Cortical feedback and gating in odor discrimination and generalization

Gaia Tavoni[1,2,*,+], David E. Chen Kersen[1,3,+], Vijay Balasubramanian[1,2,3,4]

1 Computational Neuroscience Initiative, University of Pennsylvania, United States
2 Department of Physics and Astronomy, University of Pennsylvania, United States
3 Department of Bioengineering, University of Pennsylvania, United States
4 Department of Neuroscience, University of Pennsylvania, United States

+ These authors contributed equally to this work

*** Corresponding author: gaia.tavoni@wustl.edu**

Abstract

A central question in neuroscience is how context changes perception of sensory stimuli. In the olfactory system, for example, experiments show that task demands can drive divergence and convergence of cortical odor responses, likely underpinning olfactory discrimination and generalization. Here, we propose a simple statistical mechanism for this effect based on unstructured feedback from the central brain to the olfactory bulb, which represents the context associated with an odor, and sufficiently selective cortical gating of sensory inputs. Strikingly, the model predicts that both convergence and divergence of cortical odor patterns should increase when odors are initially more similar, an effect reported in recent experiments. The theory in turn predicts reversals of these trends following experimental manipulations and in neurological conditions that increase cortical excitability.

Introduction

Contextual information has a powerful effect on perception across a range of sensory modalities [5, 15, 25, 35, 48, 79, 86, 87]. In olfaction, experiments have demonstrated the influence of context and task demands on the neural representation of odors at different levels in the olfactory pathway. In the olfactory bulb (OB), where odor information is first processed before passing to cortex, context-dependent changes in both single-neuron and collective bulb activity have been observed during and following learning [1, 20, 22, 30, 44, 56, 71, 80, 97]. Context can also reshape representations of odors in the olfactory cortex: when odors are associated with the same or different contexts, the corresponding cortical activity undergoes pattern convergence (increased response similarity) or divergence (decreased response similarity) respectively [16, 84]. The mechanisms underlying such context-induced transformations are of great interest in sensory neuroscience.

The OB and cortex are notably coupled to one another. Mitral cells (MCs) and tufted cells (TCs) from the bulb project to several higher brain regions [85]. In

particular, experiments have highlighted that the anterior piriform cortex (PC) is activated by convergent and synchronous inputs from the bulb: coincident activation of several glomeruli within a short time window [55] is required to induce spiking in cortical pyramidal neurons, a mechanism that is thought to be important for decoding complex combinations of chemical features [18, 90]. In turn, the bulb receives extensive feedback from multiple areas of the central brain, including the PC [10, 11, 58, 61, 69, 70]. This feedback predominantly targets granule cells (GCs) and other OB interneurons, enhancing or suppressing their activity [4, 11, 32, 36, 66, 68], but may also directly excite the MC/TCs [58, 81] or otherwise alter MC/TC activity via neuromodulatory factors such as acetylcholine, serotonin, and norepinephrine [26, 54].

Experiments have demonstrated that this cortical feedback plays a critical role in OB function. Activation of feedback pathways can decorrelate MC output [68] enhancing odor discrimination. Neuromodulation of the OB can also alter odor discrimination and, importantly, adjust the influence of contextual information on activity in the bulb and odor perception [20, 21, 56, 57]. Consequently, disruption of these pathways impairs both these associative [47, 49, 57] and discriminatory abilities [20, 51, 57, 66]. Theoretical and computational studies have also explored possible mechanisms by which feedback may impose these effects. Models of top-down, direct cortical feedback to the bulb which include plasticity in the feedback and in cortex have been able to reproduce odor association with visual context [56] and differences in cortical reorganization during passive vs. active learning [100]; demonstrate adaptation to specific olfactory environments and odor tasks when guided by neurogenesis [2]; and explain differential responses to the same odor under different contexts [49]. Other models have demonstrated further effects of neuromodulation on OB activity, such as normalization of output neuron response [52], increasing spike synchrony [20, 50], and general enhancement of odor discrimination [51, 52].

The potential role of such plastic feedback in driving contextual changes in odor discrimination and generalization is complicated by the apparently random projection of MC/TCs to the cortex [18, 34, 88–90]. In particular, any changes induced by plasticity in feedback to the bulb would appear to become scrambled in cortex, necessitating further plasticity at the synapses between MC/TCs and cortical cells, and within the piriform cortex, to produce targeted effects. Indeed, this dispersion of information from the bulb may underpin the distributed character of the cortical representation of odors, with different odorants activating unique, but completely dispersed groups of cortical neurons [90]. Likewise, cortical feedback fibers are distributed diffusely over the OB without any discernible spatial segregation [10, 61]. Consequently, it is difficult to predict how different cortical feedback patterns affect overall OB output and subsequent odor representation in cortex. Thus, global reorganization of the OB network through synaptic plasticity to reflect the odor environment and its associated contexts may also require significant time. This raises the question of how animals are able to effectively learn in the short term under these constraints [93].

Here, we postulate that diffuse feedback signals, carrying unstructured representations of context, can nonetheless modulate odor responses in the OB even *without* synaptic plasticity to provide a rapid trigger for effectively entraining robust convergence or divergence of odor patterns in piriform cortex (PC). These changes, which can be stabilized by plasticity in the recurrent connectivity of cortex [8, 9, 29], can in turn underpin generalization and discrimination in sensory behavior (Fig. 1A). To test this hypothesis, we constructed a statistical, analytically tractable model of the OB and its projections to the PC, which we later extended by incorporating an anatomically-faithful network of interactions between OB excitatory and inhibitory cell types along with a realistic distribution of projections from the OB to the PC. Under minimal assumptions about the statistics of odor and feedback inputs, we show that

changes in bulb firing rates driven by unstructured feedback lead to divergence or convergence of cortical odor representations. The model predicts that the extent of both pattern divergence and, counterintuitively, pattern convergence, should increase with initial odor similarity, a prediction which matches results from recent experiments [16, 83, 84]. These results are especially surprising, since we might expect that incorporation of the semantic context of stimuli would require that similar contexts lead to greater convergence in the representation of initially *dissimilar* odors, the opposite of what we predict and experiments observe. Notably, our results require that cortical units selectively gate coincident inputs, which PC pyramidal neurons are known to do [55, 90]. The model also predicts that increases in cortical excitability, either by experimental manipulation or in neurological conditions such as Alzheimer's disease [96, 98], will alter or even reverse these trends, and hence induce characteristic modifications of contextually-driven change in odor perception and behavior.

Results

A statistical mechanism for associating contexts to odors

Our model contains a sensory layer and a cortical layer, the olfactory bulb (OB) and the piriform cortex (PC) respectively, consistent with the basic anatomical features of the olfactory system (Fig. 1B). The sensory (OB) layer consists of N "modules", where each module represents the set of mitral cells (MCs) in the OB that project to the same pyramidal neuron in the cortical (PC) layer. Consequently, the cortical (PC) layer consists of N units as well, with each unit representing a pyramidal neuron receiving direct inputs from its corresponding module in the sensory layer.

Modeling odor inputs

We summarized the response of the ith OB module to an odor input by the change in the total firing of its constituent MCs R_i compared to baseline, where $0 < R_i < R_{\max}$. Thus the response for the bulb as a whole was described by the module firing rate vector

$$\mathbf{R} = (R_1, R_2, \ldots, R_N) \tag{1}$$

These responses were then transformed into a cortical response vector $\mathbf{K}$ via an element-wise nonlinear activation function $f(\mathbf{R})$, such that:

$$\mathbf{K} = (K_1, K_2, \ldots, K_N) = (f(R_1), f(R_2), \ldots, f(R_N)) \tag{2}$$

Modeling context-induced feedback inputs

We modeled context as the effect produced by cortical feedback inputs to the OB. In the OB, such feedback can lead to increases or decreases in MC firing and consequently in the strength of inputs to PC, represented in our model by the module firing rates. Cortical excitation of interneurons such as GCs and periglomerular cells reduces MC firing [10, 11, 58, 61, 69, 70, 81], while other modes of feedback, such as direct excitation of MCs [58, 61], neuromodulation of MC excitability [21, 27, 37, 53, 72], and excitation of deep short axon cells (which drives feedforward inhibition of GCs) [11, 66, 68] enhance MC activity. As a result, we represented the effect of feedback as a change in module firing rates:

$$\Delta\mathbf{R} = (\Delta R_1, \Delta R_2, \ldots, \Delta R_N) \tag{3}$$

Where $\Delta R_i < 0$ and $\Delta R_i > 0$ respectively represent decreases and increases in the overall firing rate of the ith module.

Additionally, we assumed that the similarity between changes in module responses induced by any two given feedback inputs, which we quantify using the cosine distance between the respective $\Delta\mathbf{R}$s, reflects the similarity between the contexts that elicit those feedback inputs. (The cosine distance between vectors $\mathbf{v}$ and $\mathbf{w}$ is defined as $\mathbf{v} \cdot \mathbf{w}/||\mathbf{v}||\,||\mathbf{w}||$ where $\cdot$ is the inner product and $||\mathbf{v}||$ is the norm of $\mathbf{v}$). In the most extreme case, feedback inputs that induce identical changes in the modules' responses represent identical contexts, e.g. two odors being associated with a reward at the same location (Fig. 1A, left). Conversely, those that induce uncorrelated or anticorrelated changes represent different contexts, e.g. two odors being associated with a reward at different locations (Fig. 1A, right).

Defining cortical similarity

The similarity ρ_i between cortical responses to two odors A and B before feedback was quantified by the cosine distance between the $\mathbf{K}$ vectors for the two odors. Then, the similarity ρ_f between cortical responses for those two odors after feedback was simply the cosine distance between the new $\mathbf{K}$ vectors, where $\mathbf{K} = f(\mathbf{R} + \Delta\mathbf{R})$.

Change in similarity between odor representations following feedback varies linearly with the initial similarity

To capture the overall features of the system while maintaining analytical tractability, we made the following simplifications. Firstly, since experiments have indicated that piriform neurons only respond if the amount of input they receive within a short window is above a certain threshold [18, 55], we approximated $f(\mathbf{R})$ as a step function:

$$K_i = f(R_i) = \begin{cases} 0 & \text{if } R_i < \theta_c \\ 1 & \text{if } R_i \geq \theta_c \end{cases} \tag{4}$$

where θ_c quantifies the threshold for cortical activation. Thus, in our model, the ith cortical neuron receiving input from its module was said to be "active" (i.e. $K_i = 1$) only when $R_i \geq \theta_c$. Because $f(\mathbf{R})$ is a step function, we can express the cortical similarity as:

$$\rho = \frac{C_{AB}}{\sqrt{C_A C_B}} \tag{5}$$

where, before feedback, C_{AB} is the number of PC neurons active for both odors, and C_A and C_B are the numbers of PC neurons active for odors A and B respectively (Fig. 1C).

As a second approximation, we set the module firing rate distribution to be roughly uniform above and below a value θ_m, which we denote the "response threshold". Modules with responses above this threshold are considered significantly "responsive" to that odor, but since $\theta_m \ll \theta_c$ [18, 55], a responsive module may not necessarily be cortically active, as described above. With this approximation of uniformity, we can interpret the problem geometrically (Fig. 1D). If we define N_{AB} as the number of modules responsive to both odors A and B, and N_A and N_B as the numbers of modules responsive to odor A or odor B respectively, then C_{AB}, C_A, and C_B are just rescaled versions of N_{AB}, N_A, and N_B respectively, and can be re-expressed accordingly:

