptation has recently been reported (*Getahun et al.,*
2013, 2016; Guo et al., 2017). Further, the identification of 4 subunits comprising the Orco-Or ion
 432 channel suggest that generic Or/Orco complexes may contain multiple odorant binding sites, which
 433 when included in our model supports our general findings (**Figure 3–Figure Supplement 3**).

434 Weber Law ensures that sensory systems remain in the regime of maximum sensitivity, broad-
 435 ening dynamic range and maintaining information capacity (*Wark et al., 2007*). For a single-channel
 436 system, this requires matching the midpoint of the dose-response curve to the mean ligand con-
 437 centration (*Nemenman, 2012*), a strategy which may fail in multi-channel systems with overlapping
 438 tuning curves: adaptation to one signal could inhibit identification of others, if the signals excite
 439 some but not all of the same sensors, as in **Figure 1G**. Our results show that this strategy is still
 440 largely functional. In CS decoding, this can be traced to the observation that accuracy is guaranteed
 441 when sufficiently distinct odor identities produce sufficiently distinct ORN responses, a condition
 442 known as the restricted isometry property (*Candes et al., 2006a*). Indeed, the Weber-Fechner
 443 scaling increases the likelihood that this property is satisfied, beyond that in the non-adaptive
 444 system (SI text and **Figure 3–Figure Supplement 4–Figure 3–Figure Supplement 5**). Still, restricted
 445 isometry does not require that response repertoires are *invariant* to environmental changes. That
 446 is, even if the subset of active ORNs were concentration-dependent, odors could still in principle be
 447 fully reconstructible by CS. Nonetheless, our results in t-SNE clustering (**Figure 2**), primacy coding
 448 (**Figure 4B–4C**), and odor classification (**Figure 5**) suggest that some signature of response invariance
 449 emerges as a natural byproduct of front-end adaptation. Together, this implies that Weber Law
 450 adaptation, whether required by the olfactory circuit for precise signal reconstruction (as in CS)
 451 or for developing odor associations (as in classification), can play an integral part in maintaining
 452 combinatorial codes amid changing environmental conditions.

456 Methods and Materials

457 Adaptive ORN model

458 We model an odor as an N -dimensional vector $\mathbf{s} = [s_1, \dots, s_N]$, where $s_i > 0$ are the concentrations
 459 of individual volatile molecules (odorants) comprising the odor. The olfactory sensory system is
 460 modeled as a collection of M distinct Or/Orco complexes indexed by the sub index $a = 1, \dots, M$,
 461 each of which can be bound with any one of the odorant molecules, and can be either active
 462 (firing) or inactive (quiescent). At first we assume there is one binding site per complex; this will be
 463 generalized to many sites. We consider the binding and activation processes to be in equilibrium,
 464 assigning each state a corresponding Boltzmann weight, where the zero of energy is set by the
 465 unbound, inactive state C_a . These weights are:

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

466 where ϵ_a (assumed positive) is the free energy difference between the active and inactive con-
467 formation of the unbound receptor, and E_{ai} and E_{ai}^* are the free energy differences (assumed
468 positive) between the unbound and bound state for the inactive and active receptor, respec-
469 tively. $\mu_i = \mu_0 + \beta^{-1} \log(s_i/s_0)$ is the chemical potential for odorant species i in terms of a reference
470 chemical potential μ_0 at concentration s_0 , $s_0 \exp(-\beta\mu_0) = s_i \exp(-\beta\mu_i)$, which can be traded for the
471 thermodynamic-relevant disassociation constants $K_{ai}^{-1} = K_{D,ai} = s_0 e^{\beta(-E_{ai} - \mu_0)}$.

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

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

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

474 This expression can be generalized for the case of multiple, independent binding sites through
475 some simple combinatorial factors. Consider first an odorant i which can bind one of two locations
476 on receptor a . There are then 4 possible inactive states: both sites unbound, site 1 bound, site 2
477 bound, both sites bound. Combined with the active states, there are therefore 8 states for odorant
478 i and receptor a , with energies:

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

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

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

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

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

483 To generate ORN time traces, equations 2-3 are integrated numerically using the Euler method
484 with a 2 ms time step. For ORN firing (Eq. 4), $h(t)$ is bi-lobed (**Martelli et al., 2013**): $h(t) =$
485 $A p_{\text{Gam}}(t; \alpha_1, \tau_1) - B p_{\text{Gam}}(t; \alpha_2, \tau_2)$, $A = 190$, $B = 1.33$, $\alpha_1 = 2$, $\alpha_2 = 3$, $\tau_1 = 0.012$, and $\tau_2 = 0.016$, where
486 p_{Gam} is the pdf of $\text{Gamma}(\alpha, 1/\tau)$. Nonlinearity f is modeled as a linear rectifier with 5 Hz threshold.

487

488 **Derivation of ORN gain**

489 Weber's Law states that the gain, or differential response, of the receptor activity A_a scales with the
 490 mean odor concentration \bar{s}_i . To show how this is satisfied in our model, we consider the response,
 491 Eq. 2, to a signal $s = \bar{s} + \Delta s$, where Δs consists of only a small fluctuation in the i th component
 492 $\Delta s_i < |\bar{s}_i|$ about the mean. We derive the change in response to fluctuation Δs_i for general β from 0
 493 (Weber's law) to 1 (no adaptation).

494 First we write the activity in the form:

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

495 where

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

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

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

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

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

499 whereby

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

500 For $\beta = 0$ (the fully adaptive case) and a single odorant, this expression for the gain reduces to
 501 $(1 - A_{0a}) A_{0a} / s_i$. For small β , and given $A_{0a} \simeq 0.1$ (corresponding to 30 Hz on a 300 Hz firing rate scale),
 502 the denominator is dominated by the $1 - A_{0a}$ term, giving:

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

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

506

507 **t-SNE dimensionality reduction and mutual information**

508 For t-SNE dimensionality reduction (*van der Maaten and Hinton, 2008*), ORN responses were
 509 generated for odor signal combinations consisting of 1 (among 10) distinct sparse foreground odors
 510 A atop 1 (among 50) distinct sparse background odors B, for **Figure 2B**. **Figure 2C** plots responses
 511 for 10 odors at 40 concentrations spanning 4 decades, atop a random sparse background odor of
 512 similar magnitude. For adaptive systems, ϵ_a were set to their fully adapted values to the background
 513 odor, given by Eq. 5, with $\beta = 0$.

