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

Nirag Kadakia

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

Thierry Emonet

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

Odor identity is encoded by spatiotemporal patterns of activity in olfactory receptor neurons (ORNs). In natural environments, the intensity and timescales of odor signals can span several orders of magnitude, and odors can mix with one another, potentially scrambling the combinatorial code mapping neural activity to odor identity. Recent studies have shown that in *Drosophila melanogaster* the ORNs that express the olfactory co-receptor Orco scale their gain inversely with mean odor concentration according to the Weber-Fechner Law of psychophysics. Here we use a minimal biophysical model of signal transduction, ORN firing, and signal decoding to investigate the implications of this front-end scaling law for the neural representations of odor identity. We find that Weber-Fechner scaling enhances coding capacity and promotes the reconstruction of odor identity from dynamic odor signals, even in the presence of confounding background odors and rapid intensity fluctuations. We show that these enhancements are further aided by downstream transformations in the antennal lobe and mushroom body. Thus, despite the broad overlap between individual ORN tuning curves, a mechanism of front-end adaptation, when endowed with Weber-Fechner scaling, may play a vital role in preserving representations of odor identity in naturalistic odor landscapes.

Animals identify and discriminate odors using olfactory receptors (Ors) expressed in olfactory receptor neurons (ORNs) [1–4]. Individual ORNs, which typically express a single Or, respond to many odorants, while individual odorants activate many distinct ORNs [5–8]. Odors are thus encoded by the combinatorial patterns of activity they elicit in the sensing periphery [5–7, 9–11], patterns decoded downstream into behavioral response [12]. Still, ethologically-relevant odors are often mixed with background ones [13] and intensity can vary widely and rapidly as odors are carried by the wind [14–17]. How are odors recognized reliably despite these confounds? In *Drosophila melanogaster*, ORN dose response curves exhibit similar Hill coefficients but distinct power-law distributed activation thresholds [6, 18], which together with inhibitory odorants enhance coding capacity [6, 18–20]. In antennal lobe (AL) glomeruli, mutual lateral inhibition normalizes population response, reducing the dependency of activity patterns on odor concentration [21, 22]. Further downstream, sparse connectivity to the mushroom body (MB) helps maintain neural representations of odors, and facilitates compressed sensing and associative learning schemes [23–26]. Finally, temporal features of neural responses contribute to concentration-invariant representations of odor identity [27–30].

Here we examine how short-time ORN adaptation at the very front-end of the insect olfactory circuit contributes to the fidelity of odor encoding. Our theoretical study is motivated by the recent discovery of invariances in the response dynamics of ORNs expressing the co-receptor Orco. The responses of ORNs to diverse odorants of the same concentration differ widely, due to differences in odor-receptor affinities [6, 31, 32] and stimulus dynamics [33]. However, downstream from this input nonlinearity, signal transduction and adaptation dynamics exhibit a surprising degree of invariance with respect to odor-receptor identity: reverse-correlation analysis of ORN response to fluctuating stimuli produces highly stereotyped, concentration-invariant response filters [18, 33, 34].

These properties stem in part from an apparently invariant adaptive scaling law in ORNs: gain varies inversely with mean odor concentration according to the Weber-Fechner Law of psychophysics [35, 36], irrespective of the odor-receptor combination [34, 37, 38]. The invariance of this relatively fast adaptation ($\sim$250 ms) can be traced back to feedback mechanisms in odor transduction, upstream of ORN firing [34, 37–39], which depend on the activity of the signaling pathway rather than on the identity of its receptor [39]. The generality of the adaptive scaling suggests it could be mediated by the highly conserved Orco co-receptor [40–43]. Indeed, phosphorylation sites have been recently identified on Orco, some being implicated in odor desensitization, albeit over much longer timescales [43, 44].

While in a simple system such as *E. coli* chemotaxis [45], adaptive feedback via the Weber-Fechner Law robustly maintains sensitivity over concentration changes, the implication for a multiple-channel system – which combines information from hundreds of cells with overlapping receptive fields – is less straightforward. Here
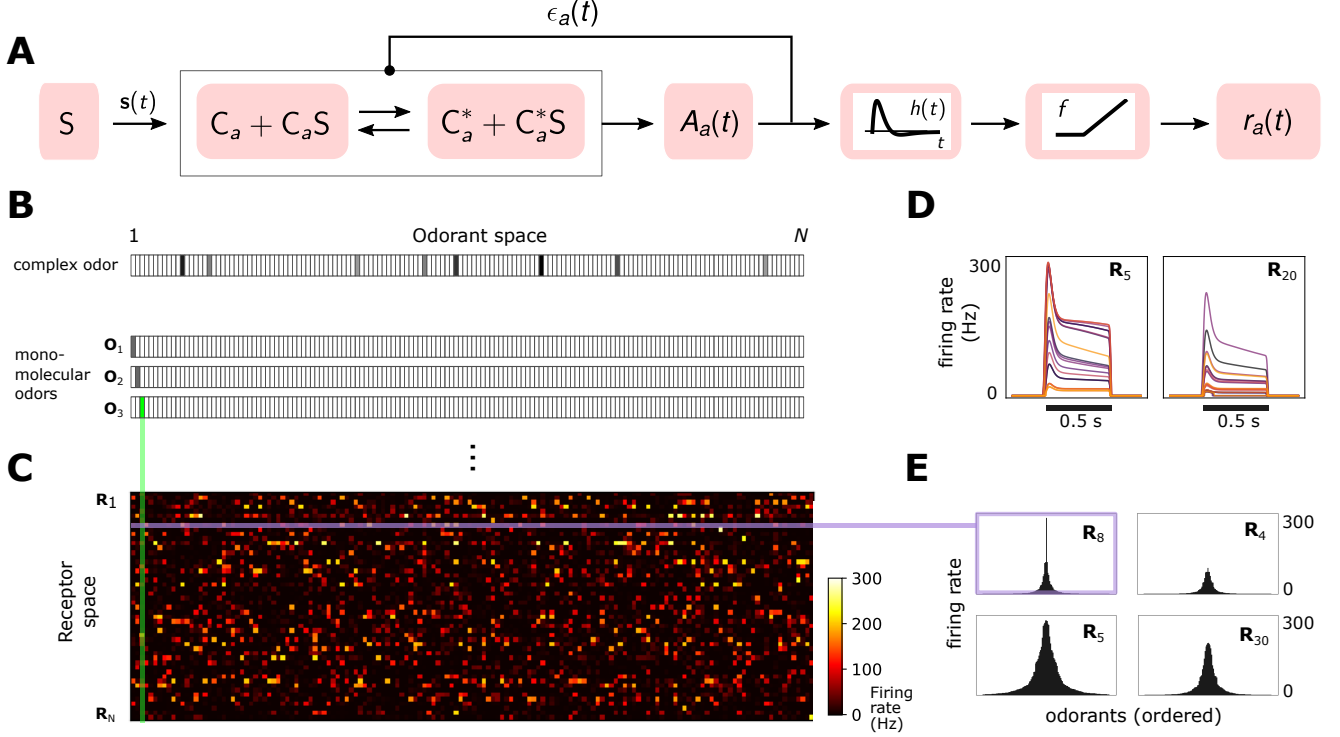
FIG. 1: **A** Simple ORN model. Or/Orco complexes C_a respond to odor concentration $\mathbf{s}(t) = (s_1(t), ..., s_i(t), ..., s_N(t))$ by binding odorant molecules (S in the diagram) of type i and concentration $s_i(t)$, where t is time. Or/Orco complexes stochastically switch between active and inactive states, where the steady-state active fraction is determined by the free energy difference (in units of $k_B T$) between active and inactive conformations in the unbound state, $\epsilon_a(t)$, and by odorant binding with dissociation constants K_{ai}^* and K_{ai} (SI text). Adaptation is mediated by a negative feedback [39] 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** Odor mixtures are represented by N-dimensional vectors $\mathbf{s}$, whose components s_i are the concentrations of the individual molecular constituents of $\mathbf{s}$. **C** Step-stimulus firing rate of 50 ORNs to the 150-possible monomolecular odors $\mathbf{s} = s_i$, given power-law distributed K_{ai}^* [18]. **D** Temporal responses of two representative ORNs to a pulse stimulus, for several monomolecular odorants (colors). **E** Representative ORN tuning curves (a single row of the response matrix in C, ordered by magnitude). Tuning curves are diverse, mimicking measured responses [6].

we combine a biophysical model of ORN adaptive response and neural firing with various sparse signal decoding frameworks to explore how ORN adaptation with Weber-Fechner scaling affects combinatorial coding and decoding of odor signals spanning varying degrees of intensity, molecular complexity, and temporal structure. We find that this front-end adaptive mechanism promotes the accurate discrimination of odor signals from backgrounds of varying molecular complexity, and aids other known mechanisms of neural processing in the olfactory circuit to maintain representations of odor identity across environmental changes.