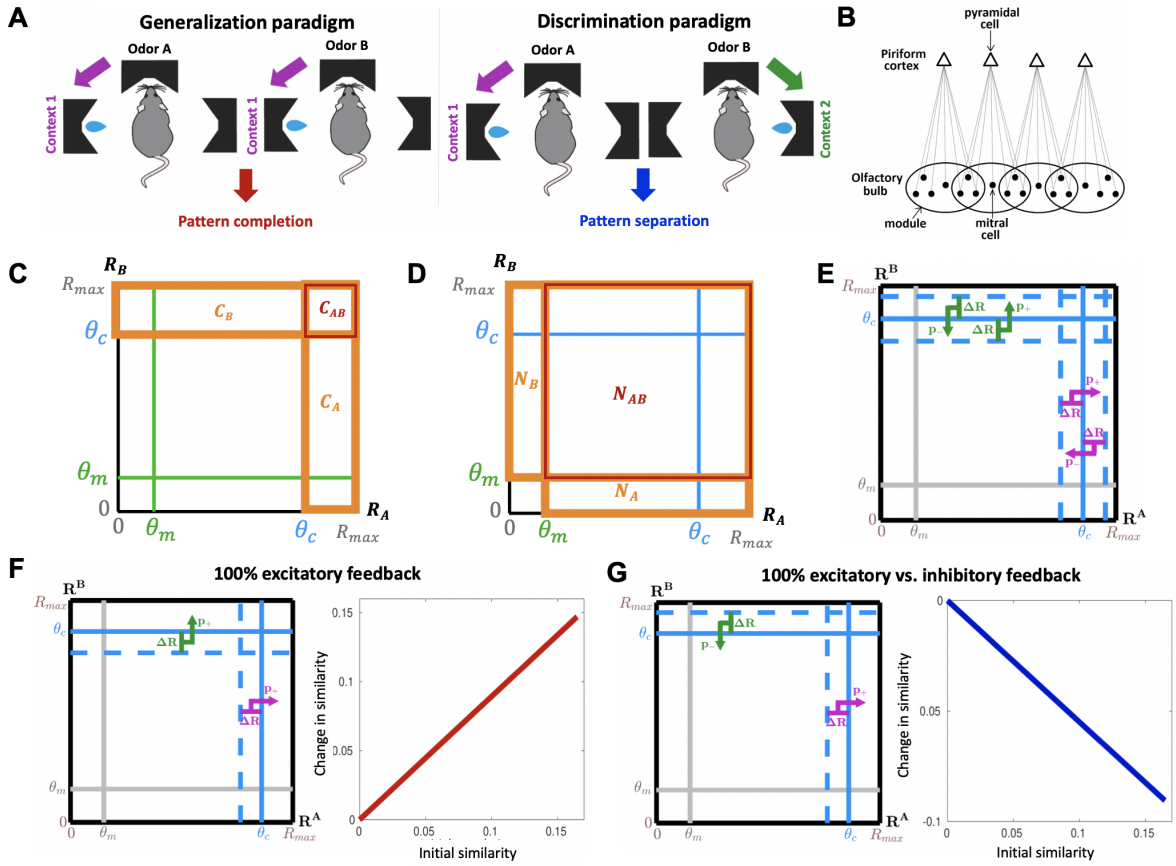

Figure 1. A statistical mechanism for context-induced pattern convergence and divergence in the olfactory system. (**A**) Context is information present in the environment that is salient to behaviour, e.g. the location of a reward associated with an odor. Left: in two-alternative forced-choice tasks, when distinct odors are associated with the same context (e.g. a left reward port), animals learn to generalize their choice (approaching the left port) across odors; after training, the neuronal ensemble responses to the odors in the PC are more correlated (pattern convergence, [16]). Right: when odors are associated with alternative contexts (e.g. opposite reward ports), animals learn to discriminate the stimuli and make different, odor-specific, choices (approaching the left versus right port); after training, the cortical ensemble responses are less correlated (pattern divergence, [16]). (**B**) In the statistical model, the olfactory bulb is represented as an ensemble of "modules" (black circles), with each module defined as the set of MCs (black dots) that project to the same cortical neuron (triangle), and the activation probability of cortical cells is a threshold function of the modules' responses. (**C**) A geometrical interpretation of Eq. 5. The square represents the space of module responses to two odors (A and B). The number of cortical cells activated by odor A (C_A in Eq. (5)) is equal to the number of modules with responses higher than the cortical activation threshold θ_c (orange rectangle labeled C_A); likewise for C_B (orange rectangle labeled C_B) and C_{AB} (red rectangle). (**D**) N_A (orange rectangle labeled N_A), N_B (orange rectangle labeled N_B), and N_{AB} (red rectangle) represent the number of modules responsive to odor A, odor B, or both odors, respectively. Because of the assumption of uniformity, C_A, C_B, and C_{AB} can be re-expressed in terms of these values. (**E**) Stochastic feedback representing contextual changes in the distribution of module responses. Modules with responses to either odor within ΔR from the cortical activation threshold and targeted by positive/negative feedback (with probabilities p_+ and p_-, respectively) are brought above/below the cortical threshold, changing similarity between the cortical odor representations (Eq. 7). (**F**) Left: When feedback is identical and positive to all bulb modules for both odors, increasing their responses by ΔR, its effect on the cortical activity is equivalent to a decrease of the cortical activation threshold by ΔR. This results in a positive change in similarity between the cortical odor representations (i.e. pattern convergence). Right: The amplitude of this effect increases linearly with initial odor similarity. (**G**) Left: When the feedback is positive to all bulb modules for odor A and negative for odor B, the change in cortical activity is equivalent to that induced by a decrease/increase of the cortical activation threshold by ΔR for odor A/B respectively. This results in a decrease in similarity between cortical odor representations, i.e. pattern divergence. Right: The amplitude of this effect increases linearly with initial odor similarity.

$$C_A = \frac{R_{\max} - \theta_c}{R_{\max} - \theta_m} N_A \tag{6a}$$

$$C_B = \frac{R_{\max} - \theta_c}{R_{\max} - \theta_m} N_B \tag{6b}$$

$$C_{AB} = \left(\frac{R_{\max} - \theta_c}{R_{\max} - \theta_m} \right)^2 N_{AB} \tag{6c}$$

As a third approximation, for modules that receive non-zero feedback, we set the magnitude of the change in the module firing rate due to feedback to be the same for all modules.

General patterns of feedback, which vary in the number of modules they reach as well as in their ratio of excitation to inhibition, require an examination of many alternative cases. In the Methods, we carry out this examination and show analytically that for a pair of odors associated to specific contexts, each of which produces unstructured cortical feedback, the general relationship between the initial similarity ρ_i of their cortical representations and the change in that similarity following feedback $\Delta\rho$ is linear:

$$\Delta\rho = Q_2\, \rho_i + Q_1 \tag{7}$$

with Q_1 and Q_2 being complicated functions of the parameters of the feedback and the response thresholds defined above (full derivation in Methods). The linear relationship can be understood intuitively as follows: feedback changes odor similarity by pushing some modules above or below the cortical threshold (depending on the feedback sign); these activated/deactivated modules are typically close to the cortical threshold and thus odor-responsive before feedback (Fig. 1E); thus, they can be expressed as a fraction of the number of odor-responsive modules. Since the response similarity ρ_i can also be re-expressed in terms of the number of odor-responsive modules (Eqs. 5, 6), one component of $\Delta\rho$ will be proportional to ρ_i, hence the $Q_2\, \rho_i$ term in Eq. 7. If feedback is sufficiently strong compared to the difference between the cortical and response thresholds, it may also activate for both odors some modules that are not odor responsive. This effect gives rise to the constant term Q_1. Note that the effects we are reporting are not a simple result of similar and different feedback producing corresponding changes in firing rates. Indeed, because we are considering unstructured feedback to the olfactory bulb, which then projects to cortex, we will see that the effects can be partially *reversed* by increasing cortical excitability, an observation that may serve as a potential experimental test.

To illustrate the general results we can examine special cases. For example, suppose all modules receive the same excitatory feedback of strength ΔR for both odors, so all modules with responses to odor A or B that are close enough to the cortical threshold ($\theta_c - \Delta R < R_i < \theta_c$) are brought above threshold by the feedback; as a result, the cortical neurons to which they project are activated by the feedback. These context-induced changes are equivalent to an effective decrease of the cortical activation threshold by ΔR for both odors (Fig. 1F). Thus, by computing the similarity via Eq. 5 before and after replacing θ_c with $\theta_c - \Delta R$ in Eqs. 6, the resulting change in similarity can be expressed:

$$\Delta\rho = \frac{\Delta R}{R_{\max} - \theta_c} \rho_i \tag{8}$$

which demonstrates pattern convergence that increases linearly in magnitude with initial odor similarity.

Similarly, we can consider a simple situation where all modules receive excitatory feedback for the first odor and inhibitory feedback for the second odor, both with

uniform strength ΔR. By an argument similar to the one above for the case of identical feedback, these effects are equivalent to an effective decrease and increase of θ_c by ΔR for the first and second odors respectively (Fig. 1G). By then computing the similarity before and after replacing θ_c with $\theta_c - \Delta R$ for the first odor and $\theta_c + \Delta R$ for the second in Eqs. 6, the change in similarity can be expressed:

$$\Delta\rho = \left(\frac{\sqrt{(R_{\max} - \theta_c)^2 - \Delta R^2}}{R_{\max} - \theta_c} - 1 \right) \rho_i \,. \tag{9}$$

which again demonstrates pattern divergence that increases linearly in magnitude with initial odor similarity.

The general scenario, with different ratios of excitatory and inhibitory feedback, different numbers of modules affected for the two odors, and different correlations between feedback patterns has an intricate parameter dependence in the linear coefficients of Eq. (7) (see Methods and Supporting Information). To illustrate these general results, we computed the slope Q_2 under a general feedback pattern in which $\sim$50% of the modules were affected by feedback, with the majority of that feedback ($\sim$75%) being negative (i.e., suppressing firing rates), and whose feedback strength ΔR was $\sim$20% of the maximum firing rate $R_{\max}$. Fig. 2A shows that under these conditions, when the cortical threshold θ_c was high (so that about 10% of the cortical units were activated by each odor, as is typical experimentally [90]), highly correlated contexts induced pattern convergence, while uncorrelated or anticorrelated contexts led to pattern divergence. Moreover, this pattern convergence and divergence both increased linearly with the degree of initial similarity in the odor responses, just as in the special cases described previously. These results held across: (i) different feedback strengths, and over larger ranges of cortical thresholds for stronger feedback, (ii) different fractions of modules targeted by feedback, (iii) different fractions of negative versus positive feedback, and (iv) different overlaps between the sets of modules targeted by feedback for the two odors (Figs. 2B-E) (although if the fraction of negative feedback was low, the transition between pattern divergence and pattern convergence could occur at an intermediate value of the feedback correlation even for low cortical thresholds (Fig. 2D, left)). These results also persisted when the module firing rate distribution was Gaussian and the nonlinearity was changed from a theta function to a sigmoid function (Supporting Information, Fig. S1), but they changed dramatically if the cortical layer and its associated nonlinearity were removed (i.e. for module i, $f(R_i) = R_i$) (Supporting Information, Fig. S2A).

Mechanistic model

Above, to facilitate mathematical analysis, we described activity in the OB and the effects of feedback in terms of independent firing rates (and changes in firing rates) of OB modules. However, anatomically, these modules comprise overlapping groups of MCs, and feedback targets individual MCs along with the GCs that inhibit them [11, 58]. Moreover, these MCs and GCs are highly interconnected in a network. Consequently, the precise effect that feedback has on MC output during combined odor driven and feedback activation is difficult to anticipate. To account for these considerations, we refined our approach with a detailed mechanistic model employing a biophysically realistic network of the OB (Fig. 3A,B).

Briefly, to create this network, we modeled each cell in terms of its dendritic tree, since interactions between MCs and GCs are dendrodendritic [24, 64, 76, 77, 95]; roughly mimicking the anatomy, MC trees were laid out in laminar discs and GC trees projected up in inverted cones [14, 33, 63, 67]. These cells were embedded in a 3-dimensional space representing the layers of the OB. The x-y location of the center of a MC dendritic disc

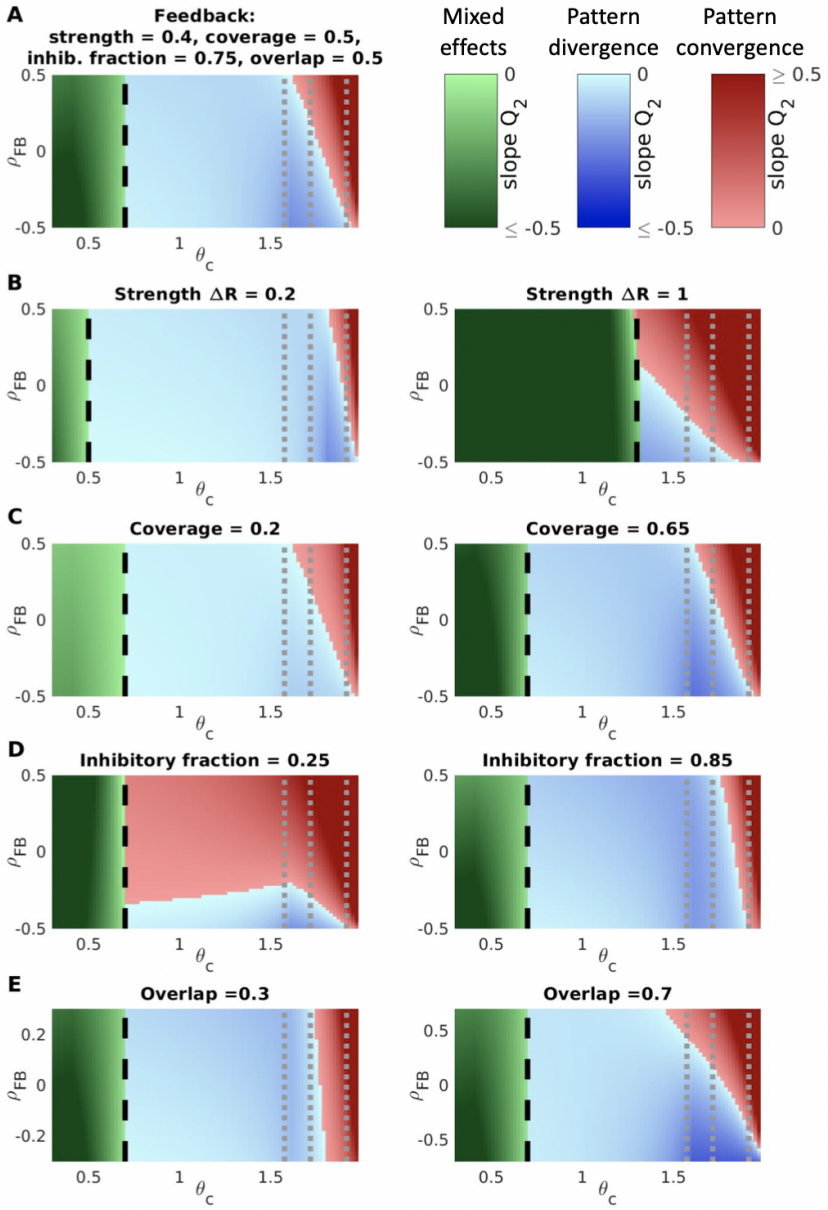

Figure 2. Pattern convergence and divergence for general feedback. (**A**) We computed the slope in Eq. 7 as a function of the cortical threshold θ_c and the feedback similarity ρ_{FB} for the following set of plausible feedback conditions: 1) feedback strength $\Delta R = R_{\max}/5$; 2) feedback coverage $p_+ + p_- = 0.5$ for both odors; 3) fraction of feedback-affected modules that receive inhibitory feedback for one of the odors $p_-/(p_+ + p_-) = 0.75$ (the inhibitory fraction for the other odor varies with the feedback correlation along the y-axis); overlap between the subsets of modules affected by the feedback for the two odors $p_{\text{both}} = 0.5$. Pattern convergence and divergence increasing with initial odor similarity arise at values of the cortical threshold overlapping with the realistic range, with correlated and anticorrelated feedback, respectively. (**B**–**E**) Results generalize to a range of feedback conditions. Panels show effects of changing one feedback parameter, while maintaining others as in **A**. Color maps represent values of slope coefficient Q_2 from Eq. 7, with darker shades corresponding to higher $|Q_2|$, as a function of cortical activation threshold θ_c and feedback similarity ρ_{FB}. **Red**: pattern convergence increasing with odor similarity ($Q_2 > 0$ and $Q_1 = 0$). **Blue**: pattern divergence increasing with odor similarity ($Q_2 < 0$ and $Q_1 = 0$). **Green**: mixed effects, i.e. pattern convergence decreasing with odor similarity and potentially turning into pattern divergence at high similarity ($Q_2 < 0$ and $Q_1 > 0$). **Gray dotted lines**: realistic range for the cortical threshold, activating 15%, 10% and 3% of cortex respectively in response to odor input [90]. **Black dashed line**: a threshold $\theta_c = \theta_m + \Delta R$ approximating the transition between pattern divergence/convergence and mixed effects. All panels: $\theta_m = 0.3$, $R_{\max} = 2$, $p_{\text{odor}} = 0.6$ (probability that a given module has a significant response to an odor input)

depended on the location of its respective glomerulus (which was in turn placed randomly) [45, 73], while its z-position depended on whether it was a Type I or Type II MC [14, 63, 67]. The GCs in our model were deep-type (since MCs primarily interact with deep GCs) [33], so while the x-y location of the vertex of a GC tree was random, the z-position of the vertex of its tree was confined to roughly the lower half of the model space. The probability of a connection between any given MC and GC was then a function of the geometric overlap between their dendritic trees (Fig. 3A, details in Methods).