514 The mutual information (MI) between signal and response quantifies how many bits of informa-
 515 tion a response contains about the stimulus. As we are interested in how much information ORN
 516 responses \mathbf{r} contain about novel foreground odors \mathbf{s} , we calculate the MI between \mathbf{s} and \mathbf{r} . This
 517 calculation requires the conditional response distribution $P(\mathbf{r}|\mathbf{s})$, where the probability distribution
 518 is over different background odors $\bar{\mathbf{s}}$. To get this distribution, we hold \mathbf{s} fixed and generate \mathbf{r} in the
 519 presence of distinct backgrounds. To these responses \mathbf{r} , we also add a small amount of Gaussian
 520 noise (mean zero and variance 1 Hz), which allows a distribution to be defined when there is no
 521 background odor. We then bin the resulting \mathbf{r} in units of $\Delta r = 1$ Hz to get a histogram representing
 522 $P(\mathbf{r}|\mathbf{s})$ (The histogram is necessary, since a sum must be taken over \mathbf{r}). If responses were completely
 523 background invariant, the resulting histogram would be highly peaked.

524 Using $P(\mathbf{r}|\mathbf{s})$, the MI is defined as

$$\text{MI} = H_{\text{response}} - H_{\text{noise}}$$

where H_{noise} is:

$$H_{\text{noise}} = - \sum_{\mathbf{s}, \mathbf{r}} P(\mathbf{s})P(\mathbf{r}|\mathbf{s}) \log_2 P(\mathbf{r}|\mathbf{s})$$

525 and H_{response} is

$$H_{\text{response}} = - \sum_{\mathbf{r}} P(\mathbf{r}) \log_2 P(\mathbf{r})$$

526 where

$$P(\mathbf{r}) = \sum_{\mathbf{s}} P(\mathbf{s})P(\mathbf{r}|\mathbf{s}).$$

527 The noise entropy H_{noise} quantifies how much variability comes from the background odors, but
 528 is not related to changes in foreground odor. The response entropy H_{response} quantifies how much
 529 variability comes from both background and foreground. The mutual information, which is their
 530 difference, is a measure of how responses differ by foreground alone.

531 **Compressed sensing decoding of ORN responses**

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

$$\mathbf{y} = \mathbf{D}\mathbf{x}, \quad (16)$$

535 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
 536 measurement matrix. Since measurements are fewer than signal components, then $M < N$,
 537 whereby \mathbf{D} is wide rectangular and so Eq. 16 cannot be simply inverted to produce \mathbf{x} . The idea of CS
 538 is to utilize the knowledge that \mathbf{x} is sparse, i.e. only K of its components, $K \ll N$ are nonzero. Both
 539 the measurements and sparsity are thus combined into a single constrained optimization routine:

$$\hat{\mathbf{x}}_i = \operatorname{argmin}_i \sum_i^N |x_i| \quad \text{such that } \mathbf{y} = \mathbf{D}\mathbf{s} \quad (17)$$

540 where \hat{x}_i are the optimal estimates of the signal components and the sum, which is known as
 541 the L_1 norm of \mathbf{x} , is a natural metric of sparsity (*Donoho, 2006*).

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

$$\begin{aligned} \Delta r_a(t) &= r_a(\mathbf{s}(t)) - r_a(\bar{\mathbf{s}}(t)) \\ \Delta r_a(t) &= \int^t d\tau h(t-\tau) \sum_i^N \frac{dA_{ai}}{ds_i} \Big|_{\bar{\mathbf{s}}} \Delta s_i \end{aligned} \quad (18)$$

547 where

$$r_a(\mathbf{s}_0) = \int^t d\tau h(t-\tau) \sum_i^N A_{0a} \quad (19)$$

548 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. 18 and 19 hold only
 549 for integrands above 5 Hz (and are zero below), as per the linear rectifier f . We assume that the
 550 neural decoder has access to background $\bar{\mathbf{s}}$, presumed learned (this assumption can be relaxed;
 551 see below), and to the linearized response matrix, Eq. 14, but must infer the excess signals Δs_i
 552 from excess ORN firing rates $\Delta r_a(t)$. Thus, this corresponds to the CS framework (Eq. 17) via $\Delta\mathbf{r} \rightarrow \mathbf{y}$,
 553 $\Delta\mathbf{s} \rightarrow \mathbf{x}$, and $dA_{ai}/ds_i \Big|_{\bar{\mathbf{s}}} \rightarrow \mathbf{D}$. We optimize the cost function in Eq. 17 using sequential least squares
 554 programming, implemented in Python through using the scientific package SciPy.

555 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
 556 $\Delta s_i \sim \mathcal{N}(s_0/3, s_0/9)$. The measurement matrix \mathbf{D} depends on the free energy differences ϵ_a . For
 557 static stimuli, ϵ_a equals the fixed point of Eq. 3 in response to the background stimulus with $\beta = 0$. For
 558 fluctuating stimuli, ϵ_a is updated in time by continuously integrating $r_a(t)$, via Eqs. 3 and 4; thus,
 559 only knowledge of the response $r_a(t)$ are needed by the decoder. To quantify decoding accuracy,
 560 we treat the zero and nonzero components of the sparse odor vector separately. We demand that
 561 the K nonzero components \hat{s}_i of the estimated sparse vector are within 25% of their true values s_i ,
 562 and that the $N - K$ zero components are estimated less than 10% of s_0 . Together, this ensures that
 563 the odorants comprising the odor mixture are estimated sufficiently close to their concentrations,
 564 and that the remaining components are sufficiently small. Odor signals \mathbf{s} are considered correctly
 565 decoded if both of these conditions are satisfied for all components s_i . The relatively lax accuracy
 566 demanded on the nonzero components is to prevent oversensitivity on the unavoidable errors
 567 introduced by linearization. Qualitatively, our findings are robust to these choices.

568 The naturalistic odor signal (*Figure 3D*) was generated by randomly varying flow rates of ethyl
 569 acetate and measuring the concentration with a photo-ionization detector (*Gorur-Shandilya et al.,
 570 2017*). Statistics mirroring a turbulent flow (*Celani et al., 2014*) were verified (*Figure 3-Figure Sup-
 571 plement 6*).