RESULTS

Model of ORN sensing repertoire

We consider a repertoire of $M = 50$ Orco-expressing ORN types modeled using a simple extension of a mini-

mal model of odor-to-ORN firing [34] that reproduces the type of Weber-Fechner adaptation and firing rate dynamics measured in individual *Drosophila* ORNs in response to Gaussian and naturalistic signals. Within ORNs of type $a = 1, ..., M$, we model Or-Orco complexes as nonselective cation channels [40] (Fig. 1a) that stochastically switch between active and inactive states, while simultaneously binding to odorants i with dissociation constants, K_{ai}^* and K_{ai}, respectively [34, 39]. Assuming these processes are faster than other reactions in the signaling pathway, the quasi-steady state active fraction A_a of channels in ORNs of type a reads:

$$A_a(t) = \left(1 + e^{\epsilon_a(t)} \frac{1 + \sum_i^N s_i(t)/K_{ai}}{1 + \sum_i^N s_i(t)/K_{ai}^*}\right)^{-1}. \quad (1)$$

$s_i(t)$ are the time-dependent concentrations of the individual monomolecular components of the odor signal $\mathbf{s}(\mathbf{t})$ at time t, and $N = 150$ is the size of the molecular odorant space (Fig. 1b). Inward currents elicited by

activating Or-Orco channels [40] eventually result in a negative feedback onto $A_a(t)$ [34, 38, 39]:

$$\tau \frac{d\epsilon_a(t)}{dt} = A_a(t) - A_0. \qquad (2)$$

Here, τ is the adaptation time and $\epsilon_a(t)$ is the free energy difference in units of $k_B T$ between active and inactive conformations of the unbound Or-Orco channel. We assume changes in free energy are limited to the finite range $\epsilon_{L,a} < \epsilon_a(t) < \epsilon_{H,a}$ [34]. Firing rate is minimally modeled by filtering the activity $A_a(t)$ with the bi-lobed filter $h(t)$ and rectifying nonlinearity f [34]:

$$r_a(t) = f\left(h \otimes A_a(t)\right), \qquad (3)$$

where $\otimes$ is convolution. When deconvolved from stimulus dynamics, the shapes of the temporal kernels of *Drosophila* ORNs that express Orco are largely receptor- and odor-independent [18, 33, 34]. Moreover, adaptation is not intrinsic to the receptor [39]. Accordingly, for simplicity τ, A_0, $h(t)$, and f are assumed independent of receptor and odorant identities.

We assume that the lower cutoffs $\epsilon_{L,a}$ are receptor-dependent and choose them from a normal distribution. This variability ensures that ORNs are activated above quiescence (set at 5 Hz) at distinct stimulus levels [33, 34]. Diversity among odor-ORN responses arises mainly from the distribution of chemical dissociation constants (Fig. 1c). For simplicity we only consider agonists, i.e. $K_{ai}^* > K_{ai}$, and assume receptors can only bind one odorant at a time. The analysis can easily be extended to include inhibitory odorants, which increases coding capacity [19]. We choose the dissociation constants from a power law distribution ($\alpha = 0.35$) recently found across ORN-odor pairs in *Drosophila* larvae [18]. For a handful of ORNs we choose a very small value for one of the K_{ai}^* to mimic high responders to private odorants relevant to innate responses [32]. These private odors do not affect the general findings.

While this phenomenological model could be extended to include further details – e.g. we could relax the quasi-steady-state assumption in Eq. 1 and use a more complex model for neural firing [34] – this minimally-parameterized form captures the key dynamical properties of Orco-expressing ORNs relevant to our study: receptor-independent adaptation [39] with Weber-Fechner scaling [34, 37, 38] that maintains response time independent of mean stimulus intensity [33, 34], along with a diversity of both steady state and temporal firing patterns in response to a panel of monomolecular odorants [6, 27–29, 31] (Fig. 1d-1e).
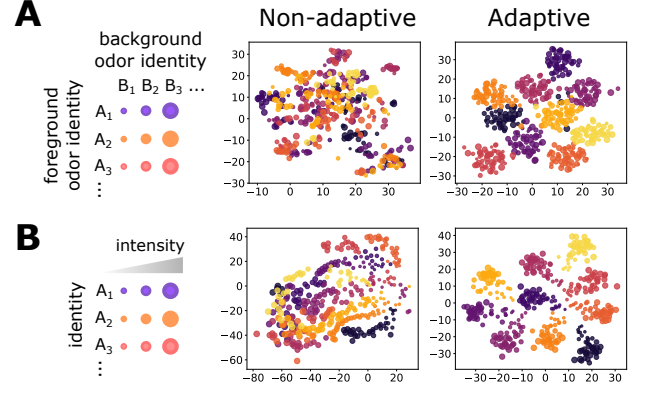

FIG. 2: Front-end adaptation maintains representations of odor identity across background and intensity confounds. **A** Abstract representation of ORN responses in a low-dimensional embedding. Each point represents the repertoire of ORN firing rates, in response to an odor environment with both a foreground (point color) and background (point size) odor. In the adaptive system, ϵ_a are set to their steady state values given odor B alone. **B** Similar to A, but now for odors whose concentrations span 4 decades (represented by point size). Here, the background odor identity is the same for all concentrations.

Concentration-invariant preservation of coding capacity and abstract representations of odor identity

To investigate how front-end Weber-Law adaptation might preserve representations of odor identity within the repertoire of ORN response, we project the 50-dimensional firing rates $\mathbf{r}$ down to a 2-dimensional space using t-distributed stochastic neighbor embedding (t-SNE) [46]. We first perform this embedding for an adaptive or non-adaptive system interacting with an odor environment containing a foreground odor A atop a background odor B (Fig. 2a). Both odors are sparse mixtures, with $K < N$ odorants of similar concentrations, odor "identity" being the particular set of odorants in the mixture. Adaptation to the background is enacted by setting ϵ_a to their steady state values (via Eq. 2) in response to odor B alone. With adaptation in place, responses cluster by the identity of odor A, suggesting that ORN responses appropriately encode the identity of novel odors irrespective of background signals – once these backgrounds have been "adapted away" (Fig. 2a). This is notable, since, due to the combinatorics of ORN response (Fig. 1c), the adapted ϵ_a distributions are themselves heavily dependent on background identity. Responses in the non-adaptive system, meanwhile, exhibit no such clustering (Fig. 2a). A similar separation by odor identity is preserved in the adaptive system if we consider responses across different intensities (Fig. 2b).

Preservation of these representations is enabled by the preservation of coding capacity. Accordingly, we calcu-

lated the mutual information between odor and ORN response in time, verifying that the the adaptive system retains coding capacity as it confronts novel odors (Fig. S1). The non-adaptive system maintains coding capacity, though in a far more limited range of odor concentration.

Front-end adaptation enhances odor discrimination in complex environments

How well does the preservation of coding capacity translate to better signal reconstruction? One potentially complicating factor is the disparity between sensor dimension and stimulus dimension: while *Drosophila* only express ~ 60 Or genes [47], the space of odorants is far greater [25]. However, many naturally-occurring odors are comprised of a small subset of odorants, which is suggestive as the theory of compressed sensing (CS) guarantees their reconstruction [48, 49]. It is unknown whether CS is implemented in the *Drosophila* olfactory circuit [50], and we use it mainly as a tool to quantify how front-end adaptation potentially affects odor decoding, later verifying our conclusions with other classification techniques that incorporate the known architecture of the olfactory system.

To incorporate the linear framework of CS, we treat the nonlinear odor encoding exactly but approximate the decoding to first order (this linearization simplifies the computation, but is not critical for our general results; see SI text and Fig. S7-S8). Odors **s** are assumed sparse, with $K \ll N$ nonzero components s_i with mean concentration s_0. We first examine how foreground odors are recognized when mixed with background odors of a distinct identity but similar intensities, quantifying decoding accuracy as the percentage of odors correctly decoded within some tolerance (Fig. 3a). Without adaptation, accuracy is maintained within the range of receptor sensitivity for monomolecular backgrounds, but is virtually eliminated as background complexity rises (Fig. 3b). The range of sensitivity is broader in the adaptive system, and is substantially more robust across odor concentration and complexity.

In realistic odor environments, the concentration and duration of individual odor whiffs vary widely [16]. 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 using a photo-ionization detector [34] (Fig. 3c) to which we randomly assign sparse identities from the N-dimensional odorant space. To mimic background confounds, we combine these signals with static odor backgrounds, and then calculate the percentage of decoded whiffs. We assume the decoder has short-term memory: detected odor signals are only retained for τ_{M}

seconds in the immediate past. 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 (Fig. 3d, blue lines). In the adaptive system, this is substantially mitigated (red lines in Fig. 3d), provided the memory duration τ_{M} is at least as long as the adaptation timescale τ (darker red lines). Because this short-term adaptation depends on the activity of the Or-Orco channel rather than on the identity of the receptor [33, 34, 39], the values of τ and A_0 were assumed the same for all ORNs; still, our results hold if these invariances are relaxed (SI, Fig. S4-S5).

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 [30]. 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 (Fig. 4a), in agreement with primacy coding. As expected for more complex odor mixtures, the full repertoire is required for accurate decoding. Primacy coding also predicts that for stronger stimuli, responses occur earlier, since the primacy set is realized quicker, which our framework replicates (Fig. S2).