To simulate network dynamics, MCs and GCs were modeled as Izhikevich neurons [40], with parameters selected to match known electrophysiological data [3, 14, 33]. Odor input was represented as an oscillatory current into the MCs, while positive and negative feedback to the system were represented as pulses of constant excitatory current to a random subset of the MCs and GCs respectively. We determined that a sufficiently large number of MCs (¿5000) was required to achieve statistically stable results (Fig. S4A). The large size (especially given that there are around an order of magnitude more GCs than MCs [13]) meant that computational limitations precluded us from directly performing the spiking simulation for all the different odor and feedback tests. We circumvented this limitation by using the spiking model to extrapolate distributions of firing rates in MCs following odor input and reciprocal feedback from GCs, both with and without external feedback to MCs and GCs (Methods and Fig. S3). This allowed us to simulate the detailed dynamics of larger numbers of MCs efficiently by summarizing the effects of the granule cell network in terms of the distribution of effects on MC firing.

To simulate cortical responses, we allowed K cells to sample randomly from a fraction q of the M MCs, with $q \sim 0.07$, consistent with experimental measurements [89]. Each group of sampled MCs thus constituted a module in the terminology used above. MCs were also partitioned into G glomerular, non-overlapping sets, each representing the set of MCs associated to the same glomerulus. We modeled the response to an odor as the evoked firing rate over a single sniff (a length of time sufficient for a rodent to distinguish between odors [94]) in a fraction f_{odor} of the glomeruli. Thus, the MCs in $f_{odor}G$ glomeruli had a non-zero probability of having a firing rate greater than 0, while the remaining MCs had vanishing firing rates, after subtracting baseline responses. For the fraction of MCs that were selected to receive odor input, the firing rate for each MC was drawn from a distribution fitted to data from the detailed spiking model of the bulb (Fig. S3).

We determined the input to each cortical unit from the sum of the firing rates of MCs that projected to it. To account for the effects of cortical balancing [91], we subtracted the mean cortical input from the input to each unit, and then passed the result through a nonlinear activation function (a sigmoid). This ensured that the strongly activated cortical units tended to receive the highest rate, and hence most coincident, inputs. This sequence of steps yielded a vector of odor-induced firing rates over a sniff as the cortical representation of the odor.

Finally, we modeled feedback as a vector of firing rate changes in a fraction f_{FB} of randomly selected mitral cells. These changes could be induced either by direct feedback to the mitral cells [58, 81] or indirect inhibitory feedback through the granule cells [11, 38, 58, 61, 68, 70], all of which have the net effect of modifying MC firing rates. The specific pattern of feedback, i.e. the affected MCs and the net change in firing rate of each MC, effectively defined the context associated to the odor presentation. Simulations in the spiking network showed that inhibitory feedback patterns in the MCs arising indirectly from excitatory feedback to the GCs were necessarily diffuse and unstructured. Specifically, targeting non-overlapping sets of about the same number of GCs nonetheless produced highly correlated changes in MC firing rates [46]. Thus, in

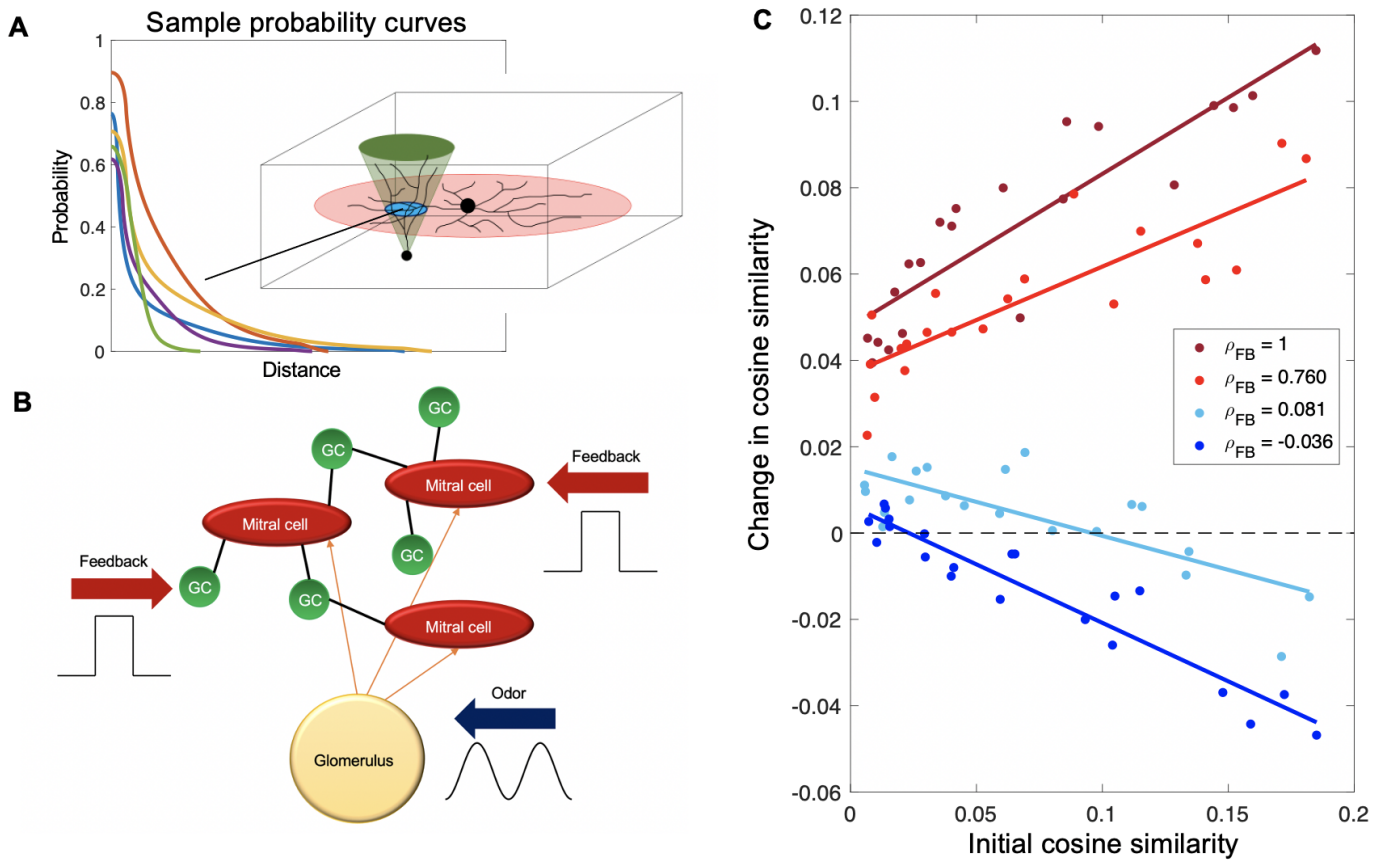

Figure 3. A mechanistic model produces trends in pattern convergence and divergence similar to those predicted by the statistical framework. (**A**) The probability of connection is a function of the overlap of the mitral cell and granule cell dendritic trees, here represented as a disc and cone, respectively. Different parameters in the mean-field distributions of these dendritic trees produce different probability curves for any given MC-GC pair. (**B**) The algorithm in **A** is used to generate a network of MCs interconnected by GCs. Network dynamics are simulated via the Izhikevich model [40], with each neuron in the network receiving external inputs in the form of odor or feedback as well as reciprocal inputs resulting from the connectivity of the network. Odor input is modeled as an oscillatory current to mimic the respiratory cycle while external feedback is modeled via a constant current input. (**C**) For high cortical thresholds, the normalized overlap between odor representations increases (pattern convergence) or decreases (pattern divergence) linearly in the initial overlap, according to the similarity in the feedback inputs. In all cases, we simulated 10,000 MCs grouped into 500 glomeruli and 100,000 cortical cells, each sampling 7% of the MCs, with odor targeting 12% of the glomeruli, positive feedback targeting 8% of the MCs, and negative feedback targeting all MCs.

our mechanistic model, the identity of the MCs which are affected by negative feedback through the GCs depends on the particular MC-GC connectivity of the given network, while the degree to which their firing rate is affected depends on the fraction of GCs targeted.

An *in silico* experiment on pattern convergence vs. pattern divergence

We tested how similarities and differences in feedback affected the cortical representation of odors in our mechanistic model. To this end, we generated pairs of odors, with each odor targeting $f_{odor}G$ glomeruli, where the two odors shared different fractions of targeted glomeruli. We then computed the cosine distance ρ between cortical responses to pairs of odors before and after addition of feedback for different feedback regimes.

We first considered the situation where the threshold of the sigmoidal activation function is high, leading to sparse firing in the cortex as seen in experiments [39, 74, 78, 90]. We presented pairs of odors with different degrees of initial similarity and different amounts of feedback similarity. We found that strongly correlated feedback led to pattern convergence (increased overlap in cortical responses). Likewise anticorrelated feedback led to pattern divergence (decreased overlap in cortical responses). Notably the pattern convergence and divergence both increased linearly with the initial overlap of cortical responses (Fig. 3C).

All told, these results recapitulated the predictions of the statistical model in a detailed mechanistic setting. One difference from the abstract analysis is that in this mechanistic model the only form of direct negative feedback goes through the GC network and is thus non-specific to particular MCs. As a result, strongly anticorrelated feedback is hard to achieve, but may be possible in the brain through targeted neuromodulatory effects or through feedback that suppresses GCs. In addition, because the MCs are embedded in a GC network, excitatory feedback to MCs necessarily induces some inhibitory feedback disynaptically through the granule cells. Thus, in this realistic mechanistic model there are constraints on the achievable forms of positive and negative feedback. Stronger feedback antisimilarity would further enhance pattern divergence.

A prediction: partial reversal of effects when cortical excitability is increased

In the normal brain, pyramidal cells in the PC have a high threshold for activation and respond only when there is coincident input from multiple MCs [55, 82, 90]. However, both experimental manipulation and neurological conditions such as Alzheimer's disease can cause increases in cortical excitability, or, equivalently, lower the cortical activation threshold [96, 98]. Our model makes striking predictions for these conditions that can be tested experimentally.

First, Fig. 2 shows that pattern convergence only arises by our proposed mechanism within a range of cortical activation thresholds that are high, and these thresholds include the typical values expected for pyramidal neurons to achieve realistic levels of cortical activation [90]. Thus we predict that increased excitability in the PC (or, equivalently, reduced activation thresholds) will impair pattern convergence and thus the behavioral ability to generalize. Specifically, if cortical excitability increases moderately, our model predicts that any feedback inputs (anticorrelated as well as correlated) will induce a weak pattern divergence effect, increasing with initial odor similarity (light blue regions in between the black and gray dashed lines in Fig. 2). If the increase in cortical excitability is sufficiently large, our model predicts more complex effects (green regions in Fig. 2), as explained below. Qualitatively, in the high-threshold regime, only modules that already have a significant response to an odor input ($R_i > \theta_c - \Delta R > \theta_m$) can exceed the cortical threshold due to feedback, whereas if the