572

573 Iterative hard thresholding (IHT) and the restricted isometry property in compressed 574 sensing

575 The purpose of response linearization (Eq. 18) is simply to apply compressed sensing reconstruction
 576 directly using linear programming, without worrying about issues of local minima in Eq. 17. This

allows us to isolate the impact of Weber Law adaptation from the particularities of the numerics. An alternate technique for compressed signal reconstruction, *iterative hard thresholding* (IHT), does not minimize the constrained L_1 norm directly, rather applying a hard threshold to an iteratively updated signal estimate (Blumensath and Davies, 2009b). IHT can be generalized straightforwardly to nonlinear constraints, and would actually dispense with the need for a learned background \bar{s} , simply initializing the iterations from $\bar{s} = \mathbf{0}$. Remarkably, this technique works quite well even for nonlinear measurements (Blumensath, 2013). We demonstrate the applicability of the IHT algorithm to our odor decoding system in **Figure 3–Figure Supplement 5**, which reproduces qualitatively the findings in the main text. For these calculations, no background odor was assumed, each iterative decoding being initialized $\bar{s} = \mathbf{0}$.

IHT provides an alternate computational technique of nonlinear CS, which could be used to both extend and verify our results. Further, it allows us to illustrate why Weber Law adaptation maintains signal reconstruction fidelity in our olfactory sensing model. Like CS using L_1 -norm minimization, IHT exhibits amenable reconstruction and convergence properties under the guarantee of the so-called restricted isometry property (RIP) (Candes et al., 2006b). Loosely, RIP measures how closely a matrix operator resembles an orthogonal transformation when acting on sparse vectors. The degree to which RIP is satisfied can be understood in terms of the spectrum of a measurement matrix \mathbf{D} . In particular, if λ_i are the eigenvalues of $\mathbf{D}_i^T \mathbf{D}_i$, where \mathbf{D}_i is any $k \times m$ submatrix of \mathbf{D} , and

$$1 - \delta_i \leq \lambda_{\min} \leq \lambda_{\max} \leq 1 + \delta_i \quad (20)$$

is satisfied for some δ_i , then \mathbf{D} satisfies the RIP with constant δ_i . Plainly, the RIP states that the eigenvalues of $\mathbf{D}_i^T \mathbf{D}_i$, when acting on k -sparse vectors, are centered around 1. Thus, to intuit why signal reconstruction breaks down in the non-adaptive sensing system, we can investigate the eigendecomposition of various linearizations of the measurement matrix. We do this now, starting with a brief description of the IHT.

In the linear setting, IHT seeks sparse signals via the following iterative procedure (Blumensath and Davies, 2009b):

$$\mathbf{s}_{i+1} = H_K(\mathbf{s}_i + \mu \mathbf{D}^T (\mathbf{s}_i + (\mathbf{y} - \mathbf{D}\mathbf{s}_i))) \quad (21)$$

where \mathbf{s}_i is the i th estimate of the sparse signal \mathbf{s} , μ is a step size for the iterations, and \mathbf{y} , \mathbf{D} are as defined above. $H_k(\cdot)$ is a thresholding function which sets all but the largest K values of its argument to zero. The nonlinear extension to IHT is (Blumensath, 2013):

$$\mathbf{s}_{i+1} = H_K(\mathbf{s}_i + \mu \mathbf{D}_{\mathbf{s}_i}^T (\mathbf{s}_i + (\mathbf{y} - D(\mathbf{s}_i)))), \quad (22)$$

where D is a nonlinear sensing function and $\mathbf{D}_{\mathbf{s}_i}$ is a linearization of D about the point \mathbf{s}_i . Reconstructibility for k -sparse signals is guaranteed if $\mathbf{D}_{\mathbf{s}_i}$ satisfies RIP for all \mathbf{s}_i and all k -sparse vectors (Blumensath and Davies, 2009b). To get a sense of how this is preserved in the adaptive system, we calculate the eigenvalues for 1000 choices of \mathbf{s}_i , acting on random signals of given sparsity K (**Figure 3–Figure Supplement 4**). Since the RIP is sensitive to constant scalings of the measurement matrix (while the actual estimation problem is not), we scaled all columns of $\mathbf{D}_{\mathbf{s}_i}$ to norm unity (Blumensath and Davies, 2009a). This normalizes the eigenvalues of $\mathbf{D}_{\mathbf{s}_i}^T \mathbf{D}_{\mathbf{s}_i}$ to center near unity before calculating the eigendecomposition, allowing us to assess the degree to which the RIP is satisfied. This scaled matrix can be used directly in Eq. 22 (Blumensath, 2013; Blumensath and Davies, 2009a). The spectra of these matrices indicates that the RIP becomes far more weakly satisfied in the non-adaptive system than in the adaptive one, for sufficient odor complexity and intensity.

615

616 **Network model and classification**

617 For the network model, the AL-to-MB connectivity matrix \mathbf{J}_1 , is chosen such that each KC connects
 618 pre-synaptically to 7 randomly chosen AL glomeruli (*Litwin-Kumar et al., 2017; Caron et al., 2013*).
 619 The results shown in *Figure 5* are an average of 10 distinct instantiations of this random topology.
 620 The $Z = 2500$ KCs are then connected by a matrix \mathbf{J}_2 to a readout layer of dimension Q , where $Q = 2$
 621 for binary and $Q = N_{ID}$ for multi-class classification. Both AL-to-MB and MB-to-readout connections
 622 are perceptron-type with rectified-linear thresholds. The weights of \mathbf{J}_1 and \mathbf{J}_2 are chosen randomly
 623 from $\sim \mathcal{N}(0, 1/\sqrt{7})$ and $\sim \mathcal{N}(0, 1/\sqrt{Z})$, respectively. Only the \mathbf{J}_2 and the MB-to-output thresholds
 624 are updated during supervised network training, via logistic regression (for binary classification) or
 625 its higher-dimensional generalization, the softmax cross entropy (for multi-class classification).

626 **Acknowledgements**

627 NK was supported by a postdoctoral fellowship through the Swartz Foundation and by an NRSA
 628 postdoctoral fellowship through the NIH BRAIN Initiative under award number 1F32MH118700.
 629 TE was supported by NIH R01 GM106189. We thank Damon Clark, John Carlson, Mahmut Demir,
 630 Srinivas Gorur-Shandilya, Henry Mattingly, and Ann Hermunstad for comments on the manuscript.