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 primacy sets are highly consistent across background confounds for all but the smallest primacy orders (Fig. 4b-4c). This also holds true for backgrounds of different concentrations (Fig. S2), suggesting a central role for front-end adaptation in reinforcing primacy codes across differing environmental conditions.
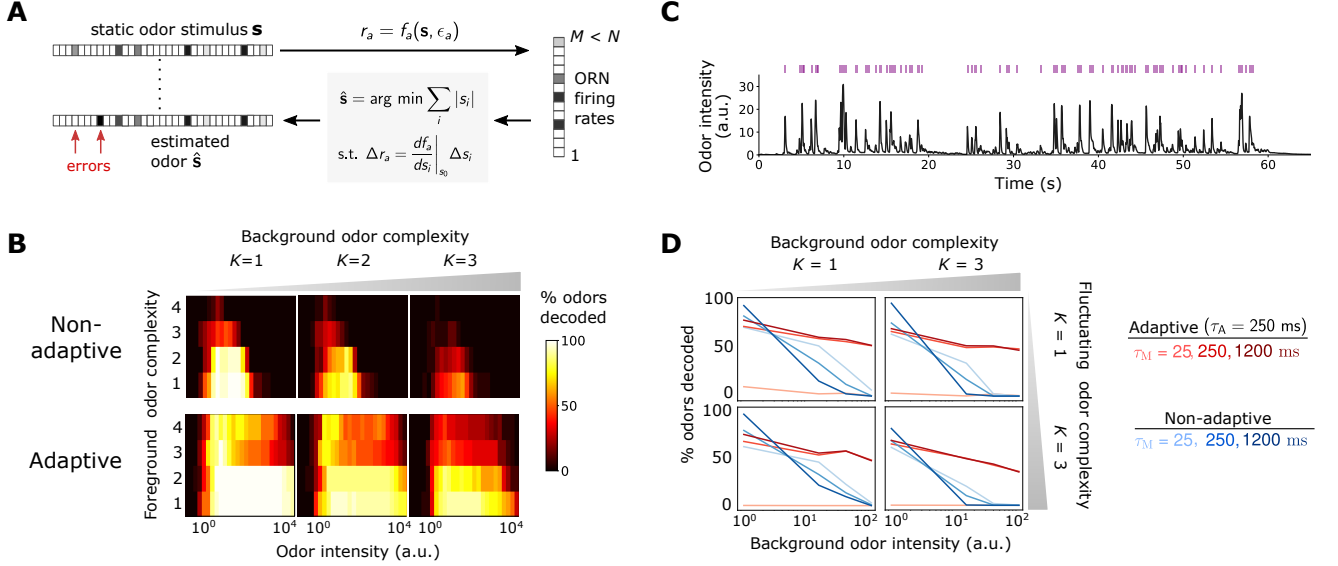
FIG. 3: Front-end adaptation promotes accurate odor decoding in static and naturalistic odor environments. **A** Odor stimuli produce ORN responses via odor-binding and activation and firing machinery, as described by Eqs. 1-3. Odors are then decoded using compressed sensing by linearizing around a background s_0 and minimizing the constrained L_1 norm of the odor signal. Odors are assumed sparse, with K nonzero components, $K \ll N$. Odors are considered accurately decoded if the K sparse components are estimated within 25% and the N-K components not in the mixture are estimated below 10% of s_0. **B** Decoding accuracy of foreground odors in the presence of background odors. **C** 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. **D** 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 t_M (line shades).

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

Signal transformations in the sensing periphery are propagated through the remainder of the olfactory circuit. How does front-end adaptation interact with these subsequent neural transformations? ORNs expressing the same OR converge to a unique AL glomerulus, where they receive lateral inhibition from other glomeruli [21, 51]. This inhibition implements a type of divisive gain control [22], normalizing the activity of output projections neurons, which then synapse onto a large number of Kenyon cells (KCs) in the mushroom body. To investigate how odor representations are affected by interactions between front-end ORN adaptation and this lateral inhibition and synaptic divergence, we extended our ORN encoding model by adding uniglomerular connections from ORNs to the antennal lobe, followed by sparse, divergent connections to 2500 KCs [23, 24, 52]. Inhibition was modeled via divisive normalization, with parameters chosen according to experiment [22]. We quantified decoding accuracy by training and testing a binary classifier on the KC activity output of sparse odors of distinct intensity and identity, randomly categorized as appetitive or aversive. For simplicity, odor signals of the same identity but differing intensity were assigned the same valence. We trained the classifier on N_{ID} sparse odor identities at intensities chosen randomly over 4 orders of magnitude, then tested the classifier accuracy on the same odor identities but of differing concentrations.

Classification accuracy degrades to chance level as N_{ID} becomes very large (Fig. 5a). When acting alone, either divisive normalization or ORN adaptation can help, although the effect of ORN adaptation is stronger. When both are active, accuracy improves further, suggesting that these distinct adaptive transformations may act jointly at different stages of neural processing in preserving representations of odor identity. As expected, these gains mostly vanish for the same odors chosen from a narrower range of concentrations (Fig. S3).

If we train the classifier to distinguish odors by identity rather than valence, the benefits conferred by divisive normalization do not appear until N_{ID} is substantial, with accuracy below 65% for $N_{\text{ID}} > 50$ (Fig. 5b). On the other hand, with ORN adaptation accuracy remains above 85% for more than 1000 odor identities, strongly implicating front-end adaptation as a key player in maintaining odor identity representations, before signals are further processed downstream.
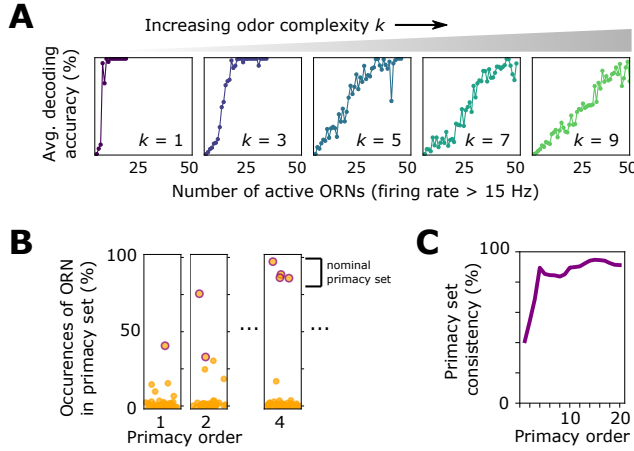
FIG. 4: **A** Decoding accuracy as a function of the number of active ORNs, for different odor complexities. The primacy set consists of those ORNs required to be active for accurate decoding. **B** Frequency of particular ORNs in primacy sets of an odor placed atop different backgrounds. Individual plots show, for given primacy order p, the percentage of backgrounds for which the primacy set of odor A contains a given ORN (dots). Those with purple borders are the p most highly occurring – i.e. a nominal background-invariant primacy set for odor A. Points are jittered horizontally for visualization. **C** Consistency of primacy sets across backgrounds, as a function of p. Consistency is defined as the likelihood that an ORN in the nominal primacy set appears in any of the individual background-dependent primacy sets, averaged over the nominal set (average of the y-values of the purple dots in B). 100% consistency means that for all backgrounds, the primacy set of odor A is always the same p ORNs.

DISCUSSION

We have found that Weber-Law adaptation at the very front-end of the insect olfactory circuit [34, 37, 38] may contribute significantly to the preservation of neural representations of odor identity amid confounding odors and intensity fluctuations. Drawing on experimental evidence for a number of ORN-invariant response features [18, 20, 33, 34, 39], we have found that this mechanism of dynamic adaptation confers significant benefits in coding fidelity, without the need for ORN-specific parameterizations. Still, our results hold when these invariances such as adaptation timescale or baseline activity are relaxed (SI Figs. S4-S5). In the olfactory periphery, front-end Weber Law adaptation therefore appears fairly robust, a consequence of controlling gain via feedback from channel activity [34, 39, 45], rather than through intrinsic, receptor-dependent mechanisms. While our framework incorporates many observed features of the *Drosphila* olfactory system – Weber-Law adaptation, power-law distributed receptor affinities, temporal filter invariance, connectivity topologies – it is minimal. We considered only one of the chemoreceptor families expressed in the fly antenna [1] and ignored possible contri-
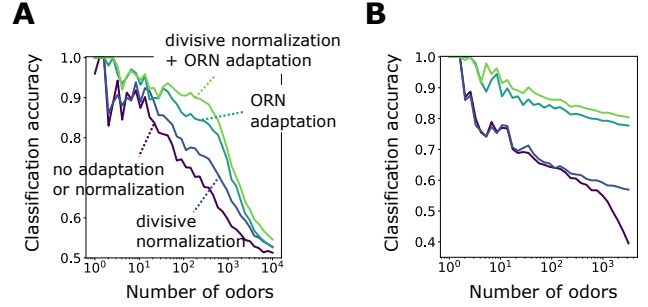

FIG. 5: **A** Accuracy of binary classification by odor valence, as a function of the number of distinct odor identities classified by the trained network (concentrations span 4 orders of magnitude), in systems with only ORN adaptation, only divisive normalization, both or neither. **B** Same as (A)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

butions of odor binding proteins [53, 54], inhibitory odorants [19], and odorant-odorant antagonism [55], which could further boost coding capacity and preserve representation sparsity. Useful extensions to our nonlinear-linear-nonlinear model might incorporate ephaptic coupling between ORNs housed in the same sensillum [56], global inhibition in the mushroom body [57], and the effects of long-term adaptation [43].

Previous studies have characterized various neural mechanisms that help preserve combinatorial codes. Lateral inhibition between glomeruli helps tame saturation and boost weak signals [22]. The sparse degree of connectivity to either the olfactory bulb (vertebrates) or mushroom body (insects) may also be precisely tuned to optimize the capacity to learn associations [24]. In this work, we find that some of these downstream features act in concert with front-end dynamic adaptation in maintaining representations of odor identity.