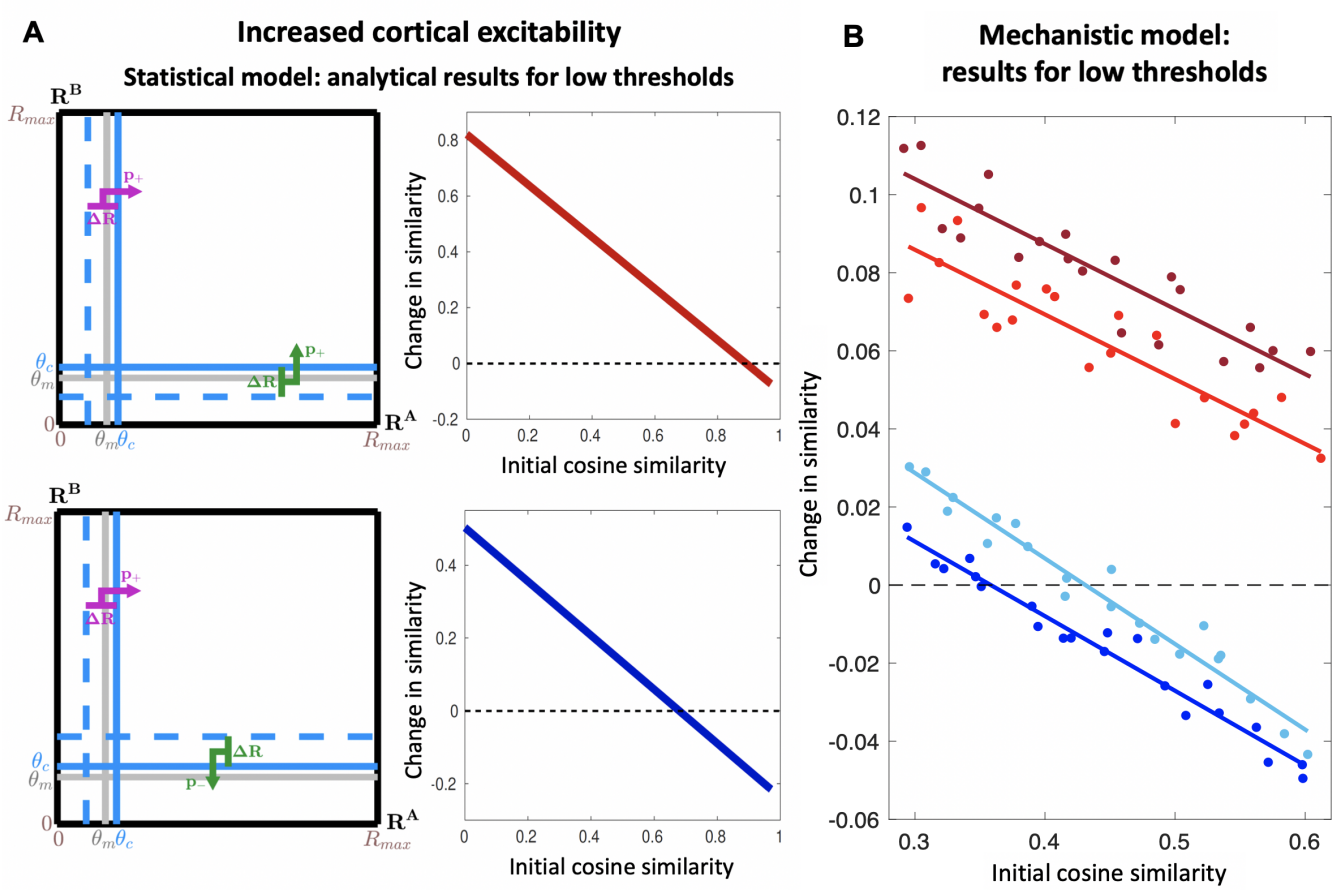

Figure 4. The mechanism for pattern convergence and divergence requires a high-threshold transfer function. Predicted trends in pattern convergence and divergence for increased cortical excitability, obtained from the statistical model and the mechanistic model with low activation thresholds. Low cortical activation thresholds yield effects that are qualitatively different from the realistic high-threshold case. (**A**) Analytical results obtained for $\theta_c = 0.35$ and the feedback scenarios of Fig. 1E,F: red indicates 100% excitatory feedback for both odors; blue indicates 100% excitatory/inhibitory feedback for the first/second odor, respectively. Since θ_c is low, some modules that are not odor-responsive ($R < \theta_m$) are pushed by feedback above θ_c (sketches to left) changing similarity between cortical odor representations (right) differently from Fig. 1E,F: correlated feedback inputs (red) yield pattern convergence decreasing with increasing odor similarity, and for high odor similarity, pattern convergence can turn into pattern divergence. Anticorrelated feedback inputs (blue) yield pattern convergence/divergence for low/high initial odor similarity, respectively. (**B**) Similar results are obtained from the mechanistic model (same simulation conditions as in Fig. 3).

threshold is sufficiently low ($\theta_c \lesssim \theta_m + \Delta R$), feedback can push some modules that were not initially odor responsive to be above the cortical threshold. This can influence the final cortical similarity and change results qualitatively (Fig. 4).

Fig. 4A shows analytical results from our statistical model at a low cortical threshold with identical (top) and opposite (bottom) feedback. We see that some of the effects at high threshold are reversed: (a) when the initial cortical similarity is high, similar contextual feedback to the bulb can lead to cortical pattern divergence instead of pattern convergence ($\Delta\rho$ becomes negative in Fig. 4A, top right), (b) when the initial cortical similarity is low, dissimilar contextual feedback to the bulb can lead to cortical pattern convergence instead of pattern divergence ($\Delta\rho$ becomes positive in Fig. 4A, bottom right), and (c) while pattern divergence for dissimilar contexts increases with increasing initial odor similarity (Fig. 4A, bottom right), pattern convergence for similar contexts decreases with increasing odor similarity (negative slope in Fig. 4A, top right), reversing the trend at high threshold. These reversals at low threshold are confirmed by analysis of more general feedback conditions and response distributions (Fig. S2).

To confirm the predictions in the mechanistic model, we considered a situation where the threshold of the sigmoidal activation function is low, leading to greater cortical excitability, and hence broad activity in the cortical units. In this case we saw the same striking reversal: pattern convergence following strongly correlated feedback decreases with initial odor overlap (Fig. 4B).

Discussion

Context profoundly shapes odor perception [16, 17, 22, 44, 92, 100], and previous studies have demonstrated the critical role of cortical feedback to the OB in the formation of odor-context associations [47, 56]. Cortical feedback can also enhance odor discrimination [2, 21, 52, 57, 68] and generalization [49]. Some studies have implicated feedback in the generation of beta oscillations, thought to be associated with olfactory learning [42, 43, 59, 60, 65]. Consequently, understanding the effect of cortical feedback on the OB is necessary to decode the broader relationship between context and perception in the olfactory system.

Our study contributes to this understanding by showing that diffuse feedback signals, carrying unstructured context representations, can modulate OB responses without synaptic plasticity to effectively entrain convergence and divergence of odor patterns in cortex despite the apparently random afferent projections from the OB. The model predicts that the resulting enhancement in discrimination or generalization should increase linearly with the initial odor similarity. This is especially surprising in the case of odor generalization, since we had expected that more dissimilar odors would show greater convergence in the presence of like contexts, in order to achieve an equal similarity in final cortical representation; in fact, initially similar odors converged more than dissimilar ones, suggesting that the cortex's "first impression" of similarity between a given odor pair plays a strong role in determining the influence of context. Moreover, our results also predict that these linear trends are critically dependent on the strong gating of the OB's projections to cortex. In fact, if the gating were weaker (i.e., if the threshold for cortical activation were lower, leading to less sparse activity), some of these effects would reverse. Our results are robust to broad changes in the statistics of odor inputs and feedback patterns, and are realized in a detailed mechanistic model of the circuit architecture of the olfactory bulb.

Generality of the proposed mechanism

Although our work focused on the olfactory system, similar mechanisms could be implemented in other areas of the brain and support pattern convergence/divergence for other sensory modalities. Indeed, all sensory cortices satisfy the key structural requirements of our statistical model: (1) convergent projections from one layer to the next; (2) increasingly selective gating of the sensory inputs (as demonstrated by a reduction in neural activation levels across the sensory hierarchy [7]); and (3) the presence of feedback carrying information about the context of the sensory stimuli towards the lower sensory layers. We showed that pattern convergence and divergence effects that emerge in neuronal networks with these features rely only on minimally constrained feedback signals that are agnostic to the specific neuronal activation patterns induced by the sensory inputs: that is, feedback is not designed to target specific sensory neurons (e.g. based on their tuning properties), and does not require synaptic plasticity. Effective forms of pattern convergence/divergence arise statistically via a mechanism that simply adjusts the level of similarity in otherwise unstructured feedback patterns to reflect the similarity between the contexts. Feedback signals with these simple properties can be implemented in many ways and in diverse sensory areas of the brain, making this mechanism applicable across modalities.

Role of feedback to the olfactory bulb

Our results demonstrated that feedback to the OB could yield cortical pattern convergence or divergence varying linearly with the initial similarity between the two odor patterns. Because this effect is only dependent on the relationship between feedback patterns induced by different contexts, rather than on the activation of particular MCs or GCs in the bulb, the cortex has much more flexibility in producing pattern convergence and divergence via this mechanism. Moreover, this feedback will automatically induce a predictable level of pattern convergence or divergence, which can presumably be further adjusted by the cortex's native ability to mold odor representations in accordance with the animal's needs [6, 8, 29, 75, 101]. For example, piriform cortex could utilize feedback to induce broad initial changes to odor representation which are then solidified via the PC's inherent recurrent circuitry, which is necessary more generally to stabilize cortical odor representations [9]. Altogether, with these advantages, it is clear that the cortex, by targeting an upstream neural structure, can efficiently induce robust odor discrimination and generalization in its own odor representations, thereby suggesting one etiology for the extensive cortical feedback to the OB.

Model predictions and experimental tests

Key behavioral predictions of our theory follow from the assumption that similarities and differences in the perception of an odor are directly related to the similarities and differences in their cortical representation. Thus, our theory suggests that perceptual learning should be strongly affected by cortical excitability, or, equivalently, the response threshold of cortical neurons. At low excitability (or high threshold), which is the normal state of sensory cortex, odors presented in related contexts should be perceived as more similar than they originally are, and the extent of this perceptual change should increase with the initial similarity. Thus the more similar the odors smell initially, the more similar they should seem after presentation in the same context. By contrast, if excitability is high (or threshold is low), the perceptual change in response to presentation in related contexts should be greatest for odors that were initially dissimilar. Thus, if the cortex of an animal performing a generalization task is highly

excitable, stimuli that are very different should be disproportionately perceived as similar and potentially confused with one another, while perceptual association of initially similar stimuli would be less effective or not achieved at all. By contrast, we predict that, at both low and high thresholds, the ability to discriminate odors associated with different contexts will increase with initial odor similarity.

Some of these predictions are already supported by experimental evidence from studies conducted in rodents. In particular, both [16] and [84] found that association of odor mixtures with different contexts decorrelated the cortical neural responses to the mixtures (pattern divergence) only in difficult discrimination tasks, i.e. when the odor mixtures were perceptually similar. This experimental finding is consistent with our prediction of a positive gain for pattern divergence as a function of the initial odor similarity. Furthermore, both studies reported an increase in cortical response similarity (pattern convergence) when odors were associated with the same context (the same reward port in [16], the same positive valence in [84]). In support of our model predictions, the perceptual effects measured in [16] also strongly suggest a positive gain for pattern convergence.

Our model also makes predictions about the different effects produced by different patterns of feedback (Fig. 2). Experimentally, these predictions could be tested by first engineering similar or different contexts for a pair of odors. Then, after training subjects to associate each odor with its given context, one could measure the statistics of the consequent changes in OB neural activity (e.g. via electrophysiology or calcium imaging), under the hypothesis that these changes reflect the influence of cortical feedback encoding context.

Overall, the predictions following from our proposed mechanism are readily falsifiable. Their rejection or further validation requires measurements of the similarity in cortical responses to odor pairs before and after odor discrimination/association training in two sets of conditions: a control set, in which the response properties of cortical neurons are not altered, and an experimental set, with heightened piriform cortical excitability (i.e. lower activation thresholds). An increase in excitability could be facilitated in a number of ways, for example via stimulation of other brain regions [102], or application of a GABA-A antagonist [28]. Our predictions could also be probed in patients showing neurological conditions of increased cortical excitability such as Alzheimer's disease [96, 98].

Materials and Methods

The statistical model

Parameters and assumptions

We derive the results of the statistical model in terms of the parameters and assumptions in Table 1. Parameters p_{both}, p_{flip}, and the distribution of ΔR together control ρ_{FB}, defined as the cosine similarity in the response changes induced by feedbacks F_A and F_B for odors A and B. Assuming constant feedback strength ΔR, ρ_{FB} (y-axis in Fig. 2) is simply:

$$\rho_{FB} = \frac{\Delta\mathbf{R}_A \cdot \Delta\mathbf{R}_B}{\|\Delta\mathbf{R}_A\|\|\Delta\mathbf{R}_B\|} = p_{\text{both}} - 2\,p_{\text{flip}} \tag{10}$$

where p_{both} is the probability that a module receives feedback for odor B given that it receives feedback for odor A, and p_{flip} is the probability that the feedback for odor B has the opposite effect on the firing rate from the feedback for odor A. The probability $p^B_{\pm}$, that feedback F_B increases/decreases the firing rate of a module, can be related to

$p^A_\pm$, the probability that feedback F_A increases/decreases the firing rate, in terms of p_{both} and p_{flip}:

$$p^B_+ = \frac{p_{\text{same}}\, p^A_+ + p_{\text{flip}}\, p^A_-}{p_{\text{both}}} \qquad p^B_- = \frac{p_{\text{same}}\, p^A_- + p_{\text{flip}}\, p^A_+}{p_{\text{both}}} \tag{11}$$

Similarly we can use Bayes' rule $\Pr(F_A < 0 | F_B > 0)\Pr(F_B > 0) = \Pr(F_B > 0 | F_A < 0)\Pr(F_A < 0)$ written as $p^A_{\text{flip}}\, p^B_+ = p_{\text{flip}}\, p^A_-$ to write the probability that the feedback for A is negative given that the feedback for B is positive:

$$p^A_{\text{flip}} = \frac{p_{\text{flip}}\, p^A_-}{p^B_+} \tag{12}$$

We will use the expressions in (11) and (12) in our analyses.