631 **References**

- 632 Asahina, K., Louis, M., Piccinotti, S., and Vosshall, L. (2009). A circuit supporting concentration-invariant odor
 633 perception in *Drosophila*. *Journal of Biology*, 8(1):9.
- 634 Blumensath, T. (2013). Compressed sensing with nonlinear observations and related nonlinear optimization
 635 problems. *IEEE Transactions on Information Theory*, 69(6).
- 636 Blumensath, T. and Davies, M. E. (2009a). How to use the iterative hard thresholding algorithm. In *Proceedings of
 637 SPARS'09 - Signal Processing with Adaptive Sparse Structured Representations* (Saint Malo, France, April 2009).
- 638 Blumensath, T. and Davies, M. E. (2009b). Iterative hard thresholding for compressed sensing. *Applied and
 639 Computational Harmonic Analysis*, 27(3):265 – 274.
- 640 Brown, S. L., Joseph, J., and Stopfer, M. (2005). Encoding a temporally structured stimulus with a temporally
 641 structured neural representation. *Nature Neuroscience*, 8:1568–1576.
- 642 Buck, L. and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor
 643 recognition. *Cell*, 65(1):175–187.
- 644 Butterwick, J. A., del Marmol, J., Kim, K. H., Kahlson, M. A., Rogow, J. A., walz, T., and Ruta, V. (2018). Cryo-em
 645 structure of the insect olfactory receptor orco. *Nature*, 560:447–452.
- 646 Cafaro, J. (2016). Multiple sites of adaptation lead to contrast encoding in the *Drosophila* olfactory system. *Physiological Reports*, 4(4):e12762.
- 648 Candes, E., Romberg, J., and Tao, T. (2006a). Stable signal recovery from incomplete and inaccurate measure-
 649 ments. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223.
- 650 Candes, E., Romberg, J., and Tao, T. (2006b). Stable signal recovery from incomplete and inaccurate measure-
 651 ments. *Communications on Pure and Applied Mathematics*, LIX:1207–1223.
- 652 Cao, L.-H., Jing, B.-Y., Yang, D., Zeng, X., Shen, Y., Tu, Y., and Luo, D.-G. (2016). Distinct signaling of *Drosophila*
 653 chemoreceptors in olfactory sensory neurons. *Proceedings of the National Academy of Sciences*, 113(7):E902–
 654 E911.
- 655 Cao, L.-H., Yang, D., Wu, W., Zeng, X., Jing, B.-Y., Li, M.-T., Qin, S., Tang, C., Tu, Y., and Luo, D.-G. (2017). Odor-evoked
 656 inhibition of olfactory sensory neurons drives olfactory perception in *Drosophila*. *Nature Communications*,
 657 8(1):1357.
- 658 Cardé, R. T. and Willis, M. A. (2008). Navigational strategies used by insects to find distant, wind-borne sources
 659 of odor. *Journal of Chemical Ecology*, 34(7):854–866.
- 660 Caron, S., Ruta, V., Abbott, L., and Axel, R. (2013). Random convergence of olfactory inputs in the *Drosophila*
 661 mushroom body. *Nature*, 497(4774):113–117.

- 662 Celani, A., Villermaux, E., and Vergassola, M. (2014). Odor landscapes in turbulent environments. *Phys. Rev. X*,
663 4:041015.
- 664 Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J., and Carlson, J. R. (1999). A novel family of divergent
665 seven-transmembrane proteins: Candidate odorant receptors in drosophila. *Neuron*, 22(2):327 – 338.
- 666 Dasgupta, S., Stevens, C. F., and Navlakha, S. (2017). A neural algorithm for a fundamental computing problem.
667 *Science*, 358(6364):793–796.
- 668 Davies, A., Louis, M., and Webb, B. (2015). A model of drosophila larva chemotaxis. *PLOS Computational Biology*,
669 11(11):1–24.
- 670 de Bruyne, M., Foster, K., and Carlson, J. R. (2001). Odor coding in the drosophila antenna. *Neuron*, 30(2):537 –
671 552.
- 672 Donoho, D. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306.
- 673 Fechner, G. T. (1860). *Elemente der psychophysik*. Breitkopf und Härtel.
- 674 Friedrich, R. W. and Korschning, S. I. (1997). Combinatorial and chemotopic odorant coding in the zebrafish
675 olfactory bulb visualized by optical imaging. *Neuron*, 18(5):737–752.
- 676 Getahun, M. N., Olsson, S. B., Lavista-Llanos, S., Hansson, B. S., and Wicher, D. (2013). Insect odorant response
677 sensitivity is tuned by metabotropic autoregulated olfactory receptors. *PLoS One*, 8(3):e58889.
- 678 Getahun, M. N., Thoma, M., Lavista-Llanos, S., Keesey, I., Fandino, R. A., Knaden, M., Wicher, D., Olsson, S. B., and
679 Hansson, B. S. (2016). Intracellular regulation of the insect chemoreceptor complex impacts odour localization
680 in flying insects. *Journal of Experimental Biology*, 219(21):3428–3438.
- 681 Gorur-Shandilya, S., Demir, M., Long, J., Clark, D. A., and Emonet, T. (2017). Olfactory receptor neurons use gain
682 control and complementary kinetics to encode intermittent odorant stimuli. *eLife*, 6:e27670.
- 683 Guo, H., Kunwar, K., and Smith, D. (2017). Odorant receptor sensitivity modulation in *Drosophila*. *The Journal of
684 Neuroscience*, 37(39):9465–9473.
- 685 Guo, H. and Smith, D. P. (2017). Odorant receptor desensitization in insects. *Journal of Experimental Neuroscience*,
686 11:1–5.
- 687 Gupta, N. and Stopfer, M. (2011). Insect olfactory coding and memory at multiple timescales. *Current Opinion in
688 Neurobiology*, 21:768–773.
- 689 Gupta, N. and Stopfer, M. (2014). A temporal channel for information in sparse sensory coding. *Neuron*, pages
690 2247–2256.
- 691 Hallem, E. and Carlson, J. (2006). Coding of odors by a receptor repertoire. *Cell*, 125(1):143–160.
- 692 Hildebrand, J. G. and Shepherd, G. M. (1997). Mechanisms of olfactory discrimination: converging evidence for
693 common principles across phyla. *Annual review of neuroscience*, 20(1):595–631.
- 694 Joseph, R. M. and Carlson, J. R. (2015). Drosophila chemoreceptors: A molecular interface between the chemical
695 world and the brain. *Trends in Genetics*, 31(12):683 – 695.
- 696 Kadakia, N. (2019). ORN-WL-gain-control. GitHub. <https://github.com/emonetlab/ORN-WL-gain-control.git>.
697 926cfa7.
- 698 Keene, A. C. and Waddell, S. (2007). *Drosophila* olfactory memory: single genes to complex neural circuits. *Nature
699 Reviews Neuroscience*, 8:341–354.
- 700 Krishnamurthy, K., Hermundstad, A. M., Mora, T., Walczak, A. M., and Balasubramanian, V. (2017). Disorder and
701 the neural representation of complex odors: smelling in the real world. *bioRxiv*, doi:10.1101/160382.
- 702 Larsson, M. C., Domingos, A. I., Jones, W. D., Chiappe, M., Amrein, H., and Vosshall, L. B. (2004). Or83b encodes a
703 broadly expressed odorant receptor essential for drosophila olfaction. *Neuron*, 43(5):703 – 714.
- 704 Litwin-Kumar, A., Harris, K. D., Axel, R., Sompolinsky, H., and Abbott, L. (2017). Optimal degrees of synaptic
705 connectivity. *Neuron*, 93(5):1153 – 1164.e7.
- 706 Malnic, B., Hirono, J., Sato, T., and Buck, L. B. (1999). Combinatorial receptor codes for odors. *Cell*, 96(5):713–723.