Other studies have implicated the unique temporal patterns of neural response as signatures of odor identity [27–29, 58]. ORN and projection neuron time traces form distinct trajectories in low-dimensional projections, and cluster by odor identity, much as we have found here (Fig. 2). In our framework, temporal coding is implicit: because the input nonlinearity depends on the diversity of binding affinities, odor signals are naturally formatted into temporal patterns that are both odor- and ORN-specific (Figs. 1d-1e). Further, the short required memory timescales ($\tau_M \sim \tau \sim 250$ ms) suggest that only brief time windows are needed for accurate odor identification, consistent with previous findings [27]. Moreover, we find that front-end adaptation enhances the robustness of other combinatorial coding schemes, such as primacy coding [30], which relies on the temporal order of ORN activation but not absolute firing rate (Fig. 4).

In the well-characterized chemosensory system of bacterial chemotaxis, Weber Law adaptation is enacted

through a feedback loop from the output activity of the receptor-kinase complexes onto the enzymes modifying receptor sensitivity [45]. It is interesting that some aspects of this logic are also present in ORNs: although the molecular players are different (and still largely unknown), it has been shown that transduction activity feeds back onto Or-Orco cation channel opening, enabling the Weber-Fechner relation [34, 38, 39]. That this adaptation mechanism appears to act similarly across ORNs [33, 34, 38] suggests the possible involvement of the universal co-receptor Orco, whose role in long-term adaptation has recently been reported [41–43]. Further, the identification of 4 subunits comprising the Orco-Or ion channel suggest that generic Or/Orco complexes may contain multiple odorant binding sites, which when included in our model supports our general findings (Fig. S6).

Weber Law ensures that sensory systems remain in the regime of maximum sensitivity, broadening dynamic range and maintaining information capacity [59]. For a single-channel system, this requires matching the midpoint of the dose-response curve to the mean ligand concentration [60], a strategy which may fail in multi-channel systems with overlapping tuning curves: adaptation to one signal could inhibit identification of others, if the signals excite some but not all of the same sensors. Our results show that this strategy is still largely functional. This can be traced to the observation that in CS, accuracy is guaranteed when sufficiently distinct odor identities produce sufficiently distinct ORN responses, a condition known as the restricted isometry property [49]. Indeed, the Weber-Fechner scaling increases the likelihood that this property is satisfied, beyond that in the non-adaptive system (SI text and Figs. S7-S8). Still, restricted isometry does not require that response repertoires are *invariant* to environmental changes. That is, even if the subset of active ORNs were concentration-dependent, odors could still in principle be fully reconstructible by CS. Nonetheless, our results in t-SNE clustering (Fig. 2), primacy coding (Fig. 4b-4c), and odor classification (Fig. 5) suggest that such response invariance is a natural byproduct of front-end adaptation. Together, this implies that Weber Law adaptation, whether required by the olfactory circuit for precise signal reconstruction (as in CS) or for developing odor associations (as in classification), can play an integral part in maintaining combinatorial codes amid changing environmental conditions.

METHODS

Equations 1-2 are integrated numerically using the Euler method with a 2 ms time step. For ORN firing (Eq. 3), $h(t)$ is bi-lobed [33]: $h(t) = Ap_{\mathrm{Gam}}(t; \alpha_1, \tau_1) - Bp_{\mathrm{Gam}}(t; \alpha_2, \tau_2)$, $A = 190$, $B = 1.33$, $\alpha_1 = 2$, $\alpha_2 = 3$, $\tau_1 = 0.012$, and $\tau_2 = 0.016$, where p_{Gam} is the pdf of $\mathrm{Gamma}(\alpha, 1/\tau)$. Nonlinearity f is modeled as a linear rectifier with 5 Hz threshold. For t-SNE dimensionality reduction, ORN responses were generated for odor signal combinations consisting of 1 (among 10) distinct foreground odors A atop 1 (among 50) distinct background odors B, each of complexity $K = 5$, for Fig. 2a. Fig. 2b plots responses for 10 odors ($K = 5$) at 40 concentrations spanning 4 decades, atop a random sparse background odor ($K = 5$) of similar magnitude.

For compressed sensing decoding, sparse components s_i are chosen as $s_i = s_0 + \Delta s_i$ where s_0 is set as the center of linearization and $\Delta s_i \sim \mathcal{N}(s_0/3, s_0/9)$. Reconstructed signal components $\hat{s}_i = s_0 + \Delta s_i$ are computed by minimizing $\sum_i |\Delta s_i|$ subject to $\Delta r_a = \sum_i dr_a/ds_i\big|_{s_0} \Delta s_i$ where $\Delta r_a = r_a(\mathbf{s}) - r_a(s_0)$ are the "excess ORN firing rates about the linearization point. For static stimuli, ϵ_a equals the fixed point of Eq. 2 in response to the background stimulus. For fluctuating stimuli, ϵ_a is updated in time by continuously integrating $r_a(t)$, via Eqs. 2 and 3; thus, only knowledge of $r_a(t)$ is needed by the decoder. The naturalistic odor signal (Fig. 3d) was generated by randomly varying flow rates of ethyl acetate and measuring the concentration with a photo-ionization detector [34]. Statistics mirroring a turbulent flow [16] were verified (Fig. S9).

For the network model, the AL-to-MB connectivity matrix $\mathbf{J}_1$, is chosen such that each KC connects presynaptically to 7 randomly chosen AL glomeruli [23, 24]. The results shown in Fig. 5 are an average of 10 distinct instantiations of this random topology. The $Z = 2500$ KCs are then connected by a matrix $\mathbf{J}_2$ to a readout layer of dimension Q, where $Q = 2$ for binary and $Q = N_{\mathrm{ID}}$ for multi-class classification. Both AL-to-MB and MB-to-readout connections are perceptron-type with rectified-linear thresholds. The weights of $\mathbf{J}_1$ and $\mathbf{J}_2$ are chosen randomly 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 are updated during supervised network training, via logistic regression (for binary classification) or its higher-dimensional generalization, the softmax cross entropy (for multi-class classification).

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

[2] L. Buck and R. Axel, Cell **65**, 175 (1991).

[3] P. J. Clyne, C. G. Warr, M. R. Freeman, D. Lessing, J. Kim, and J. R. Carlson, Neuron **22**, 327 (1999).

[4] L. B. Vosshall, H. Amrein, P. S. Morozov, A. Rzhetsky, and R. Axel, Cell **96**, 725 (1999).

[5] R. W. Friedrich and S. I. Korsching, Neuron **18**, 737 (1997).

[6] E. Hallem and J. Carlson, Cell **125**, 143 (2006).

[7] G. Wang, A. F. Carey, J. R. Carlson, and L. J. Zwiebel, Proceedings of the National Academy of Sciences **107**, 4418 (2010).

[8] K. Nara, L. R. Saraiva, X. Ye, and L. B. Buck, Journal of Neuroscience **31**, 9179 (2011).

[9] B. Malnic, J. Hirono, T. Sato, and L. B. Buck, Cell **96**, 713 (1999).

[10] J. G. Hildebrand and G. M. Shepherd, Annual review of neuroscience **20**, 595 (1997).

[11] M. de Bruyne, K. Foster, and J. R. Carlson, Neuron **30**, 537 (2001).

[12] R. I. Wilson, Annual Review of Neuroscience **36**, 217 (2013).

[13] M. Renou, V. Party, A. Rouyar, and S. Anton, Biosystems **136**, 35 (2015).

[14] J. Murlis, Annual Review of Entymology **37**, 505 (1992).

[15] M. Weissburg, The Biological Bulletin **198**, 188 (2000), pMID: 10786940.

[16] A. Celani, E. Villermaux, and M. Vergassola, Phys. Rev. X **4**, 041015 (2014).

[17] R. T. Cardé and M. A. Willis, Journal of Chemical Ecology **34**, 854 (2008).

[18] G. Si, J. K. Kanwal, Y. Hu, C. J. Tabone, J. Baron, M. E. Berck, G. Vignoud, and A. D. Samuel, bioRxiv , 208538 (2017).

[19] L.-H. Cao, D. Yang, W. Wu, X. Zeng, B.-Y. Jing, M.-T. Li, S. Qin, C. Tang, Y. Tu, and D.-G. Luo, Nature Communications **8**, 1357 (2017).

[20] C. F. Stevens, Proceedings of the National Academy of Sciences **113** (2016).

[21] K. Asahina, M. Louis, S. Piccinotti, and L. Vosshall, Journal of Biology **8**, 9 (2009).

[22] S. R. Olsen, B. Vikas, and R. I. Wilson, Neuron **66**, 287 (2010).

[23] S. Caron, V. Ruta, L. Abbott, and R. Axel, Nature **497**, 113 (2013).

[24] A. Litwin-Kumar, K. D. Harris, R. Axel, H. Sompolinsky, and L. Abbott, Neuron **93**, 1153 (2017).

[25] K. Krishnamurthy, A. M. Hermundstad, T. Mora, A. M. Walczak, and V. Balasubramanian, bioRxiv **doi:10.1101/160382** (2017), 10.1101/160382.

[26] S. Dasgupta, C. F. Stevens, and S. Navlakha, Science **358**, 793 (2017).

[27] S. L. Brown, J. Joseph, and M. Stopfer, Nature Neuroscience **8**, 1568 (2005).

[28] B. Raman, J. Joseph, J. Tang, and M. Stopfer, Journal of Neuroscience **30**, 1994 (2010).

[29] N. Gupta and M. Stopfer, Neuron , 2247 (2014).

[30] C. D. Wilson, G. O. Serrano, A. A. Koulakov, and D. Rinberg, Nature Communications **8**, 1477 (2017).