Parameters	
Thresholds	
θ_m	module response threshold
θ_c	cortical response threshold ($\theta_c > \theta_m$)
$\Delta\theta = \theta_c - \theta_m$	difference in thresholds
Module responses	
R_{max}	maximum module response
$\Delta R_c = R_{max} - \theta_c$	range of responses above θ_c
$\Delta R_m = R_{max} - \theta_m$	range of responses above θ_m
N	total # of modules
$N_\mu \quad (\mu = A, B, AB)$	# of modules responsive (response ¿ θ_m) for odor A, odor B, or both ($\mu = AB$)
$C_\mu \quad (\mu = A, B, AB)$	# of modules activated (response ¿ θ_c) for odor A, odor B, or both ($\mu = AB$)
ΔR	magnitude of change in firing rate for modules affected by feedback
Module densities in response space (derived in terms of the parameters above)	
$\rho(\mu) = N_\mu/\Delta R_m \ (\mu = A, B)$	density of modules (modules/response range) responsive to odor A or B
$\rho(\bar{\mu}) = (N - N_\mu)/\theta_m \ (\mu = A, B)$	density of modules (modules/response range) *not* responsive to odor A or B
$\rho(AB) = N_{AB}/(\Delta R_m)^2$	density of modules in response space (modules/area) responsive to *both* A and B
$\rho(\bar{A}\bar{B}) = (N - N_A - N_B + N_{AB})/\theta_m^2$	density of modules in response space (modules/area) responsive to *neither* A and B
$\rho(A\bar{B}) = (N_A - N_{AB})/(\Delta R_m \times \theta_m)$	density of modules in response space (modules/area) responsive to A but *not* to B
$\rho(\bar{A}B) = (N_B - N_{AB})/(\Delta R_m \times \theta_m)$	density of modules in response space (modules/area) responsive to B but *not* to A
Feedback	
$F_\mu \quad (\mu = A, B)$	feedback associated to odor A or B
$p^\mu_\pm \quad (\mu = A, B)$	fraction of modules that increase or decrease their response due to feedback F_μ
p_{both}	probability a module is affected by feedback for B given it is also affected by feedback for A
p_{flip}	probability the feedback for B affects a module oppositely from the feedback for A
$p_{\text{same}} = p_{\text{both}} - p_{\text{flip}}$	probability the two feedbacks both increase or decrease the firing rate for a module
Assumptions	
response distribution (analytical model)	marginal and joint response distributions of responses to odors A and B are separately uniform below and above θ_m
appendiceal numerical analysis	response distribution and response change after feedback (ΔR) are taken to be Gaussian
feedback coverage	both feedbacks target the same number of modules $p^A_+ + p^A_- = p^B_+ + p^B_-$
feedback independence 1	probability that one feedback increases/decreases a module's firing rate is independent of whether the other feedback affects the module
feedback independence 2	probability that one feedback has no effect on a module given the second feedback does affect it is the same regardless of whether the latter increases or decreases the firing rate

Table 1. Parameters and assumptions of the analytical model.

Analytical calculation of the cortical similarity

The assumption of a uniform response distribution (see Table 1 and Fig. 1C and D in the main paper) allows us to write the number of modules responding over the cortical

threshold θ_c as a proportion of the number of modules that are odor responsive (response higher than θ_m):

$$C_A = \frac{\Delta R_c}{\Delta R_m} N_A \qquad C_B = \frac{\Delta R_c}{\Delta R_m} N_B \qquad C_{AB} = \left(\frac{\Delta R_c}{\Delta R_m}\right)^2 N_{AB} \tag{13}$$

Hence, the initial similarity between the cortical responses to odors A and B is a fraction of the initial similarity in the odor-responsive OB modules (Eq. (5)):

$$\rho_i = \frac{C_{AB}}{\sqrt{C_A C_B}} = Q_0 \frac{N_{AB}}{\sqrt{N_A N_B}} \quad \text{with} \quad Q_0 = \frac{\Delta R_c}{\Delta R_m} \tag{14}$$

We want to derive an analogous expression for the similarity between cortical responses after receiving feedbacks $F_{A,B}$ associated to odors A and B (Fig. 1E). To do so we have to ask how F_A and F_B affect the numbers of modules C_A, C_B and C_{AB} that produce cortical responses to odors A, B, and both.

Feedback-driven changes in modules responding to odor A or to odor B. Let us start with the C_A modules producing cortical responses to A. (The modules responding to B can be treated similarly.) We have to consider two cases: (ii) modules whose responses increase due to feedback, and (ii) modules whose responses decrease due to feedback. C_A is increased by modules whose responses increase if their initial firing rate R is less than, but within ΔR of, the cortical threshold ($\theta_c - \Delta R \leq R \leq \theta_c$). C_A is reduced by modules whose responses decrease if their initial firing rate R is greater than, but within ΔR of, the cortical threshold ($\theta_c + \Delta R \geq R \geq \theta_c$). For simplicity, we have assumed an initially uniform distributions of the N_A module responses above θ_m, and of the $N - N_A$ module responses below θ_m. This means that we can write the change in the number of cortically responsive modules in terms of the proportions of modules within the two ranges indicated above. If the change in the module firing rates after feedback ΔR is smaller than the difference between the cortical activation and module response thresholds $\Delta\theta$, then only the N_A modules with initial firing rate greater than θ_m matter for the analysis. This gives the change in the number of modules responding to A as:

$$\delta C_A = \left[p_+^A \ \min(\Delta R, \Delta\theta) - p_-^A \ \min(\Delta R, \Delta R_c)\right] \ \rho(A) \tag{15}$$

Here $p_\pm^A$ is the probability that a given module increases or decreases its firing rate, $\rho_A = N_A/(R_{\max} - \theta_m)$ is the density of responsive modules (modules/firing rate range), $\min(\Delta R, \Delta\theta)$ is the size of the interval of rates that can contribute to increase in C_A, and $\min(\Delta R, R_{\max} - \theta_c)$ is the size of the interval of rates that can contribute to decrease in C_A. If $\Delta R > \Delta\theta$, some modules that were previously unresponsive (below threshold θ_m) can contribute to the cortical response after feedback, and we have to add this contribution to get:

$$\delta C_A^{\text{total}} = \delta C_A + p_+^A \min(\Delta R - \Delta\theta, \theta_m)\, \rho(\bar{A})\, H(\Delta R - \Delta\theta) \ ; \ H(x) = 0 \text{ if } x < 0 \text{ and equals 1 otherw} \tag{16}$$

where $\rho(\bar{A}) = \frac{(N - N_A)}{\theta_m}$ is the density of modules below the response threshold, and $\min(\Delta R - \Delta\theta, \theta_m) = \min(\theta_m - (\theta_c - \Delta R), \theta_m)$ is the size of the interval below θ_m from which modules could become cortically activated after feedback. We included the step function $H(\Delta R - \Delta\theta)$ to indicate that the second term is only present if the effect of feedback is larger than the difference between thresholds. Through similar reasoning

$$\delta C_B^{\text{total}} = \delta C_B + p_+^B \min(\Delta R - \Delta\theta, \theta_m)\, \rho(\bar{B})\, H(\Delta R - \Delta\theta) \tag{17}$$

is the change in cortical responses to B after feedback.

Feedback-driven changes in modules responding to both odors. A module may increase or decrease its firing rate after feedback following one or both odors, and this feedback may be correlated. To calculate the resulting change in the number of modules reaching the cortical threshold for both odors, δC_{AB}, we must consider different possibilities for the initial firing rates as in the analysis of responses to single odors above. For example, if the response to odor A is just below the cortical threshold θ_c, while odor B lies just above, then positive feedback after A combined with the absence of negative feedback after B triggers a cortical response for both odors. By contrast, if the odor B response lies just below θ_c and the response to A lies just above, then the pattern of feedback must be reversed to get cortical responses after both odors. Reasoning in this manner we can distinguish the 15 cases enumerated in Table 2.

The first four columns in Table 2 list the various relevant ranges for the initial firing rates of a module responding to odors A and B in which increase or decrease of firing by an amount ΔR after feedback can move a module above or below the cortical threshold θ_c. The results for the contribution to δC_{AB} presented in the last column all have the same general form:

$$\begin{aligned}&\Pr\textit{(module increases/decreases response for one odor)}\\ &\times \Pr\textit{(same or opposite for the other odor)}\\ &\times \textit{(Area of the region in the response space)}\\ &\times \textit{(Density of modules in this region).}\end{aligned}$$

The results are written in terms of the probability that both feedbacks affect the module (p_{both}), the probability that the two feedbacks have same sign effects (p_{same}), and the probability that the two feedbacks have opposite sign effects (p_{flip}). See Table 1 for definitions of other parameters.

For example, in Row 1 of Table 2, we start initially with a module whose responses to both odors lie above the response threshold and below the cortical threshold ($\theta_m \leq R \leq \theta_c$). We then multiply the probability that feedback causes the response to both odors to increase ($p_+^A \times p_{\text{same}}$) with the number of modules that lie close enough to the cortical threshold such that feedback will take them over this threshold. This latter is evaluated by multiplying by the density of modules in the joint A-B response space ($\rho(AB) = N_{AB}/(\Delta R_m)^2$) (modules per unit response space area) times the area of the region that lies close enough to the cortical threshold to cross over after feedback. Since the feedback is taken to to have magnitude ΔR, the latter area is the square of the minimum of ΔR and the interval between the thresholds $\Delta\theta = \theta_c - \theta_m$.

In Row 2 of Table 2 we consider a module which initially responds below the cortical threshold to odor A $\theta_m < R < \theta_c$, but above the threshold for odor B. Thus, before feedback, this module does not contribute to C_{AB}, the count of modules responding to both odors. After feedback the module will contribute to this count if the response to odor A increases, and the response to odor B does not decrease. The probability of this happening is $p_+^A\,(1 - p_{\text{flip}})$, accounting for the first two factors in the result. To count the number of modules that experience this situation, as with the Row 1 result, we have to multiply by the density of modules in response space by the area in response space which lies close enough to the cortical threshold. This accounts for the last two factors in the result.

Proceeding similarly, we arrive at the results in each row of Table 2. Note that the results in some rows (e.g. rows 6, 7, and 8) are negative, and contain factors of p_-^A or p_-^B because they represent the effects of negative feedback that cause modules which are initially above the cortical threshold for both odors to cease to be so.

$\delta C_{AB}^{\text{total}} = \sum_{i=1}^{15} \delta C_{AB}^{(i)}$					
(i)	$(\theta_c - \Delta R, \theta_m)$	(θ_m, θ_c)	$(\theta_c, \theta_c + \Delta R)$	$(\theta_c + \Delta R, R_{\max})$	Contribution $\delta C^{(i)}$ to δC_{AB}
1		A,B			$p_+^A\, p_{\text{same}}\, [\min(\Delta R, \Delta\theta)]^2\ \rho(AB)$
2		A	B		$p_+^A\ (1 - p_{\text{flip}})\ [\min(\Delta R, \Delta\theta) \min(\Delta R, \Delta R_c)]\ \rho(AB)$
3		B	A		$p_+^B\ \left(1 - p_{\text{flip}}^A\right)\ [\min(\Delta R, \Delta\theta) \min(\Delta R, \Delta R_c)]\ \rho(AB)$
4		A		B	$p_+^A\ [\min(\Delta R, \Delta\theta) \max(0, \Delta R_c - \Delta R)]\ \rho(AB)$
5		B		A	$p_+^B\ [\min(\Delta R, \Delta\theta) \max(0, \Delta R_c - \Delta R)]\ \rho(AB)$
6			A,B		$-\left(p_-^A + p_-^B - p_-^A\, p_{\text{same}}\right)\ [\min(\Delta R, \Delta R_c)]^2\ \rho(AB)$
7			A	B	$-p_-^A\ [\min(\Delta R, \Delta R_c) \max(0, \Delta R_c - \Delta R)]\ \rho(AB)$
8			B	A	$-p_-^B\ [\min(\Delta R, \Delta R_c) \max(0, \Delta R_c - \Delta R)]\ \rho(AB)$
9	A,B				$p_+^A\, p_{\text{same}}\ [\min(\Delta R - \Delta\theta, \theta_m)]^2\ \rho(\bar{A}\bar{B}) H(\Delta R - \Delta\theta)$
10	A	B			$p_+^A\, p_{\text{same}}\ [\min(\Delta R - \Delta\theta, \theta_m) \times \Delta\theta]\ \rho(\bar{A}B) H(\Delta R - \Delta\theta)$
11	B	A			$p_+^A\, p_{\text{same}}\ [\min(\Delta R - \Delta\theta, \theta_m) \times \Delta\theta]\ \rho(A\bar{B}) H(\Delta R - \Delta\theta)$
12	A		B		$p_+^A\, (1 - p_{\text{flip}})\, [\min(\Delta R - \Delta\theta, \theta_m) \min(\Delta R, \Delta R_c)]\, \rho(\bar{A}B) H(\Delta R - \Delta\theta)$
13	B		A		$p_+^B (1 - p_{\text{flip}}^A)\, [\min(\Delta R - \Delta\theta, \theta_m) \min(\Delta R, \Delta R_c)]\, \rho(A\bar{B}) H(\Delta R - \Delta\theta)$
14	A			B	$p_+^A\ [\min(\Delta R - \Delta\theta, \theta_m) \max(0, \Delta R_c - \Delta R)]\ \rho(\bar{A}B) H(\Delta R - \Delta\theta)$
15	B			A	$p_+^B\ [\min(\Delta R - \Delta\theta, \theta_m) \max(0, \Delta R_c - \Delta R)]\ \rho(A\bar{B}) H(\Delta R - \Delta\theta)$

Table 2. All fifteen cases for feedback-driven changes in modules responding to both odors. The first four columns are indexed by ranges for the initial response rates of the modules before feedback. For example, (θ_m, θ_c) indicates that the initial response lies between the response threshold θ_m and the cortical threshold θ_c. Likewise $(\theta_c - \Delta R, \theta_m)$ indicates that the initial response lies below the response threshold θ_m but is sufficiently high that a feedback-driven addition of ΔR to the response rate would push the module above the cortical threshold. This case is only possible if the cortical threshold is sufficiently low and the strength of the feedback is sufficiently high. The letters in each row indicate the range in which the response lies for each of the two odors. The results in rows 9-15 include the step function $H(x)$ which is defined as $H(x) = 0$ if $x < 0$ and $H(x) = 1$ if $x \geq 0$.