- 707 Martelli, C., Carlson, J. R., and Emonet, T. (2013). Intensity invariant dynamics and odor-specific latencies in
708 olfactory receptor neuron response. *Journal of Neuroscience*, 33(15):6285–6297.
- 709 Menuz, K., Larter, N. K., Park, J., and Carlson, J. R. (2014). An rna-seq screen of the drosophila antenna identifies
710 a transporter necessary for ammonia detection. *PLoS Genetics*, 10(11):e1004810.
- 711 Montague, S. A., Mathew, D., and Carlson, J. R. (2011). Similar odorants elicit different behavioral and physiological
712 responses, some supersustained. *Journal of Neuroscience*, 31(21):7891–7899.
- 713 Murlis, J. (1992). Odor plumes and how insects use them. *Annual Review of Entomology*, 37:505–532.
- 714 Nagel, K. and Wilson, R. (2011). Biophysical mechanisms underlying olfactory receptor neuron dynamics. *Nature
715 Neuroscience*, 14:208–216.
- 716 Nara, K., Saraiva, L. R., Ye, X., and Buck, L. B. (2011). A large-scale analysis of odor coding in the olfactory
717 epithelium. *Journal of Neuroscience*, 31(25):9179–9191.
- 718 Nemenman, I. (2012). Information theory and adaptation. In Wall, M. E., editor, *Quantitative Biology: From
719 Molecular to Cellular Systems*, chapter 4, pages 73–91. CRC Press, USA.
- 720 Olsen, S. R., Vikas, B., and Wilson, R. I. (2010). Divisive normalization in olfactory population codes. *Neuron*,
721 66:287–299.
- 722 Olsen, S. R. and Wilson, R. I. (2008). Lateral presynaptic inhibition mediates gain control in an olfactory circuit.
723 *Nature*, 452:952–960.
- 724 Papadopoulou, M., Cassenaer, S., Nowotny, T., and Laurent, G. (2011). Normalization for sparse encoding of
725 odors by a wide-field interneuron. *Science*, 332(6030):721–725.
- 726 Pehlevan, C., Genkin, A., and Chklovskii, D. B. (2017). A clustering neural network model of insect olfaction. In
727 *2017 51st Asilomar Conference on Signals, Systems, and Computers*, pages 593–600.
- 728 Raman, B., Joseph, J., Tang, J., and Stopfer, M. (2010). Temporally diverse firing patterns in olfactory receptor
729 neurons underlie spatiotemporal neural codes for odors. *Journal of Neuroscience*, 30(6):1994–2006.
- 730 Reddy, G., Zak, J., Vergassola, M., and Murthy, V. N. (2018). Antagonism in olfactory receptor neurons and its
731 implications for the perception of odor mixtures. *eLife*, 7:344958.
- 732 Renou, M., Party, V., Rouyar, A., and Anton, S. (2015). Olfactory signal coding in an odor background. *Biosystems*,
733 136:35 – 45.
- 734 Saha, D., Leong, K., Peterson, S., Siegel, G., and Raman, B. (2013). A spatiotemporal coding mechanism for
735 background-invariant odor recognition. *Nature Neuroscience*, 16:1830–1839.
- 736 Schulze, A., Gomez-Marin, A., Rajendran, V. G., Lott, G., Musy, M., Ahammad, P., Deogade, A., Sharpe, J., Riedl, J.,
737 Jarriault, D., Trautman, E. T., Werner, C., Venkadesan, M., Druckmann, S., Jayaraman, V., and Louis, M. (2015).
738 Dynamical feature extraction at the sensory periphery guides chemotaxis. *eLife*, 4:e06694.
- 739 Si, G., Kanwal, J., Hu, Y., Tabone, C. J., Baron, J., Berck, M. E., Vignoud, G., and Samuel, A. D. (2019). Invariances in
740 a combinatorial olfactory receptor code. *Neuron*, 101:950–962.
- 741 Stensmyr, M. C., Dweck, H. K., Farhan, A., Ibba, I., Strutz, A., Mukunda, L., Linz, J., Grabe, V., Steck, K., Lavista-
742 Llanos, S., Wicher, D., Sachse, S., Knaden, M., Becher, P. G., Seki, Y., and Hansson, B. S. (2012). A conserved
743 dedicated olfactory circuit for detecting harmful microbes in drosophila. *Cell*, 151(6):1345 – 1357.
- 744 Stevens, C. F. (2016). A statistical property of fly odor responses is conserved across odors. *Proceedings of the
745 National Academy of Sciences*, 113(24).
- 746 Su, C.-Y., Menuz, K., Reisert, J., and Carlson, J. (2012). Non-synaptic inhibition between grouped neurons in an
747 olfactory circuit. *Nature*, 492(7427):76–71.
- 748 van der Maaten, L. and Hinton, G. (2008). Visualizing high-dimensional data using t-sne. *Journal of Machine
749 Learning Research*, 9:2579–2605.
- 750 Vogt, R. G. and Riddiford, L. M. (1981). Pheromone binding and inactivation by moth antennae. *Nature*,
751 293(5828):161.