[31] S. A. Montague, D. Mathew, and J. R. Carlson, Journal of Neuroscience **31**, 7891 (2011).

[32] M. Stensmyr, H. Dweck, A. Farhan, I. Ibba, A. Strutz, L. Mukunda, J. Linz, V. Grabe, K. Steck, S. Lavista-Llanos, D. Wicher, S. Sachse, M. Knaden, P. Becher, Y. Seki, and B. Hansson, Cell **151**, 1345 (2012).

[33] C. Martelli, J. R. Carlson, and T. Emonet, Journal of Neuroscience **33**, 6285 (2013).

[34] S. Gorur-Shandilya, M. Demir, J. Long, D. A. Clark, and T. Emonet, eLife **6**, e27670 (2017).

[35] E. H. Weber, *EH Weber on the tactile senses* (Psychology Press, 1996).

[36] G. T. Fechner, *Elemente der psychophysik* (Breitkopf und Härtel, 1860).

[37] J. Cafaro, Physiological Reports **4**, e12762 (2016).

[38] L.-H. Cao, B.-Y. Jing, D. Yang, X. Zeng, Y. Shen, Y. Tu, and D.-G. Luo, Proceedings of the National Academy of Sciences **113**, E902 (2016).

[39] K. Nagel and R. Wilson, Nature Neuroscience **14**, 208216 (2011).

[40] J. A. Butterwick, J. del Mármol, K. H. Kim, M. A. Kahlson, J. A. Rogow, T. walz, and V. Ruta, Nature **560**, 447 (2018).

[41] M. N. Getahun, S. B. Olsson, S. Lavista-Llanos, B. S. Hansson, and D. Wicher, PLoS One **8**, e58889 (2013).

[42] M. N. Getahun, M. Thoma, S. Lavista-Llanos, I. Keesey, R. A. Fandino, M. Knaden, D. Wicher, S. B. Olsson, and B. S. Hansson, Journal of Experimental Biology **219**, 3428 (2016).

[43] H. Guo, K. Kunwar, and D. Smith, The Journal of Neuroscience **37**, 9465 (2017).

[44] H. Guo and D. P. Smith, Journal of Experimental Neuroscience **11**, 1 (2017).

[45] A. J. Waite, N. W. Frankel, and T. Emonet, Annual Review of Biophysics **47**, 595 (2018), pMID: 29618219, https://doi.org/10.1146/annurev-biophys-062215-010954.

[46] L. van der Maaten and G. Hinton, Journal of Machine Learning Research **9**, 2579 (2008).

[47] L. B. Vosshall, A. M. Wong, and R. Axel, Cell **102**, 147 (2000).

[48] D. Donoho, IEEE Transactions on Information Theory **52**, 1289 (2006).

[49] E. Candes, J. Romberg, and T. Tao, Communications on Pure and Applied Mathematics **59**, 1207 (2006).

[50] C. Pehlevan, A. Genkin, and D. B. Chklovskii, in *2017 51st Asilomar Conference on Signals, Systems, and Computers* (2017) pp. 593–600.

[51] S. R. Olsen and R. I. Wilson, Nature **452**, 952 (2008).

[52] A. C. Keene and S. Waddell, Nature Reviews Neuroscience **8**, 341 (2007).

[53] R. G. Vogt and L. M. Riddiford, Nature **293**, 161 (1981).

[54] K. Menuz, N. K. Larter, J. Park, and J. R. Carlson, PLoS Genetics **10**, e1004810 (2014).

[55] G. Reddy, J. Zak, M. Vergassola, and V. N. Murthy, bioRxiv , 204354 (2017).

[56] C.-Y. Su, K. Menuz, J. Reisert, and J. Carlson, Nature **492**, 76 (2012).

[57] M. Papadopoulou, S. Cassenaer, T. Nowotny, and G. Laurent, Science **332**, 721 (2011).

[58] N. Gupta and M. Stopfer, Current Opinion in Neurobiology **21**, 768 (2011).

[59] B. Wark, B. N. Lundstrom, and A. Fairhall, Current Opinion in Neurobiology **17**, 423 (2007).

[60] I. Nemenman, in *Quantitative Biology: From Molecular to Cellular Systems*, edited by M. E. Wall (CRC Press, USA, 2012) Chap. 4, pp. 73–91.

Supporting Information Text

Mathematical model

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

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

where ϵ_a (assumed positive) is the free energy difference between the active and inactive conformation of the unbound receptor, and E_{ai} and E_{ai}^* are the free energy differences (assumed positive) between the unbound and bound state for the inactive and active receptor, respectively. $\mu_i = \mu_0 + \beta^{-1}\log(s_i/s_0)$ is the chemical potential for odorant species i in terms of a reference 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 thermodynamic-relevant disassociation constants $K_{ai} = s_0 e^{\beta(-E_{ai} - \mu_0)}$. Adding up contributions from all i odorants, the active fraction is:

$$
A_a = \frac{C_a^* + \sum_i C_a^* s_i}{C_a + \sum_i C_a s_i + C_a^* + \sum_i C_a^* s_i} = \left(1 + e^{\epsilon_a}\frac{1 + \sum_i^N \frac{s_i}{K_{ai}}}{1 + \sum_i^N \frac{s_i}{K_{ai}^*}}\right)^{-1},
\tag{2}
$$

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

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

$$
\begin{aligned}
\text{active} & \qquad \{1, \quad -E_{ai} - \mu_i, \quad -E_{ai} - \mu_i, \quad -2E_{ai} - 2\mu_i\} \\
\text{inactive} & \qquad \{\epsilon_a, \quad -(E_{ai}^* - \epsilon_a) - \mu_i, \quad -(E_{ai}^* - \epsilon_a) - \mu_i, \quad -(2E_{ai}^* - \epsilon_a) - 2\mu_i\}
\end{aligned}
\tag{3}
$$

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

$$
A_a(\text{odorant } i, \text{2 binding sites}) = \left(1 + e^{\epsilon_a}\frac{(1 + \frac{s_i}{K_{ai}})^2}{(1 + \frac{s_i}{K_{ai}^*})^2}\right)^{-1}.
\tag{4}
$$

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

$$
A_a(N \text{ odorants}, R \text{ binding sites}) = \left[1 + e^{\epsilon_a}\left(\frac{1 + \sum_i^N \frac{s_i}{K_{ai}}}{1 + \sum_i^N \frac{s_i}{K_{ai}^*}}\right)^R\right]^{-1}.
\tag{5}
$$

Finally, firing rate dynamics are assumed linear-nonlinear:

$$
r_a(t) = f\left(\int^t h(t - \tau)A(\tau)d\tau\right),
\tag{6}
$$

where $h(t)$ and f are a temporal filter and rectifying linear unit (with threshold $\theta = 5$ Hz) as noted in the main text.

Compressed sensing decoding

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

$$
\mathbf{y} = \mathbf{R}\mathbf{x},
\tag{7}
$$

where $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^M$ are vectors of signals and responses, respectively, and $\mathbf{R}$ is the measurement matrix. Since measurements are fewer than signal components, then $M < N$, whereby $\mathbf{R}$ is wide rectangular and so Eq. 7 cannot be simply

inverted to produce $\mathbf{x}$. The idea of CS is to utilize the knowledge that $\mathbf{x}$ is sparse, i.e.g only K of its components, $K \ll N$ are nonzero. Both the measurements and sparsity are thus combined into a single constrained optimization routine:

$$\hat{x}_i = \operatorname{argmin} \sum_i^N |x_i| \quad \text{such that} \quad \mathbf{y} = \mathbf{R}\mathbf{s} \tag{8}$$

where $\hat{x}_i$ are the optimal estimates of the signal components and the sum, which is known as the L_1 norm of $\mathbf{x}$, is a natural metric of sparsity (1).

The L_1 norm is a convex operation and the constraints are linear, so the optimization has a unique global minimum. To incorporate the nonlinear response of our encoding model into this linear framework, we assume that the responses are generated through the full nonlinear steady state response, Eq. 2- 6, but that the measurement matrix $\mathbf{R}$ needed for decoding uses a linear approximation of this transformation. Expanding Eq. 6 around $\mathbf{s}_0 = \mathbf{s} - \Delta\mathbf{s}$ gives

$$\Delta r_a(t) = r_a(\mathbf{s}(t)) - r_a(\mathbf{s}_0(t)) \tag{9}$$

$$\Delta r_a(t) = \int^t d\tau h(t-\tau) \sum_i^N \frac{dA_{ai}}{ds}\Big|_{\mathbf{s}_0} \Delta s_i \tag{10}$$

$$r_a(\mathbf{s}_0) = \int^t d\tau h(t-\tau) \sum_i^N A_{a0} \tag{11}$$

$$\frac{dA_{ai}}{ds}\Big|_{\mathbf{s}_0} = A_{a0}(1-A_{a0})\left[\frac{1}{K_{ai}}\left(1+\sum_j \frac{s_{0,j}}{K_{aj}}\right)^{-1} - \frac{1}{K_{ai}^*}\left(1+\sum_j \frac{s_{0,j}}{K_{aj}^*}\right)^{-1}\right] \tag{12}$$

where $A_{a0} = A(\mathbf{s}_0)$ and where Eqs. 10 and 11 hold only for integrands above 5 Hz (and are zero below), as per the linear rectifier f. We assume that the neural decoder has access to background $\mathbf{s}_0$, presumed learned (this assumption can be relaxed; see below), and to the linearized response matrix, Eq. 12, but must infer the excess signals Δs_i from excess ORN firing rates $\Delta r_a(t)$. Thus, this corresponds to the CS framework (Eq. 8) via $\Delta\mathbf{r} \to \mathbf{y}$, $\Delta\mathbf{s} \to \mathbf{x}$, and $A'_{ai}\big|_{\mathbf{s}_0} \to \mathbf{R}$. We optimize the cost function in Eq. 8 using sequential least squares programming, implemented in Python through using the scientific package SciPy.