Overall equation

As described in the main paper, the cortical similarity is given by $\rho = C_{AB}/\sqrt{C_A C_B}$, which is the ratio of the number of modules that respond to both odors to the geometric mean of the numbers of modules that respond to each of A and B. Thus the similarity after feedback is:

$$\rho_f = \frac{C_{AB} + \delta C_{AB}^{\text{total}}}{\sqrt{\left(C_A + \delta C_A^{\text{total}}\right)\left(C_B + \delta C_B^{\text{total}}\right)}} = Q_3 \frac{N_{AB}}{\sqrt{N_A N_B}} + Q_1 \tag{18}$$

where in the second equation we are separating out all terms proportional to N_{AB}, having noting that, from (6), C_{AB} is proportional to N_{AB}, and that, from the expressions for $\delta C_{AB}^{\text{total}}$ in Table 2 and for the densities $\rho(AB)$, $\rho(\bar{A}B)$, $\rho(A\bar{B})$, and $\rho(\bar{A}B)$ in Table 1, every term in $\delta C_{AB}^{\text{total}}$ also contains one piece proportional to N_{AB}. Upon including the explicit expressions for $\delta C_A^{\text{total}}$, $\delta C_B^{\text{total}}$, and $\delta C_{AB}^{\text{total}}$ all factors of N_A, N_B and N_{AB} except those explicitly displayed in (18) cancel out, so that Q_3 and Q_1 only depend on the remaining model parameters such as the thresholds and the probabilities of positive and negative feedback. Recalling from (14) that the initial cortical similarity before feedback is $\rho_i = Q_0 N_{AB}/\sqrt{N_A N_B}$, we can express the change

in similarity as

$$\Delta\rho = Q_2\,\rho_i + Q_1 \quad \text{with} \quad Q_2 = \frac{Q_3}{Q_0} - 1 \tag{19}$$

This is the main result of our analysis.

Pattern convergence and divergence: analytical model

Figs. 1 and 2 of the main paper illustrate these general results for a wide variety of feedback parameters. The general finding is that if the cortical activation threshold is sufficiently high, then $Q_1 = 0$ and $Q_2 > 0$ when the feedback similarity is high, while $Q_2 < 0$ when the feedback similarity is low. Thus high and low feedback similarity give rise to, respectively, pattern convergence and divergence increasing with initial odor similarity. On the contrary, if the cortical threshold is low, so that $\theta_c < \theta_m + \Delta R$, then $Q_1 \neq 0$ and $Q_2 < 0$ for both low and high feedback similarities (partial reversal of effects, Fig. 4 and green regions in Fig. 2). Here, we discuss in more detail the intercept Q_1 and the sign of the slope Q_2 in high and low cortical-threshold regimes.

Intercept Q_1 and slope Q_2 in the high-threshold regime The coefficient $Q2$ in (19) arises from the terms in $\delta C_{AB}^{\text{total}}$ that are proportional to N_{AB}, the number of modules that are responsive to both odors. Every contribution $\delta C^{(i)}$ to $\delta C_{AB}^{\text{total}}$ that is listed in Table 2 contains such a term via the expressions in Table 1 for the densities of modules that are responsive/unresponsive to the two odors, $\rho(AB)$, $\rho(\bar{A}B)$, $\rho(A\bar{B})$, and $\rho(\bar{A}\bar{B})$. By contrast, the intercept Q_1 comes only from the contributions 9-15 in Table 2 that arise from situations where modules are are non-responsive before feedback for at least one of the two odors ($R < \theta_m$). These contributions only matter if the effect of feedback, ΔR, is larger than the difference between the cortical and response thresholds, $\Delta\theta$ (hence the step function $H(\Delta R - \Delta\theta)$ in each of these contributions). Thus all of these contributions vanish if the cortical activation threshold is sufficiently high, i.e. $\theta_c > \theta_m + \Delta R$, so that the $Q_1 = 0$ in these cases (Figs. 1 and 2 of the main text). In the main text, we have also provided explanations for the sign of Q_2 in the high cortical-threshold regime, in the simple situations of identical ($Q_2 > 0$) and opposite ($Q_2 < 0$) feedback. The only exception to these situations is the case of $\sim 100\%$ inhibitory feedback for both odors (i.e. $p_-/(p_+ + p_-) \sim 1$). In this singular case, identical feedback yields pattern divergence increasing with odor similarity ($Q_2 < 0$). As a simple illustration, consider the special case that the inhibitory feedback targets all the modules ($p_-^{F1} = p_-^{F2} = 1$): the effect of such feedback on the cortical similarity is equivalent to an increase in the cortical activation threshold for both odors, which corresponds to flipping the sign of ΔR in Eq. (8) of the main text. Thus $Q_2 = -\Delta R/\left(R_{\max} - \theta_c\right) < 0$. As another special case, the slope is also negative when feedback does not target all the modules and affects disjoint subsets for the two odors.

Intercept Q_1 and slope Q_2 in the low-threshold regime Now consider a situation where cortex has increased excitability and thus has a low threshold relative to the strength of feedback: $\theta_c < \theta_m + \Delta R$. Then, from Eq. (19) we can derive exact expressions for Q_1 and Q_2 in the special case $p_+^A = 1$, $p_{\text{both}} = 1$, $p_{\text{flip}} = 0$, i.e., when feedback targets all the modules and is positive for both odors (main text Fig. 4A, top). In this case, $p_+^B = 1$ as well and $p_-^A = p_-^B = 0$. Because of this, only the p_+^A and p_+^B terms in Eqs. (15), (16), and (17) contribute to the shifts in the number of cortically responsive modules $\delta C_A^{\text{total}}$ and $\delta C_B^{\text{total}}$ that appear in the denominator of the cortical similarity after feedback (18).

Every contribution to $\delta C_{AB}^{\text{total}} = \sum_{i=1}^{15} \delta C^{(i)}$ in the numerator of the similarity equation (18) contains a term that is *proportional* to N_{AB} because there is such a term

in all module densities ρ listed in Table 1; we will call these terms $\delta C^{(i,p)}$. Meanwhile the contributions $\delta C^{(i)}$ with $i = 9, \cdots 15$ also contain terms that are *not proportional* to N_{AB}; we will call these $\delta C^{(i,n)}$. In terms of these quantities, after some algebra we can write the intercept Q_1 and slope Q_2 in (18) as:

$$Q_1 = \frac{\sum_9^{15} \delta C^{(i,n)}}{\sqrt{(C_A + \delta C_A^{\text{total}})(C_B + \delta C_B^{\text{total}})}}$$

$$Q_2 = \frac{\sqrt{N_A N_B}}{Q_0 N_{AB}} \frac{\left(C_{AB} + \sum_1^{15} \delta C^{(i,p)}\right)}{\sqrt{(C_A + \delta C_A^{\text{total}})(C_B + \delta C_B^{\text{total}})}} - 1$$

Notice that the intercept Q_1 only takes contributions from modules that do not respond to at least one odor before feedback (cases $i = 9 \cdots 15$ in Table 2), and thus requires a low cortical threshold with $\Delta\theta < \Delta R$. Explicit computation shows that $Q_1 > 0$. Meanwhile, in the expression for Q_2 there are three qualitatively different sets of terms:

- The part of the sum for Q_2 from $i = 1$ to $i = 8$ represents modules that initially respond to both odors (see Table 2), and these give positive contributions of Q_2, just like they do in the high cortical threshold regime.
- The term in Q_2 with $i = 9$ comes from modules that do not respond to either odor before feedback (see Table 2), and therefore only makes a contribution to Q_2 in the low threshold regime where the difference in thresholds is smaller than the effect of feedback, i.e., $\Delta\theta < \Delta R$. This term is positive because the number of modules in this set increases as odor similarity increases.
- The part of the sum in Q_2 from $i = 10$ to $i = 15$ arises from modules that respond to only one of the two odors before feedback (see Table 2). Therefore it also only makes a contribution to Q_2 when the cortical threshold is low, i.e., $\Delta\theta < \Delta R$. This contribution is negative because the number of modules in this set increases as the odor similarity *decreases*.

Overall, the third contribution dominates so that the slope is negative ($Q_2 < 0$) when $\theta_c \leq \max(\theta_m, \Delta R)$, regardless of the other parameter values. This is a sufficient condition to have $Q_2 < 0$, but for most values of the model parameters, higher θ_c can also yield negative slopes, as long as $\theta_c < \theta_m + \Delta R$.

A similar argument can be applied to the case $p_+^A = 1$, $p_{\text{both}} = 1$, $p_{\text{flip}} = 1$, i.e. when feedback targets all the modules and is positive for odor A, and negative for odor B (Fig. 4A, bottom). In this case the analysis shows that in this feedback condition $Q_2 < 0$ for any $\theta_c < \theta_m + \Delta R$ and $Q_1 > 0$ for any $\theta_c < \min(\theta_m + \Delta R, R_{\max} - \Delta R)$.

The mechanistic model

Connectivity

Each cell was modeled by its distribution of lateral dendrites (for the MCs) or its dendritic spines (for the GCs), since the interface between these two cell types occurs at the junction of these neuronal processes. To determine the average number of synapses between a specific GC and MC then, we wished to calculate the average number of spines from that GC that would be present at a sufficiently close space to the MC lateral dendrites, under the assumption that spines which are sufficiently close to the dendrites will form synapses. Thus, in geometric terms, we wished to calculate, for a given MC-GC pair, 1) the amount of lateral dendrite present in the overlap between the

MC lateral dendritic and GC spine spatial distributions and then 2) the number of spines sufficiently close to those lateral dendrites in the overlap.

To treat this mathematically, we first developed mean-field distributions of MC lateral dendritic length density (in μm of lateral dendrite per unit area) and GC spine density (number of spines per unit volume) based on the experimental literature. Since the MC lateral dendritic distribution was defined on a disk and assumed to be radially symmetric, it was simply dependent on the distance away from the center of the cell, with the specific shape of this distribution determined from data from [14, 63, 67]. Meanwhile, the GC spine distribution was defined on an inverted cone. The spine density at a given height was expressed as the number of spines per unit height N_s divided by the area of the circular cross-section of the cone at that height. We assumed no radial dependency, since the dendrites carrying the spines pass roughly perpendicular to the MC dendrites, but N_s was height-dependent: this height dependence was modeled based on results from [63] and [33]. Thus, the MC and GC distributions could be expressed as $\rho_m(r)$ and $\rho_g(z)$, respectively, where r is the distance away from the center of the MC and z is the height along the GC cone (main paper, Fig. 3A).

With these distributions, we could then calculate the number of expected synapses between a given MC-GC pair. First, we determined how much length of lateral dendrite (in μm) existed in the overlap between the dendritic distributions of the two cells via the integral:

$$l = \int_{A_{\text{overlap}}} \rho_m(r) dA \tag{20}$$

To calculate the number of spines apposed to these dendrites (and therefore synapses), we needed to account for the 3-dimensional nature of potential synapses between MC and GC by determining the total volume of MC-GC interaction around these dendrites. To do this, we first converted the total dendritic length in the overlap to an equivalent volume of lateral dendrite by assuming the lateral dendrites to be cylinders with diameter 1 μm. Then the volume is simply the area of the cross-section multiplied by the total length in the overlap calculated in 20, $\pi r^2_{\text{dendrite}} \times l$. We then defined the volume of interaction as a cylindrical shell surrounding the lateral dendrite with a thickness d_{shell} equal to 1 μm, about the effective diameter of a spine [99] (Fig. S3A). To calculate this volume, we simply subtracted the lateral dendrite volume from the combined volume of the shell and lateral dendrite as follows:

$$V = \pi((d_{\text{shell}} + r_{\text{dendrite}})^2 - r^2_{\text{dendrite}})l = q\pi l \tag{21}$$

Where $q = 2.32\ \mu m^2$ [14]. The density of GC spines in this volume was determined by the height where the MC disk intersects the GC cone, under the simplifying assumption that this density is roughly constant across the height of the volume of interest. Finally, we can multiply this spine density by the previous volume V to find the average number of spines and thus synapses present between MC and GC:

$$\lambda = \rho_g(z_{\text{intersect}})V = \rho_g(z_{\text{intersect}}) \left(q\pi \int_{A_{\text{overlap}}} \rho_m(r) dA \right) \tag{22}$$

Since a given MC-GC pair generally has only one synapse [99], we used a Poisson distribution to calculate the probability that at least one synapse was formed, with the assumption that if more than one synapse was formed, it would still be counted as a single synapse. Thus:

$$P(\text{synapse}) = P(N_{\text{synapses}} \neq 0) = 1 - \exp(-\lambda) \tag{23}$$

From this algorithm, we were able to generate a network of 3,550 MCs and 53,250 GCs for simulation.

Neuronal and network dynamics

The internal dynamics of individual neurons were modeled via the Izhikevich equations [40, 41]:

$$C\frac{dv}{dt} = k(v - v_r)(v - v_t) - u + I \quad ; \quad \frac{du}{dt} = a(b(v - v_r) - u)\,,$$

with spike reset:

$$\text{If } v \geq v_c, \text{then } \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} ,$$

where v is the voltage, u is a recovery current, v_r is the resting potential, v_t is the threshold potential, v_c is the cutoff voltage, and I is external current. a, b, c, d, and k are parameters to be chosen.

For mitral cells, the parameters were selected to account for their class II behavior [31] and to establish a realistic $f - I$ curve [14]. Granule cells were assumed to be class I neurons; in that case, in the Izhikevich model, b and k could be calculated as follows for $b < 0$ [41]:

$$b = \frac{v_r - v_t + 4R\rho}{4R^2\rho} \quad ; \quad k = \frac{1}{4R^2\rho}\,,$$

where R is input resistance and ρ is the rheobase. The remaining parameters were chosen to match a realistic $f - I$ curve [33]. The mean parameter values are given below:

Parameters (mean values)		
Variable	Mitral cell	Granule cell
k (nS(mV)$^{-1}$)	2.5	0.067
a (ms^{-1})	0.02	0.01
b (nS)	12	-0.133
c (mV)	-70	-75
d (pA)	13	2
v_r (mV)	-58	-71
v_t (mV)	-49	-39
v_c (mV)	30	25
C (pF)	191	48

Parameter values were drawn from a normal distribution with standard deviation equal to 1/10 of the given mean, except for b and k for granule cells, which were drawn from normal distributions with standard deviations equal to 2/3 of the mean and constrained such that 1) $b < 0$, 2) the resulting rheobase was between 10 and 70 pA, and 3) the input resistance was between 0.25 and 1.5 GΩ.