- 752 Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., and Axel, R. (1999). A spatial map of olfactory receptor
753 expression in the drosophila antenna. *Cell*, 96(5):725 – 736.
- 754 Vosshall, L. B., Wong, A. M., and Axel, R. (2000). An olfactory sensory map in the fly brain. *Cell*, 102(2):147 – 159.
- 755 Waite, A. J., Frankel, N. W., and Emonet, T. (2018). Behavioral variability and phenotypic diversity in bacterial
756 chemotaxis. *Annual Review of Biophysics*, 47(1):595–616. PMID: 29618219.
- 757 Wang, G., Carey, A. F., Carlson, J. R., and Zwiebel, L. J. (2010). Molecular basis of odor coding in the malaria vector
758 mosquito anopheles gambiae. *Proceedings of the National Academy of Sciences*, 107(9):4418–4423.
- 759 Wark, B., Lundstrom, B. N., and Fairhall, A. (2007). Sensory adaptation. *Current Opinion in Neurobiology*,
760 17(4):423–429.
- 761 Weber, E. H. (1996). *EH Weber on the tactile senses*. Psychology Press.
- 762 Weissburg, M. (2000). The fluid dynamical context of chemosensory behavior. *The Biological Bulletin*, 198(2):188–
763 202. PMID: 10786940.
- 764 Wilson, C. D., Serrano, G. O., Koulakov, A. A., and Rinberg, D. (2017). A primacy code for odor identity. *Nature
765 Communications*, 8(1):1477.
- 766 Wilson, R. I. (2013). Early olfactory processing in *Drosophila*: mechanisms and principles. *Annual Review of
767 Neuroscience*, 36(1):217–241.
- 768 Zhou, Y. and Sharpee, T. O. (2018). Using global t-SNE to preserve inter-cluster data structure. *bioRxiv*,
769 doi:10.1101/331611.

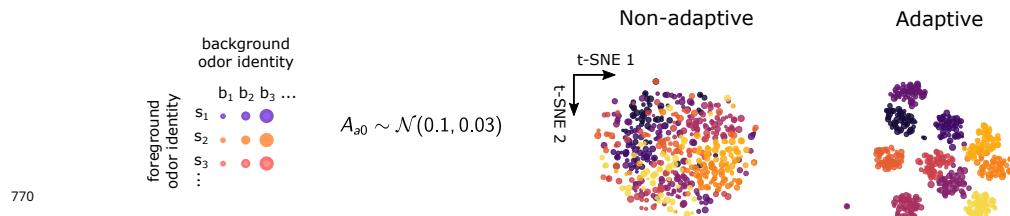


Figure 2-Figure supplement 1. t-SNE projections for non-adaptive ($\beta = 1$) and adaptive ($\beta = 0$) systems, when background firing rates depend on ORN identity. Background active fractions A_{0a} are chosen normally with mean 0.1 and deviation 0.03, corresponding to background firing rates around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

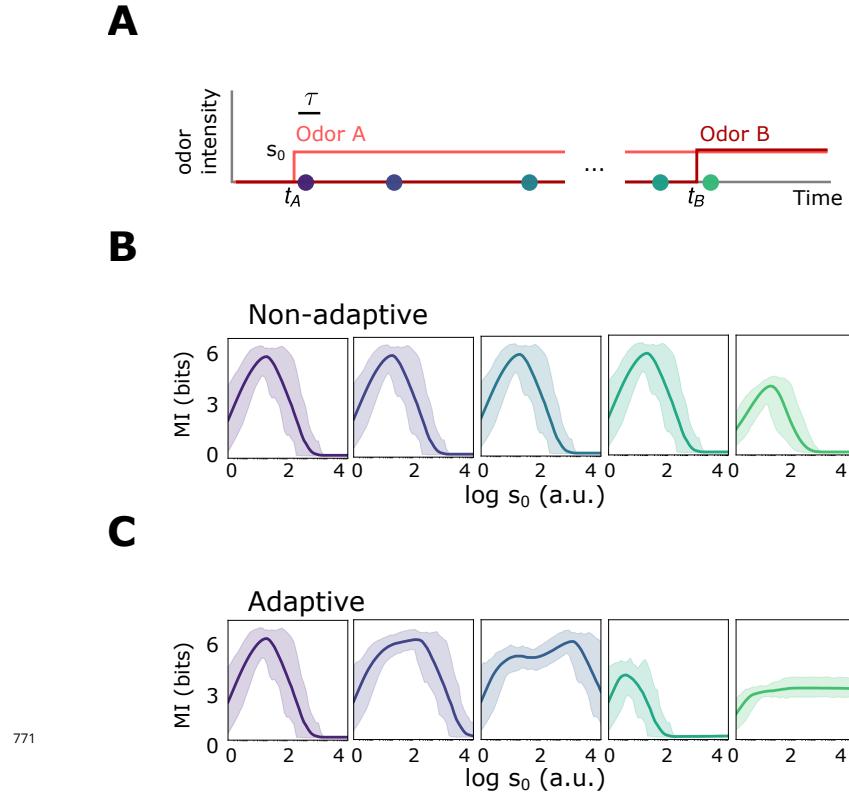


Figure 2–Figure supplement 2. Front-end adaptive feedback preserves information capacity of the ORN sensing repertoire. **A** Mutual information between signal $s(t) = s_A(t) + s_B(t)$ and response $r(t)$ is calculated at various points in time t for an odor environment consisting of two step odors, A and B. Odor A, with concentration $s_A(t)$, turns on at time t_A and a odor B, with concentration $s_B(t)$, 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 ~ 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).

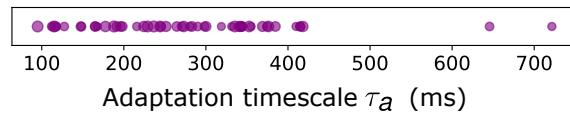
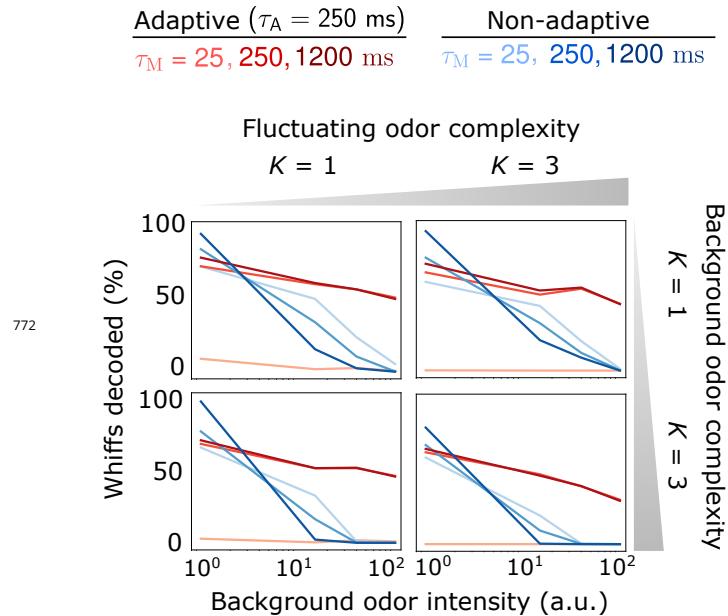
A**B**

Figure 3-Figure supplement 1. Decoding accuracy for system with ORN-dependent adaptive timescales τ_a . **A** Distribution of timescales for all ORNs a (purple dots). Here, $\tau_a \sim \tau 10^X$ 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 τ_M (line shades). Plots are arrayed by the complexity of the naturalistic signal (column-wise) and the complexity of the background odor (row-wise).