Iterative Hard Tresholding (IHT) and the Restricted Isometry Property in compressed sensing. We stress that the purpose of response linearization is simply to apply compressed sensing reconstruction directly using linear programming, without worrying about issues of local minima in Eq. 8. 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 (2). IHT can be generalized straightforwardly to nonlinear constraints, and would actually dispense with the need for a learned background $\mathbf{s}_0$, simply initializing the iterations from $\mathbf{s}_0 = \mathbf{0}$. Remarkably, this technique works quite well even for non-linear measurements (3). We demonstrate the applicability of the IHT algorithm to our odor decoding system in Fig. S8, which reproduces qualitatively the findings in the main text. For these calculations, no background odor was assumed, each iterative decoding being initialized at the zero vector.

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) (4). Loosely, RIP measures the closeness of matrix operator to 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 $\mathbf{A}$. In particular, if λ_i are the eigenvalues of $\mathbf{A}_k^T \mathbf{A}_k$, where $\mathbf{A}_k$ is any $k \times m$ submatrix of $\mathbf{A}$, and

$$1 - \delta_k \leq \lambda_{min} \leq \lambda_{max} \leq 1 + \delta_k$$

is satisfied for some δ_k, then $\mathbf{A}$ satisfies the RIP with constant δ_k. Plainly, the RIP states that the eigenvalues of $\mathbf{A}_k^T \mathbf{A}_k$, 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 (2):

$$\mathbf{x}_{k+1} = H_K(\mathbf{x}_k + \mu\mathbf{R}^T(\mathbf{x}_k + (\mathbf{y} - \mathbf{R}\mathbf{x}_k))) \tag{13}$$

where $\mathbf{x}_k$ is the kth estimate of the sparse signal $\mathbf{x}$, μ is a step size for the iterations, and $\mathbf{y}$, $\mathbf{R}$ 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 (3):

$$\mathbf{x}_{k+1} = H_K(\mathbf{x}_k + \mu\mathbf{A}_{\mathbf{x}_n}^T(\mathbf{x}_k + (\mathbf{y} - A(\mathbf{x}_k)))), \tag{14}$$

where A is a nonlinear sensing function and $\mathbf{A}_{\mathbf{x}_n}$ is a linearization of A about the point $\mathbf{x}_n$. Reconstructibility for k-sparse signals is guaranteed if $\mathbf{A}_{\mathbf{x}_n}$ satisfies RIP for all $\mathbf{x}_n$ and all k-sparse vectors (2). To get a sense of how this is preserved in the adaptive system, we calculate the eigenvalues for 1000 choices of $\mathbf{x}_n$, acting on random signals of given sparsity K (Fig. S7). Since the RIP is sensitive to constant scalings of the measurement matrix (while the actual estimation problem is not), we scaled all columns of $\mathbf{A}_{\mathbf{x}_n}$ to norm unity (5). This normalizes the eigenvalues of $\mathbf{A}_{\mathbf{x}_n}^T \mathbf{A}_{\mathbf{x}_n}$ 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. 14 (3, 5). The spectra of these matrices indicates that the RIP becomes far more weakly satisifed in the non-adaptive system than in the adaptive one, for sufficient odor complexity and intensity.
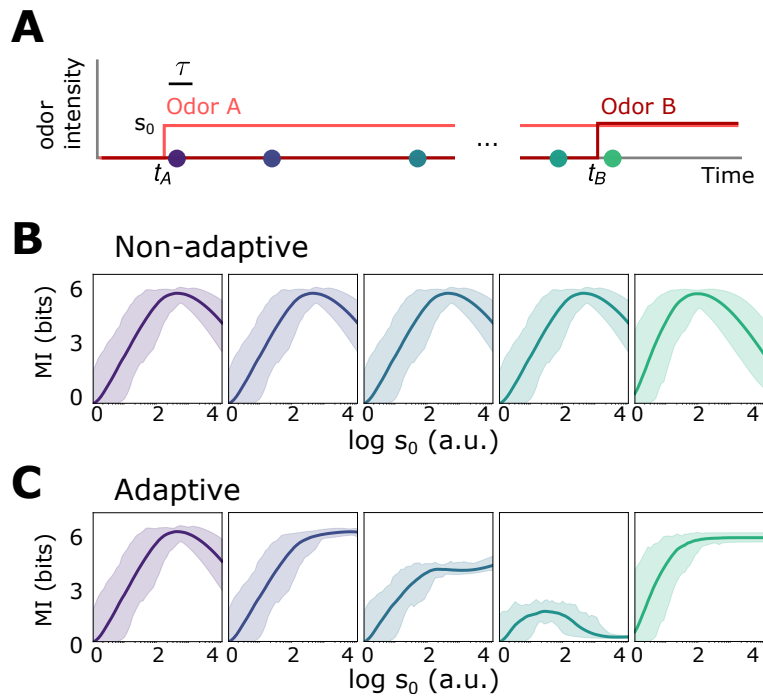
Fig. S1. Front-end adaptive feedback preserves information capacity of the ORN sensing repertoire. Mutual information between signal $\mathbf{s}(t) = \mathbf{s}_A(t) + \mathbf{s}_B(t)$ and response $\mathbf{r}(t)$ is calculated at various points in time t for an odor environment consisting of two step odors, A and B. **A** Odor A, with concentration $\mathbf{s}_A(t)$, turns on at time t_A and a odor B, with concentration $\mathbf{s}_B(t)$, turns on at some later time t_B. Both odors have similar intensities $\sim s_0$ and similar molecular complexity ($k = 4$). **B** Mutual information as a function of s_0 for the non-adaptive system, respectively, at different time points after t_A, corresponding to the dots in A. The mutual information carried by distinct ORNs is represented by the shaded region; their average is plotted by the heavy line. In the non-adaptive system, the mutual information peaks in the regime of high sensitivity after the arrival of odor A (purple, blue), and shifts leftward with the onset of odor B (teal, green). The leftward shifts occurs since stronger signals are more prone to response saturation (compromising information transfer) as odor B arrives. **C** Same as B, now for the adaptive system. The MI mimics the non-adaptive case at the onset of odor A, before adaptation has kicked in (purple). As the system adapts and responses decrease toward baseline, previously saturating signal intensities now cross the regime of maximal sensitivity, which therefore shifts rightward to higher s_0 (dark blue). Much later, but before the arrival of odor B, the ORNs that responded now fire at a similar adapted firing rate ~ 30 Hz, irrespective of odor identity, so the mutual information drops to zero. However, having now adjusted its sensitivity to the presence of odor A, the system can respond appropriately to odor B: the MI at t_B is nearly 6 bits across decades of concentration immediately following t_B (green).
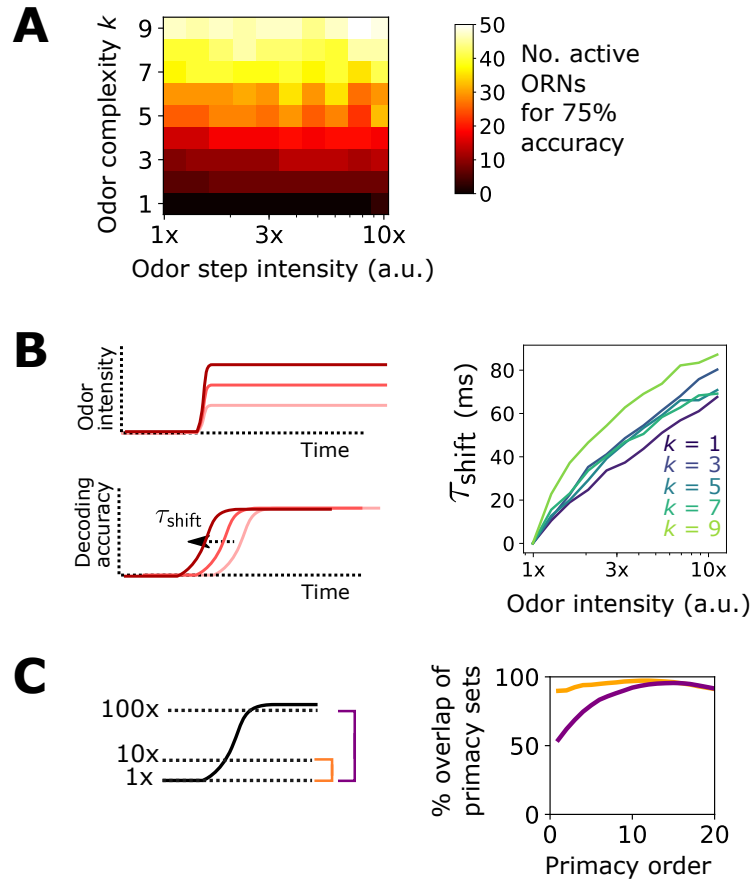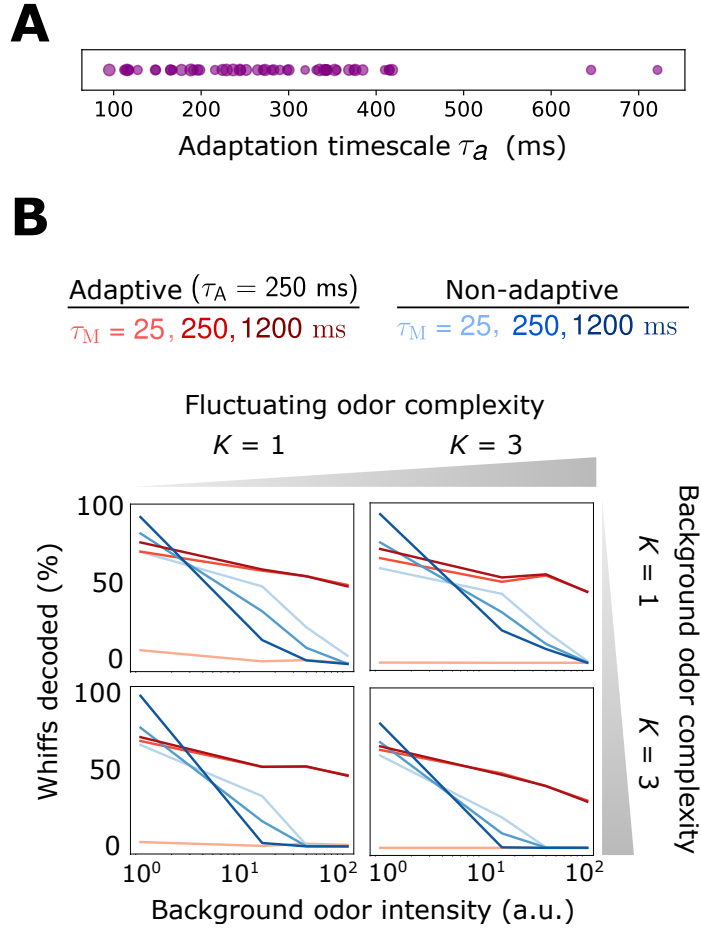
Fig. S2. Additional results pertaining to the primacy coding hypothesis. **A** Percent of active ORNs required for 75% accuracy of a steep sigmoidal odor step, as a function of odor step intensity and odor complexity. For low complexities, a primacy set of fewer ORNs may be sufficient to decode the full 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 (6). We also find this shift in our compressed sensing decoding framework (right plot), which rises monotonically with step height for various odor complexities, in agreement with primacy coding. **C** The consistency of a primacy code across changes in background odor concentration, in a system with Weber Law adaptation. We calculate the primacy set for odor A (step odor; black) in the presence of either a weak, medium, or strong background (dotted lines; 1x, 10x, 100x a.u.), assuming the system has adapted its response to the background as described in the main text. Averaged across odor A identities, primacy sets for odor A when in the 1x background are nearly identical to those when odor A is in the 10x background (right plot; yellow). The same holds true when comparing the 1x and 100x backgrounds, for sufficiently large primacy order, above 8 or so right plot; purple). This indicates that Weber Law adaptation preserves primacy codes across disparate environmental conditions.
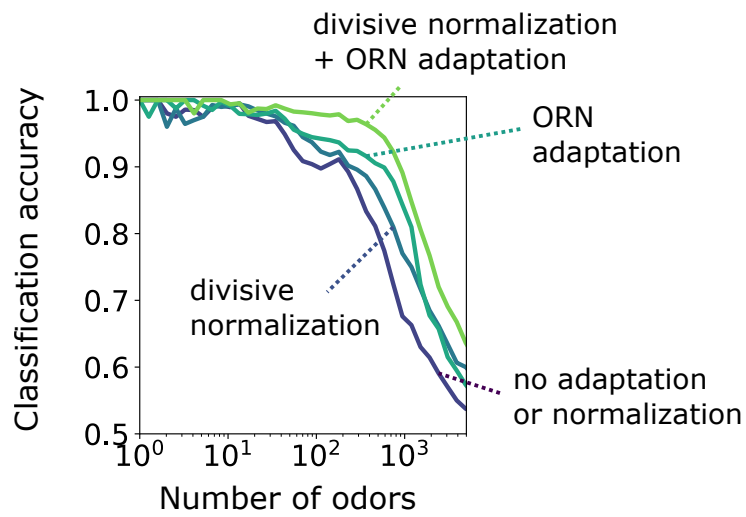
Fig. S3. Accuracy of binary classification by odor valence, for odors whose concentrations span a narrow range of concentrations (1 order of magnitude). Accuracy is plotted as a function of the number of distinct odor identities classified by the trained network, in systems with only ORN adaptation, 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.