The dendrodendritic synapses were modeled as NMDA and AMPA receptors on GCs and GABA receptors on MCs. The synaptic current was modeled as the following for AMPA receptors:

$$I_{\text{AMPA}}(t) = s(t)g_{\text{AMPA}}(V(t) - E)\,,$$

where, for a particular receptor, g is the conductance, s is a gating variable representing the fraction of open channels at the synapse, $V(t)$ is the voltage of the recipient cell, and $E_e = 0$ mV is the reversal potential.

For GABA receptors, the current was modeled as:

$$I_{\text{GABA}}(t) = s(t)g_{\text{GABA}}\exp(-\frac{L}{\lambda})(V(t) - E_i)\,,$$

where $E_i = -70$ mV is the inhibitory reversal potential, L is the distance between the center of the MC and the synapse location, and λ is a length constant. Since inhibitory signals from the cell periphery degrade as they propagate to the soma [62], we modeled this degradation as a simple exponential decay.

Additionally, for NMDA receptors [19]:

$$I_{\mathrm{NMDA}}(t) = s(t)\frac{g(V(t)-E)}{1+\frac{[\mathrm{Mg}^{2+}]\exp(-0.062V(t))}{3.57}}\,,$$

where the additional term describes the magnesium block, with $[\mathrm{Mg}^{2+}]$ assumed to be 1 mM.

The time evolution of the gating variables for each receptor was modeled following [12]:

$$\frac{ds_{\mathrm{AMPA}}}{dt} = \frac{-s_{\mathrm{AMPA}}}{\tau_{\mathrm{AMPA}}} \quad ; \quad \frac{ds_{\mathrm{GABA}}}{dt} = \frac{-s_{\mathrm{GABA}}}{\tau_{\mathrm{GABA}}}$$

$$\frac{ds_{\mathrm{NMDA}}}{dt} = \frac{-s_{\mathrm{NMDA}}}{\tau_{NMDA_{decay}}} + \alpha n(t)(1-s_{\mathrm{NMDA}}(t)) \quad ; \quad \frac{dn}{dt} = \frac{-n}{\tau_{NMDA_{rise}}}$$

When a MC spikes, the gating variables of NMDA and AMPA receptors at its synapses (which provide NMDA and AMPA current to the connected GCs) are updated as follows:

$$s_{\mathrm{AMPA}} \leftarrow s_{\mathrm{AMPA}} + W(1-s_{\mathrm{AMPA}}) \quad ; \quad n \leftarrow n + W(1-n)\,,$$

where $W = 0.5$.

Additionally, the gating variables of all MCs indirectly connected to the spiking MC via GCs have the gating variables of their GABA receptors at their synapses with those GCs updated as follows to account for spike-independent GC spiking [23]:

$$s_{\mathrm{GABA}} \leftarrow s_{\mathrm{GABA}} + \kappa W(1-s_{\mathrm{GABA}})\,,$$

where $0 < \kappa < 1$.

When a GC spikes, the gating variables of GABA receptors at its synapses (providing GABA current to all connected MC) are updated as follows:

$$s_{\mathrm{GABA}} \leftarrow s_{\mathrm{GABA}} + W(1-s_{\mathrm{GABA}})$$

The values of the time constants τ and α were taken or derived from [82], while the synaptic conductances and κ were tuned in a specialized network, whose dimensions were meant to mimic (within computational constraints), the length and width of a brain slice, in order to reproduce as faithfully as possible lateral inhibition results from [3]. The length constant λ was calculated from the formula in [62] for a 1.26 μm diameter dendrite. The mean parameter values are as follows:

Synaptic parameters (mean values)	
Variable	Value
g_{AMPA} (nS)	0.73
g_{NMDA} (nS)	0.84
g_{GABA} (nS)	0.13
κ	0.006
τ_{AMPA} (ms)	5.5
$\tau_{\mathrm{NMDA\,rise}}$ (ms)	10
$\tau_{\mathrm{NMDA\,decay}}$ (ms)	80
τ_{GABA} (ms)	18
λ (μm)	675
α (ms^{-1})	0.1

Parameter values were drawn from a normal distribution with standard deviation equal to 1/10 of the given mean.

Firing rate distributions

Firing rate distributions were acquired by simulating a network of 3,550 mitral cells and 53,250 granule cells for one sniff cycle ($\frac{1}{6}$ of a second assuming a 6 Hz sniffing rate) for a set of 8 different odors each targeting 20 of 178 glomeruli, with the strength of the odor being drawn from a uniform distribution between 300 and 500 pA. The resulting firing rates were fit with a generalized Pareto distribution with parameters $k = -0.281$, $\sigma = 3.331$, and $\theta = -0.4$ (Fig. S3B). Excitatory feedback current was randomly added to a different subset of the MCs for each odor (in each case, feedback targeted $\frac{1}{5}$ of all MCs with current drawn from between 0 and 300 pA), and the simulations were rerun. The resulting changes in firing rates due to feedback were fit by a zero-elevated skew normal distribution, with the probability of being drawn from the distribution equal to 0.31 and the consequent skew normal distribution having the parameters $\alpha = 4.232$, $\omega = 2.687$, and $\xi = 0.7785$ (Fig. S3C). For excitatory feedback to GCs, a simulation was done for a set of 8 odors with the a different feedback pattern for each odor (targeting $\frac{1}{8}$ of all GCs with current drawn from between 0 and 100 pA), and the change in firing among the odor-receiving cells was then determined. Note that the distribution was determined among MCs with spike count greater than or equal 2, as the maximum number of spikes lost due to GC feedback was found to be 2. The resulting changes in firing rate were fit with a lognormal distribution with parameters $\mu = 0.7395$, $\sigma = 0.2515$, which was then appropriately shifted and flipped to produce values with the desired range and shape (Fig. S3D).

Acknowledgments

We thank Jay A. Gottfried for fruitful discussions. DK would also like to thank Minghong Ma and Graeme Lowe for their advice on olfactory bulb anatomy, Eugenio Piasini for use of his server, as well as Max Kelz, Joseph Aicher, Rebecca Li, and Matthew Kersen for useful discussions and feedback. VB and DK were supported by the Simons Foundation through MMLS grant 400425. GT was supported by the Swartz Foundation.

References

1. N. M. Abraham, R. Vincis, S. Lagier, I. Rodriguez, and A. Carleton. Long term functional plasticity of sensory inputs mediated by olfactory learning. *eLife*, 3:e02109, 2014.

2. W. Adams, J. N. Graham, X. Han, and H. Riecke. Top-down inputs drive neuronal network rewiring and context-enhanced sensory processing in olfaction. *PLoS Computational Biology*, 15:1–27, 2019.

3. A. C. Arevian, V. Kapoor, and N. N. Urban. Activity–dependent gating of lateral inhibition in the mouse olfactory bulb. *Nature Neuroscience*, 11(1):80–87, 2008.

4. R. Balu, R. T. Pressler, and B. W. Strowbridge. Multiple modes of synaptic excitation of olfactory bulb granule cells. *The Journal of Neuroscience*, 27(21):5621–5632, 2007.

5. L. F. Barrett, B. Mesquita, and M. Gendron. Context in emotion perception. *Current Directions in Psychological Science*, 20:286–290, 2011.

6. J. M. Bekkers and N. Suzuki. Neurons and circuits for odor processing in the piriform cortex. *Trends in Neuroscience*, 36(7):429–438, 2013.

7. M. Beyeler, E. L. Rounds, K. D. Carlson, N. Dutt, and J. L. Krichmar. Neural correlates of sparse coding and dimensionality reduction. *PLoS Computational Biology*, 15(6):e1006908, 2019.

8. K. A. Bolding and K. M. Franks. Complementary codes for odor identity and intensity in olfactory cortex. *eLife*, 6:1–26, 2017.

9. K. A. Bolding, S. Nagappan, B. Han, F. Wang, and K. M. Franks. Recurrent circuitry is required to stabilize piriform cortex odor representations across brain states. *eLife*, 9:1–23, 2020.

10. A. M. Boyd, H. K. Kato, T. Komiyama, and J. S. Isaacson. Broadcasting of cortical activity to the olfactory bulb. *Cell Reports*, 10(7):1032–1039, 2015.

11. A. M. Boyd, J. F. Sturgill, C. Poo, and J. S. Isaacson. Cortical feedback control of olfactory bulb circuits. *Neuron*, 76(6):1161–1174, 2012.

12. N. Brunel and X. J. Wang. Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience*, 11(1):63–85, 2001.

13. S. D. Burton. Inhibitory circuits of the mammalian main olfactory bulb. *Journal of Neurophysiology*, 118(4):2034–2051, 2017.

14. S. D. Burton and N. N. Urban. Greater excitability and firing irregularity of tufted cells underlies distinct afferent–evoked activity of olfactory bulb mitral and tufted cells. *The Journal of Physiology*, 592(10):2097–2118, 2014.

15. C. Chambers, S. Akram, V. Adam, C. Pelofi, M. Sahani, S. Shamma, and D. Pressnitzer. Prior context in audition informs binding and shapes simple features. *Nature Communications*, 8:1–11, 2017.

16. J. Chapuis and D. A. Wilson. Bidirectional plasticity of cortical pattern recognition and behavioral sensory acuity. *Nature Neuroscience*, 15(1):155, 2012.

17. T. Cleland, A. Morse, E. L. Yue, and C. Linster. Behavioral models of odor similarity. *Behavioral Neuroscience*, 116:222–231, 2002.

18. I. G. Davison and M. D. Ehlers. Neural circuit mechanisms for pattern detection and feature combination in olfactory cortex. *Neuron*, 70(1):82–94, 2011.

19. A. Destexhe, Z. F. Mainen, and T. J. Sejnowski. *Methods in Neuronal Modeling*. MIT Press, 1998.

20. W. Doucette, D. H. Gire, J. Whitesell, V. Carmean, M. T. Lucero, and D. Restrepo. Associative cortex features in the first olfactory brain relay station. *Neuron*, 69(6):1176–1187, 2011.

21. W. Doucette, J. Milder, and D. Restrepo. Adrenergic modulation of olfactory bulb circuitry affects odor discrimination. *Learning & Memory*, 14(8):539–547, 2007.

22. W. Doucette and D. Restrepo. Profound context-dependent plasticity of mitral cell responses in olfactory bulb. *PLoS Biology*, 6(10):e258, 2008.

23. V. Egger, K. Svoboda, and Z. F. Mainen. Dendrodendritic synaptic signals in olfactory bulb granule cells: Local spine boost and global low-threshold spike. *The Journal of Neuroscience*, 25(14):3521–3530, 2005.

24. V. Egger and N. N. Urban. Dynamic connectivity in the mitral cell–granule cell microcircuit. *Seminars in Cell & Developmental Biology*, 17(4):424–432, 2006.

25. D. Ellingsen, S. Leknes, G. Loseth, J. Wessberg, and H. Olausson. The neurobiology shaping affective touch: Expectation, motivation, and meaning in the multisensory context. *Frontiers in Psychology*, 6:1–16, 2016.

26. M. Ennis, A. C. Puche, T. Holy, and M. T. Shipley. The olfactory system. In G. Paxinos, editor, *The Rat Nervous System*, chapter 27, pages 761–803. Elsevier, Inc., 2015.

27. M. L. Fletcher and W. R. Chen. Neural correlates of olfactory learning: critical role of centrifugal neuromodulation. *Learning & Memory*, 17(11):561–570, 2010.

28. M. Forti, G. Biella, S. Caccia, and M. de Curtis. Persistent excitability changes in the piriform cortex of the isolated guinea-pig brain after transient exposure to bicuculline. *European Journal of Neuroscience*, 9(3):435–451, 1997.

29. K. M. Franks, M. J. Russo, D. L. Sosulski, A. A. Mulligan, S. A. Siegelbaum, and R. Axel. Recurrent circuitry dynamically shapes the activation of piriform cortex. *Neuron*, 72:49–56, 2011.

30. R. A. Fuentes, M. I. Aguilar, M. L. Aylwin, and P. E. Maldonado. Neuronal activity of mitral-tufted cells in awake rats during passive and active odorant stimulation. *Journal of Neurophysiology*, 100(1):422–430, 2008.

31. R. F. Galán, N. Fourcaud-Trocmé, G. B. Ermentrout, and N. U. Urban. Correlation-induced synchronization of oscillations in olfactory bulb neurons. *The Journal of Neuroscience*, 26(14):3646–3655, 2005.

32. Y. Gao and B. W. Strowbridge. Long-term plasticity of excitatory inputs to granule cells in the rat olfactory bulb. *Nature Neuroscience*, 12:731–733, 2009.

33. M. A. Geramita, S. D. Burton, and N. N. Urban. Distinct lateral inhibitory circuits drive parallel processing of sensory information in the mammalian olfactory bulb. *eLife*, 5(e16039):1–22, 2016.

34. S. Ghosh, S. D. Larson, H. Hefzi, Z. Marnoy, T. Cutforth, K. Dokka, and K. K. Baldwin. Sensory maps in the olfactory cortex defined by long-range viral tracing of single neurons. *Nature*, 472(7342):217–220, 2011.

35. A. R. Girshick, M. S. Landy, and E. Simoncelli. Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, 14:926–932, 2011.

36. C. M. Gray and J. E. Skinner. Centrifugal regulation of neuronal activity in the olfactory bulb of the waking rabbit as revealed by reversible cryogenic blockade. *Experimental Brain Research*, 69:378–386, 1988.