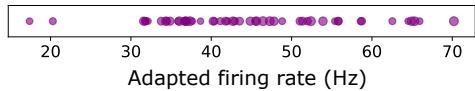
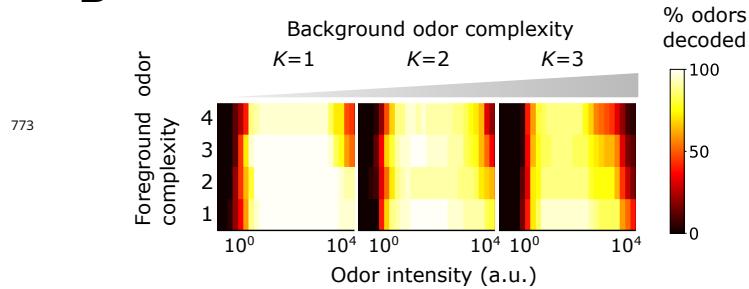
A**B**

Figure 3-Figure supplement 2. Benefits conferred by Weber-Fechner adaptation remain for a broader distribution of baseline adapted activity levels 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.

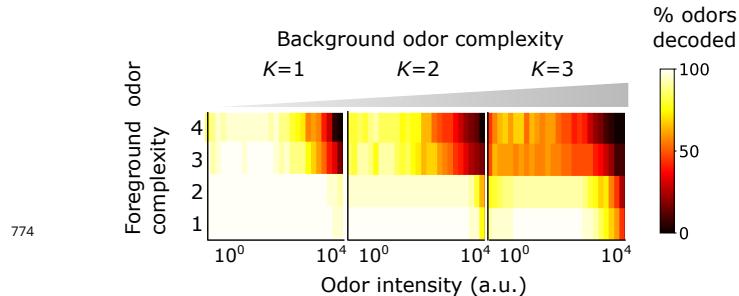


Figure 3-Figure supplement 3. Benefits conferred by Weber-Fechner adaptation remain for 2 binding sites per receptor. This might conceivably occur in insect olfactory receptors, heterotramers consisting of 4 Orco/Or subunits that gate a central ion channel pathway (*Butterwick et al., 2018*). Plotted is the decoding accuracy of foreground odors in the presence of background odors.

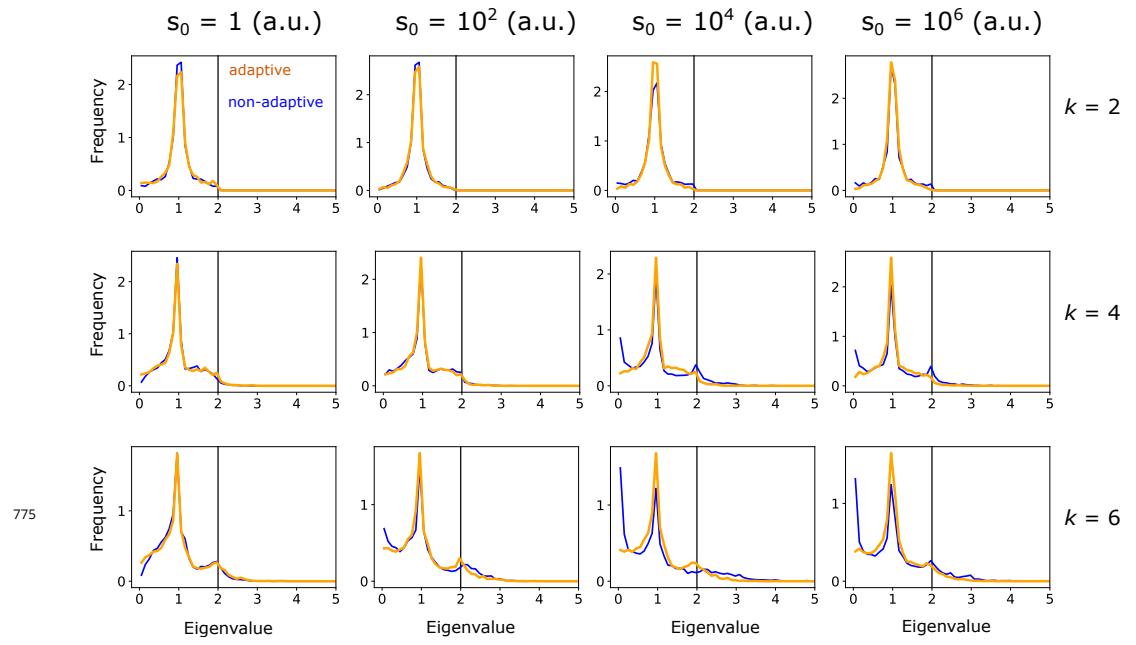


Figure 3–Figure supplement 4. Eigenvalue distribution of $\mathbf{D}_s^T \mathbf{D}_s$, where \mathbf{D}_s is a $m \times k$ submatrix of the column-normalized linearized ORN response matrix \mathbf{D} , evaluated at the linearization point s . Note that s is k -sparse, but its components do not necessarily 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 s 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.