Nirag Kadakia and Thierry Emonet

A

100 200 300 400 500 600 700

Adaptation timescale τ_a (ms)

B

Adaptive $\left(\tau_A = 250 \text{ ms}\right)$ | Non-adaptive
$\tau_M = 25, 250, 1200 \text{ ms}$ | $\tau_M = 25, 250, 1200 \text{ ms}$

Fluctuating odor complexity

$K = 1$ $\qquad$ $K = 3$

Whiffs decoded (%)

Background odor complexity

$K = 1$ $\qquad$ $K = 3$

100

50

0

100

50

0

10^0 $\quad$ 10^1 $\quad$ 10^2 10^0 $\quad$ 10^1 $\quad$ 10^2

Background odor intensity (a.u.)

Fig. S4. Decoding accuracy for system with broader distribution of adaptive timescales τ. **A** Distribution of timescales for all ORNs a (purple dots). Here, $\tau_a \sim = 10^X$ where $\tau = 250$ ms as in the main text and $X \sim \mathcal{N}(0, 0.2)$. **B** Individual plots show the percent of accurately decoded odor whiffs (same fluctuating odor signal used in the main text) as a function of background odor intensity, for the non-adaptive (blue) and adaptive (red) systems, for different 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
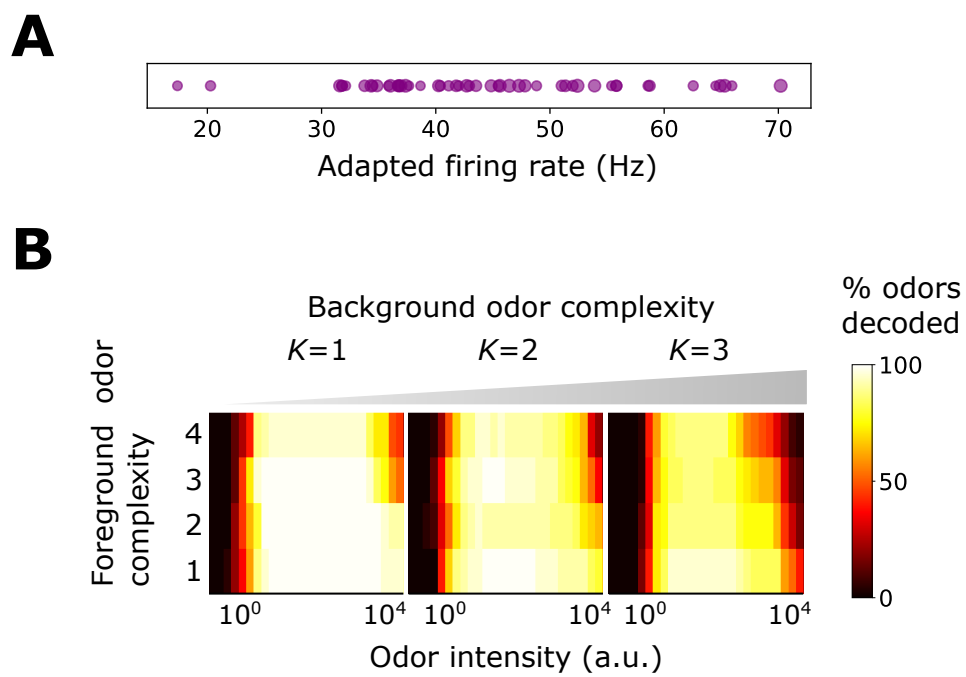
Fig. S5. Benefits conferred by Weber-Fechner adaptation remain for a broader distribution of baseline firing rates A_{a0}, now assumed to be ORN-dependent and chosen from a normal distribution. **A** Distribution of A_{a0}. **B** Decoding accuracy of foreground odors in the presence of background odors.
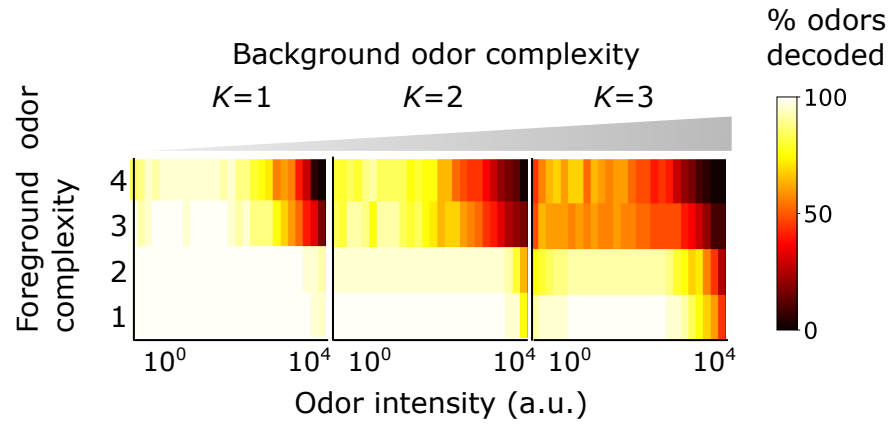
Fig. S6. Benefits conferred by Weber-Fechner adaptation remain for 2 binding sites per receptor. This might conceivably occur in insect olfactory receptors, heterotetramers consisting of 4 Orco/Or subunits that gate a central ion channel pathway (7). Plotted is the decoding accuracy of foreground odors in the presence of background odors.
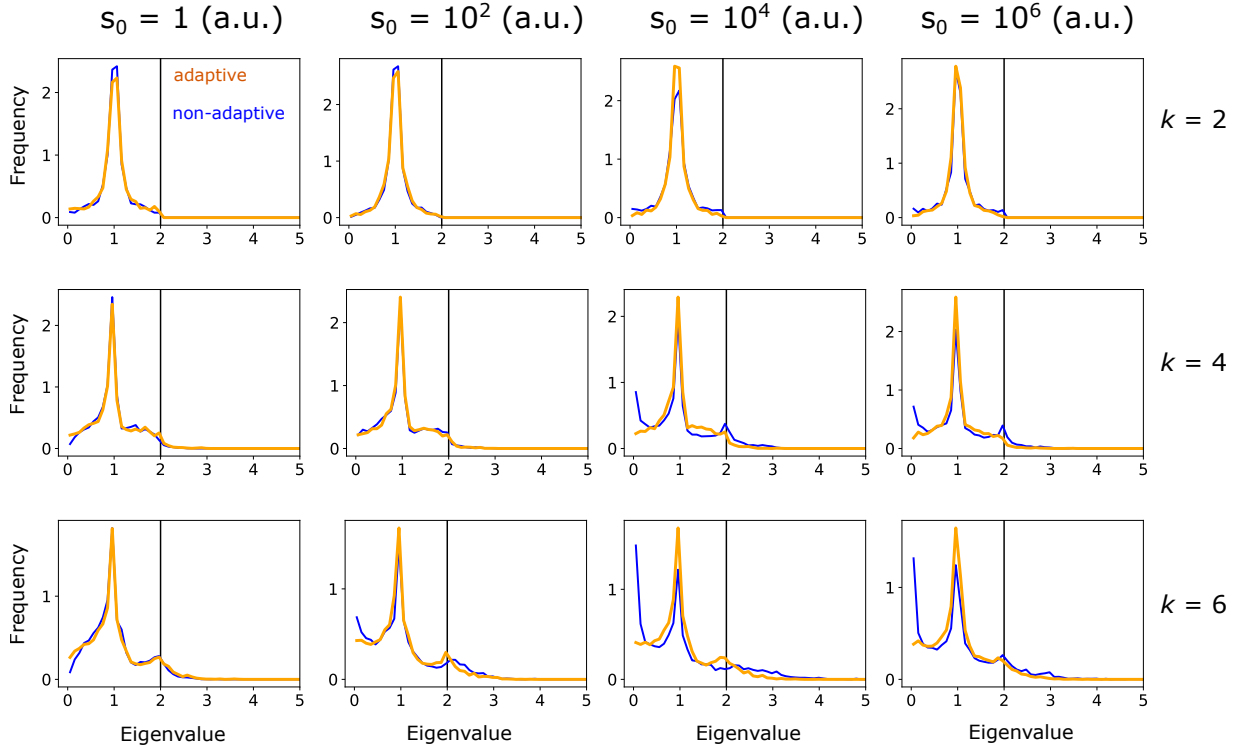
Fig. S7. Eigenvalue distribution of $\mathbf{A}_{\mathbf{x}_n}^T \mathbf{A}_{\mathbf{x}_n}$, where $\mathbf{A}_{\mathbf{x}_n}$ is a $m \times k$ submatrix of the column-normalized linearized ORN response matrix $\mathbf{A}$, evaluated at the linearization point $\mathbf{x}_n$. Note that $\mathbf{x}_n$ is k-sparse, but its components do not necessarily align with the k columns chosen for the sub-matrix. Eigenvalues are calculated for the adaptive (orange) and non-adaptive (blue) systems, for 1000 randomly chosen linearization points $\mathbf{x}_n$ and submatrices. Plots are arranged for various odor sparsities (by row) and odor intensities (by column). The restricted isometry property is satisfied when the eigenvalues lie between 0 and 2 (black vertical line), and is more strongly satisfied the more centered the distribution is around unity. The increase in near-zero eigenvalues for the non-adaptive system at higher odor complexities and intensities (lower right plots) indicates the weaker fulfillment of the restricted isometry property for thhse signals, and leads to higher probability of failure in compressed sensing signal reconstruction.