37. A. Hayar, P. M. Heyward, T. Heinbockel, M. T. Shipley, and M. Ennis. Direct excitation of mitral cells via activation of α1-noradrenergic receptors in rat olfactory bulb slices. *Journal of Neurophysiology*, 86(5):2173–2182, 2001.

38. T. E. Holy. Olfactory bulb connectomics: a silver lining. *Nature Neuroscience*, 19(6):767–768, 2016.

39. I. Ito, R. C. Ong, B. Raman, and M. Stopfer. Sparse odor representation and olfactory learning. *Nature Neuroscience*, 11:1177–1184, 2008.

40. E. M. Izhikevich. Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14(6):1569–1572, 2003.

41. E. M. Izhikevich. *Dynamical Systems in Neuroscience*. MIT Press, 2005.

42. L. M. Kay. Circuit oscillations in odor perception and memory. *Progress in Brain Research*, 208:223–251, 2014.

43. L. M. Kay and W. J. Freeman. Bidirectional processing in the olfactory-limbic axis during olfactory behavior. *Behavioral Neuroscience*, 112:541–553, 1998.

44. L. M. Kay and G. Laurent. Odor-and context-dependent modulation of mitral cell activity in behaving rats. *Nature Neuroscience*, 2(11):1003, 1999.

45. M. T. Ke, S. Fujimoto, and T. Imai. SeeDB: a simple and morphology–preserving optical clearing agent for neuronal circuit reconstruction. *Nature Neuroscience*, 16(8):1154–1161, 2013.

46. D. E. C. Kersen, G. Tavoni, and V. Balasubramanian. The influence of connectivity on function in a large-scale model of the olfactory bulb. *In preparation*, 2021.

47. C. L. Kiselycznyk, S. Zhang, and C. Linster. Role of centrifugal projections to the olfactory bulb in olfactory processing. *Learning & Memory*, 13(5):575–579, 2006.

48. K. Kveraga, A. S. Ghuman, and M. Bar. Top-down predictions in the cognitive brain. *Brain and Cognition*, 65:145–168, 2007.

49. M. Levinson, J. P. Kolenda, G. J. Alexandrou, O. Escanilla, T. A. Cleland, D. M. Smith, and C. Linster. Context-dependent odor learning requires the anterior olfactory nucleus. *Behavioral Neuroscience*, 134:332–343, 2020.

50. G. Li and T. A. Cleland. A two-layer biophysical model of cholinergic neuromodulation in olfactory bulb. *The Journal of Neuroscience*, 33(7):3037–3058, 2013.

51. C. Linster and T. A. Cleland. Cholinergic modulation of sensory representations in the olfactory bulb. *Neural Networks*, 15:709–717, 2002.

52. C. Linster and M. Hasselmo. Modulation of inhibition in a model of olfactory bulb reduces overlap in the neural representation of olfactory stimuli. *Behavioural Brain Research*, 84:117–127, 1997.

53. C. Linster, Q. Nai, and M. Ennis. Nonlinear effects of noradrenergic modulation of olfactory bulb function in adult rodents. *Journal of Neurophysiology*, 105(4):1432–1443, 2011.

54. K. M. Lizbinski and A. M. Dacks. Intrinsic and extrinsic neuromodulation of olfactory processing. *Frontiers in Cellular Neuroscience*, 11(424):1–11, 2018.

55. V. M. Luna and N. E. Schoppa. GABAergic circuits control input–spike coupling in the piriform cortex. *Journal of Neuroscience*, 28(35):8851–8859, 2008.

56. N. Mandairon, F. Kermen, C. Charpentier, J. Sacquet, C. Linster, and A. Didier. Context-driven activation of odor representations in the absence of olfactory stimuli in the olfactory bulb and piriform cortex. *Frontiers in Behavioral Neuroscience*, 8:138, 2014.

57. N. Mandairon, S. Peace, A. Karnow, J. Kim, M. Ennis, and C. Linster. Noradrenergic modulation in the olfactory bulb influences spontaneous and reward-motivated discrimination, but not the formation of habituation memory. *European Journal of Neuroscience*, 27(5):1210–1219, 2008.

58. F. Markopoulos, D. Rokni, D. H. Gire, and V. N. Murthy. Functional properties of cortical feedback projections to the olfactory bulb. *Neuron*, 76(6):1175–1188, 2012.

59. C. Martin, R. Gervais, E. Hugues, B. Messaoudi, and N. Ravel. Learning modulation of odor-induced oscillatory responses in the rat olfactory bulb: A correlate of odor recognition? *The Journal of Neuroscience*, 24:389–397, 2004.

60. C. Martin, R. Gervais, B. Messaoudi, and N. Ravel. Learning-induced oscillatory activities correlated to odour recognition: a network activity. *European Journal of Neuroscience*, 23:1801–1810, 2006.

61. S. Matsutani. Trajectory and terminal distribution of single centrifugal axons from olfactory cortical areas in the rat olfactory bulb. *Neuroscience*, 169(1):436–448, 2010.

62. A. B. R. McIntyre and T. A. Cleland. Biophysical constraints on lateral inhibition in the olfactory bulb. *Journal of Neurophysiology*, 115(6):2937–2949, 2016.

63. K. Mori, K. Kishi, and H. Ojima. Distribution of dendrites of mitral, displaced mitral, tufted, and granule cells in the rabbit olfactory bulb. *The Journal of Comparative Neurology*, 219(3):339–355, 1983.

64. S. Nagayama, R. Homma, and F. Imamura. Neuronal organization of olfactory bulb circuits. *Frontiers in Neural Circuits*, 8(98):1–19, 2014.

65. K. R. Neville and L. B. Haberly. Beta and gamma oscillations in the olfactory system of the urethane-anesthetized rat. *Journal of Neurophysiology*, 90:3921–3930, 2003.

66. A. Nunez-Parra, R. K. Maurer, K. Krahe, R. S. Smith, and R. C. Araneda. Disruption of centrifugal inhibition to olfactory bulb granule cells impairs olfactory discrimination. *Proceedings of the National Academy of Sciences*, 110(36):14777–14782, 2013.

67. E. Orona, E. C. Rainer, and J. W. Scott. Dendritic and axonal organization of mitral and tufted cells in the rat olfactory bulb. *The Journal of Comparative Neurology*, 226(3):346–356, 1984.

68. G. H. Otazu, H. Chae, M. B. Davis, and D. F. Albeanu. Cortical feedback decorrelates olfactory bulb output in awake mice. *Neuron*, 86(6):1461–1477, 2015.

69. K. Padmanabhan, F. Osakada, A. Tarabrina, E. Kizer, E. M. Callaway, F. H. Gage, and T. J. Sejnowski. Diverse representations of olfactory information in centrifugal feedback projections. *Journal of Neuroscience*, 36(28):7535–7545, 2016.

70. K. Padmanabhan, F. Osakada, A. Tarabrina, E. Kizer, E. M. Callaway, F. H. Gage, and T. J. Sejnowski. Centrifugal inputs to the main olfactory bulb revealed through whole brain circuit-mapping. *Frontiers in Neuroanatomy*, 12, 2018.

71. J. Pager. Unit responses changing with behavioral outcome in the olfactory bulb of unrestrained rats. *Brain Research*, 289:87–98, 1983.

72. S. Pandipati, D. H. Gire, and N. E. Schoppa. Adrenergic receptor-mediated disinhibition of mitral cells triggers long-term enhancement of synchronized oscillations in the olfactory bulb. *Journal of Neurophysiology*, 104(2):665–674, 2010.

73. S. L. Pomeroy, A. Lamantia, and D. Purves. Postnatal construction of neural circuitry in the mouse olfactory bulb. *Journal of Neuroscience*, 10(6):1952–1966, 1990.

74. C. Poo and J. Isaacson. Odor representations in olfactory cortex: "sparse" coding, global inhibition, and oscillations. *Neuron*, 62:850–861, 2009.

75. C. Poo and J. S. Isaacson. A major role for intracortical circuits in the strength and tuning of odor-evoked excitation in olfactory cortex. *Neuron*, 72:41–48, 2011.

76. R. T. Pressler and B. R. Strowbridge. Direct recording of dendrodendritic excitation in the olfactory bulb: Divergent properties of local and external glutamatergic inputs govern synaptic integration in granule cells. *The Journal of Neuroscience*, 37:11774–11788, 2017.

77. W. Rall and G. M. Shepherd. Theoretical reconstruction of field potentials and dendrodendritic synaptic interactions in olfactory bulb. *Journal of Neurophysiology*, 31:884–915, 1968.

78. R. L. Rennaker, C. F. Chen, A. M. Ruyle, A. M. Sloan, and D. A. Wilson. Spatial and temporal distribution of odorant-evoked activity in the piriform cortex. *The Journal of Neuroscience*, 27:1534–1542, 2007.

79. G. E. Riccio, E. J. Martin, and T. A. Stoffregen. The role of balance dynamics in the active perception of orientation. *Journal of Experimental Psychology: Human Perception and Performance*, 18:624–644, 1992.

80. J. M. Ross and M. L. Fletcher. Learning-dependent and -independent enhancement of mitral/tufted cell glomerular odor responses following olfactory fear conditioning in awake mice. *The Journal of Neuroscience*, 38:4623–4640, 2018.

81. M. Rothermel and M. Wachowiak. Functional imaging of cortical feedback projections to the olfactory bulb. *Frontiers in Neural Circuits*, 8(73):1–14, 2014.

82. N. E. Schoppa, M. J. Kinzie, Y. Sahara, T. P. Segerson, and G. L. Westbrook. Dendrodendritic inhibition in the olfactory bulb is driven by NMDA receptors. *Journal of Neuroscience*, 18(17):6790–6802, 1998.

83. A. M. D. Shakhawat, A. Gheidi, Q. Hou, S. K. Dhillon, D. F. Marrone, C. W. Harley, and Q. Yuan. Visualizing the engram: learning stabilizes odor representations in the olfactory network. *Journal of Neuroscience*, 34(46):15394–15401, 2014.

84. A. M. D. Shakhawat, C. W. Harley, and Q. Yuan. Arc visualization of odor objects reveals experience-dependent ensemble sharpening, separation, and merging in anterior piriform cortex in adult rat. *Journal of Neuroscience*, 34(31):10206–10210, 2014.

85. M. T. Shipley and M. Ennis. Functional organization of olfactory system. *Journal of Neurobiology*, 30(1):123–176, 1996.

86. M. Sigman, G. A. Cecchi, C. D. Gilbert, and M. O. Magnasco. On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences*, 98(4):1935–1940, 2001.

87. F. W. Smith and L. Muckli. Nonstimulated early visual areas carry information about surrounding context. *Proceedings of the National Academy of Sciences*, 107:20099–20103, 2010.

88. D. L. Sosulski, M. L. Bloom, T. Cutforth, R. Axel, and S. R. Datta. Distinct representations of olfactory information in different cortical centres. *Nature*, 472(7342):213–215, 2011.

89. S. Srinivasan and C. F. Stevens. The distributed circuit within the piriform cortex. *The Journal of Comparative Neurology*, 526(17):2725–2741, 2018.

90. D. D. Stettler and R. Axel. Representations of odor in the piriform cortex. *Neuron*, 63(6):854–864, 2009.

91. M. L. S. Tantirigama, H. H.-Y. Huang, and J. M. Bekkers. Spontaneous activity in the piriform cortex extends the dynamic range of cortical odor coding. *Proceedings of the National Academy of Sciences*, 114(9):2407–2412, 2017.

92. T. Tesileanu, S. Cocco, R. Monasson, and V. Balasubramanian. Adaptation of olfactory receptor abundances for efficient coding. *eLife*, 8:1–39, 2019.

93. M. T. Tong, S. T. Peace, and T. A. Cleland. Properties and mechanisms of olfactory learning and memory. *Frontiers in Behavioral Neuroscience*, 8(238):1–16, 2014.

94. N. Uchida and Z. F. Mainen. Speed and accuracy of olfactory discrimination in the rat. *Nature Neuroscience*, 6(11):1224–1229, 2003.

95. N. Urban. Lateral inhibition in the olfactory bulb and in olfaction. *Physiology and Behavior*, 77:607–612, 2002.

96. D. W. Wesson, A. H. Borkowksi, G. E. Landreth, R. A. Nixon, E. Levy, and D. A. Wilson. Sensory network dysfunction, behavioral impairments, and their reversibility in an alzheimer's -amyloidosis mouse model. *The Journal of Neuroscience*, 31(44):15962–15971, 2011.

97. D. A. Wilson, R. M. Sullivan, and M. Leon. Single-unit analysis of postnatal olfactory learning: modified olfactory bulb output response patterns to learned attractive odors. *Journal of Neuroscience*, 7(10):3154–3162, 1987.

98. D. A. Wilson, W. Xu, B. Sadrian, E. Courtiol, Y. Cohen, and D. C. Barnes. Cortical odor processing in health and disease. In *Progress in Brain Research*, volume 208. Elsevier, 2014.

99. T. B. Woolf, G. M. Shepherd, and C. A. Greer. Serial reconstructions of granule cell spines in the mammalian olfactory bulb. *Synapse*, 7(3):181–192, 1991.

100. Y. Yamada, K. Bhaukaurally, T. J. Madarasz, A. Pouget, I. Rodriguez, and A. Carleton. Context- and output layer-dependent long-term ensemble plasticity in a sensory circuit. *Neuron*, 93:1198–1212, 2017.

101. C. Zhan and M. Luo. Diverse patterns of odor representation by neurons in the anterior piriform cortex of awake mice. *The Journal of Neuroscience*, 30(49):16662–16672, 2010.

102. L. A. Zimmer, M. Ennis, and S. M. T. Diagonal band stimulation increases piriform cortex neuronal excitability *in vivo*. *NeuroReport*, 10(10):2101–2105, 1999.