775

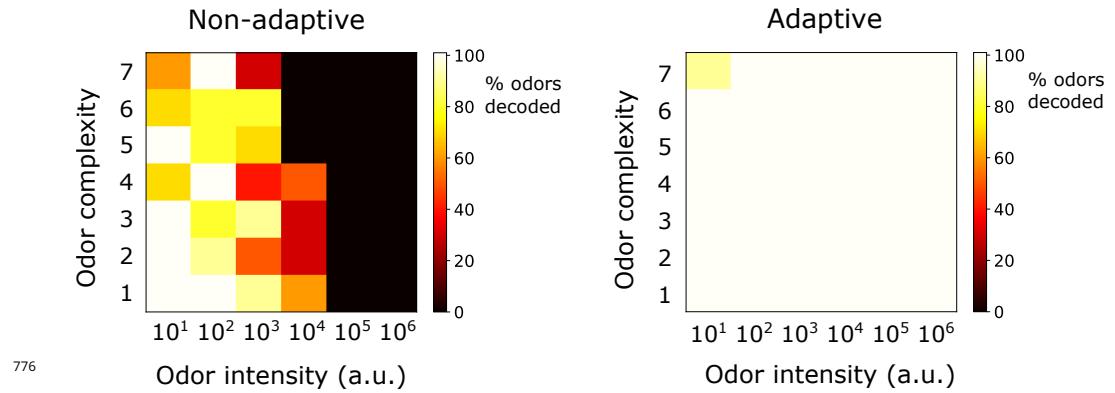


Figure 3-Figure supplement 5. Decoding of odor signals (no background odors) using the IHT algorithm (*Blumensath and Davies, 2009b; Blumensath, 2013*) qualitatively reproduces the results from the main text, which used traditional CS with background linearization. In the adaptive case, IHT actually 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. The iterative algorithm was initialized at $\hat{s} = \mathbf{0}$ and run forward until \hat{s} was stationary, or 10000 iterations were reached. Step size μ in Eq. 22 was set to $s_0/20$. At each step, the linearized response used in determining \mathbf{x}_{i+1} (\mathbf{D}_{s_i} in Eq. 22) was evaluated at the result of i th iteration, \mathbf{x}_i . IHT also requires an assumption on the number of components in the mixture (which defines $H_K(\cdot)$ in Eq. 22); here, that was set to twice the actual sparsity of true signal.

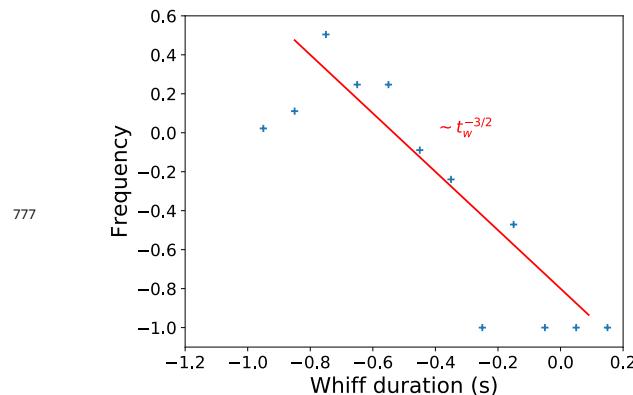


Figure 3-Figure supplement 6. Distribution of whiff durations in naturalistic stimulus, compared to the theoretical prediction (*Celani et al., 2014*)

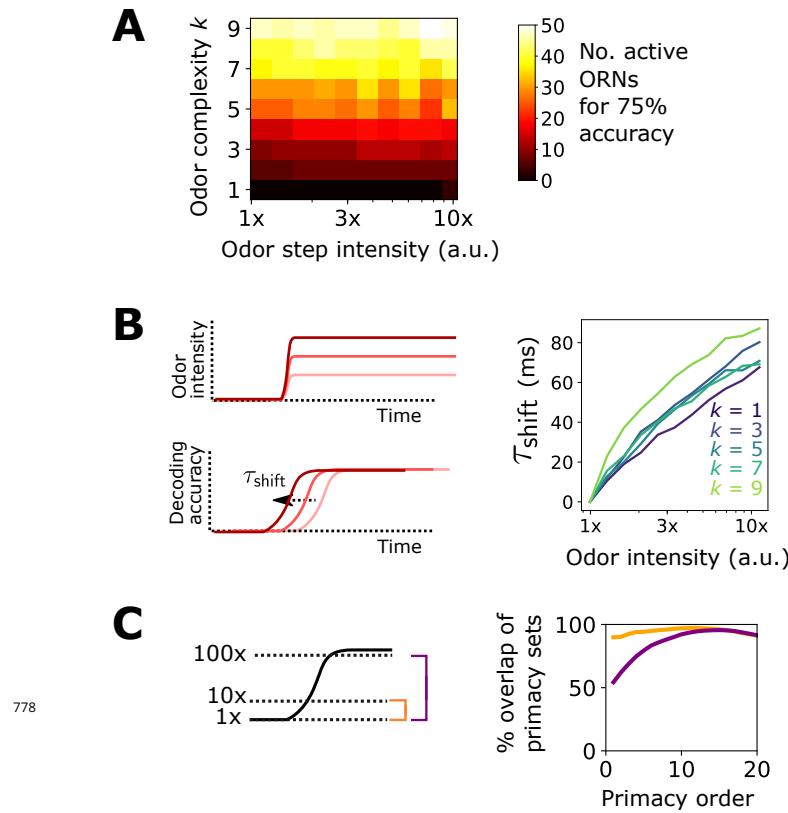


Figure 4-Figure supplement 1. Additional results pertaining to the primacy coding hypothesis. **A** Percent of active ORNs required for 75% accuracy of a steep sigmoidal odor step, as a function of odor step intensity and odor complexity. For low complexities, a primacy set of fewer ORNs may be sufficient 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 (Wilson et al., 2017). 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.

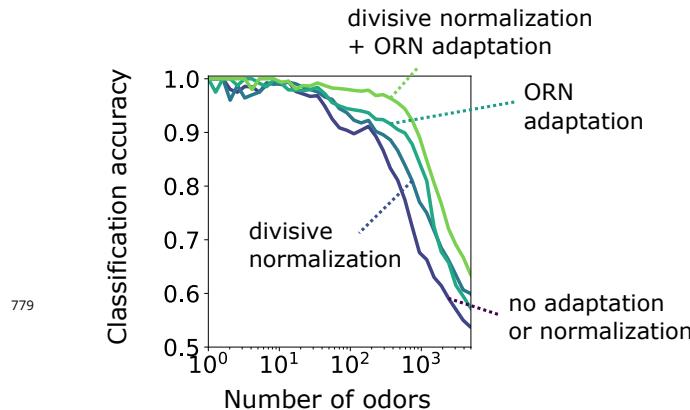


Figure 5–Figure supplement 1. Accuracy of binary classification by odor valence, for odors whose concentrations span a narrow range of concentrations (1 order of magnitude). Accuracy is plotted as a function of the number of distinct odor identities classified by the trained network, in systems with 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.