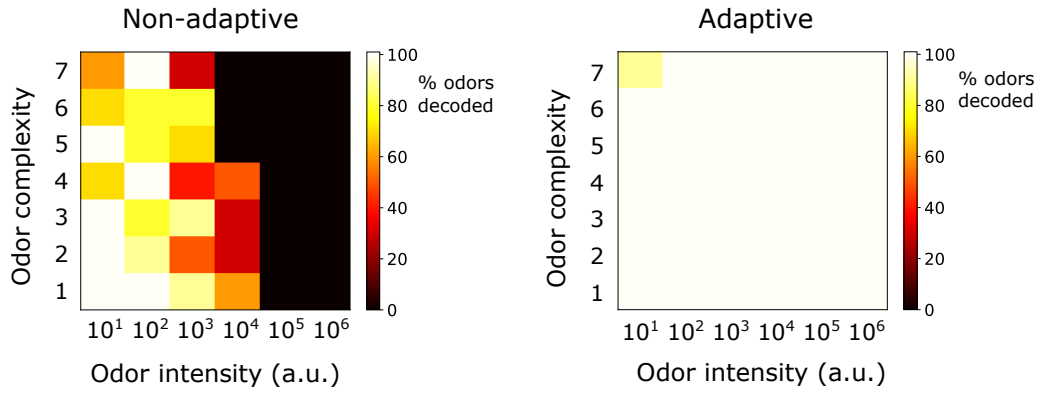
Fig. S8. Decoding of odor signals (no background odors) using the IHT algorithm (2, 3) qualitatively reproduces the results from the main text, which used traditional CS with background linearization. In the adaptive case, IHT actually exhibits superior accuracy to traditional CS, though IHT demands more compute time. The results here show odor decoding accuracy for sparse odor signals of given complexity and intensity, averaged over 10 distinct identities. Odors are considered accurately decoded if the K sparse components are estimated within 25% and the components not in the mixture are estimated below 10% of s_0. The iterative algorithm was initialized at $\hat{\mathbf{x}} = \mathbf{0}$ and run forward until $\hat{\mathbf{x}}$ was stationary, or 10000 iterations were reached. Step size μ in Eq. 14 was set to $s_0/20$. At each step, the linearized response ($\mathbf{A}_{\mathbf{x}_n}$ in Eq. 14) was evaluated at the result of the previous iteration. IHT also requires an assumption on the number of components in the mixture (which defines $H_K(\cdot)$ in Eq. 14); here, that was set to twice the actual sparsity of true signal.
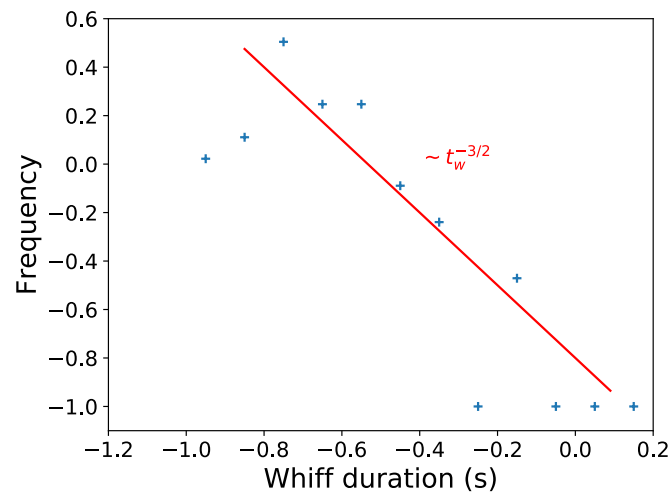
Fig. S9. Distribution of whiff durations in naturalistic stimulus, compared to the theoretical prediction (8).

References

1. Donoho D (2006) Compressed sensing. *IEEE Transactions on Information Theory* 52(4):1289–1306.
2. Blumensath T, Davies ME (2009) Iterative hard thresholding for compressed sensing. *Applied and Computational Harmonic Analysis* 27(3):265 – 274.
3. Blumensath T (2013) Compressed sensing with nonlinear observations and related nonlinear optimization problems. *IEEE Transactions on Information Theory* 69(6).
4. Candes E, Romberg J, Tao T (2006) Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics* LIX:1207–1223.
5. Blumensath T, Davies ME (2009) How to use the iterative hard thresholding algorithm in *Proceedings of SPARS'09 - Signal Processing with Adaptive Sparse Structured Representations (Saint Malo, France, April 2009)*.
6. Wilson CD, Serrano GO, Koulakov AA, Rinberg D (2017) A primacy code for odor identity. *Nature Communications* 8(1):1477.
7. Butterwick JA, et al. (2018) Cryo-EM structure of the insect olfactory receptor Orco. *Nature* 560:447–452.
8. Celani A, Villermaux E, Vergassola M (2014) Odor landscapes in turbulent environments. *Phys. Rev. X* 4(4):041015.