Biological Cybernetics manuscript No.
(will be inserted by the editor)

Self-organization in the olfactory system:

One shot odor recognition in insects

Thomas Nowotny[1], **Ramón Huerta**[1], **Henry D. I. Abarbanel**[1 2], **Mikhail I. Rabinovich**[1]

[1] Institute for Nonlinear Science, UCSD, 9500 Gilman Dr., La Jolla, CA 92093-0402
[2] Department of Physics and Marine Physical Laboratory (Scripps Institution of Oceanography), UCSD, La Jolla, CA 93093-0402

The date of receipt and acceptance will be inserted by the editor

Abstract We show in a model of spiking neurons that synaptic plasticity in the mushroom bodies in combination with the general fan-in, fan-out properties of the early processing layers of the olfactory system might be sufficient to account for its efficient recognition of odors. For a large variety of initial conditions the model system consistently finds a working solution without any fine-tuning, and is, therefore, inherently robust. We demonstrate that gain control through the known feedforward inhibition of lateral horn interneurons increases the capacity of the system but is not essential for its general function. We also predict an upper limit for the number of odor classes *Drosophila* can discriminate based on the number and connectivity of its olfactory neurons.

Key words Olfaction, pattern recognition, synaptic convergence, information coding, fan-in, fan-out

1 Introduction

Odor space is in contrast to visual or auditory space not endowed with an obvious structure. Accordingly it is not obvious how odor space should be represented in the olfactory brain structures. In insects, the information collected by receptor cells in the antenna is projected to glomeruli in the antennal lobe (AL). Each receptor cell expresses one receptor type [48], and all cells with the same receptor type project to the same glomeruli [31, 46]. As a result, olfactory information seems to be encoded as combinatorial activation patterns of glomeruli in the AL [21,40,11,29]. In locust, these patterns are transformed to spatio-temporal patterns of active projection neurons (PNs) [22,25,50,24,39] which improve the separation of representations of similar odors [42,18, 10]. The spatio-temporal code of the AL is transmitted to the mushroom body (MB) as discrete snapshots of activity [35]. These snapshots are reminiscent of sniffing behavior in mammals even though on a somewhat different time scale. It has been shown in rats that odor recognition can be very fast, often performed within a single sniff [47]. To investigate such rapid odor recognition we concentrate on information processing in a single snapshot. On the single snapshot level, the system needs to perform a one shot pattern recognition task with noisy patterns. For other insects, like Drosophila, honeybee or moth, where the spatio-temporal coding in the AL and snapshot transmission to the MB has not been demonstrated, the pattern recognition task addressed here can be interpreted as the rapid recognition of an initial activity pattern in the AL in response to an odor.

We consider the activity of the PNs in the AL in a given short time window as the input to a classification system consisting of the intrinsic Kenyon cells (iKC) of the MB and the extrinsic Kenyon cells (eKC) of the MB lobes. Classification systems in abstract neural networks, e.g., linear classificators or support vector machines, typically have a very specific connectivity and are trained to respond to given inputs with prescribed outputs, i.e., to represent given inputs in a prescribed way. In this work we explore an alternative computing strategy based entirely on random connectivity and self-organization through local learning and competition. We demonstrate that the known facts about the olfactory system of insects are consistent with a classification scheme that does not require tutoring, prescribed representation of information, or genetically or algorithmically determined special connectivity.

Some aspects of the resulting model resemble the ideas of support vector machines [6,5], however, with major differences:

1. The connectivity is unspecific, i.e. the system uses a random kernel function
2. The final representation of classes is self-organized through a type of spike timing dependent plasticity and mutual inhibition

3. the model system is built with realistic spiking neurons.

The classification procedure is split into two stages. The first stage is a nonlinear transformation from the AL to the MB, which separates the patterns in the PNs into sparse patterns in the iKCs realized by a non-specific connectivity matrix between AL and MB [12], which is consistent with observations in *Drosophila* [49]. The second stage is linear classification of iKC activity patterns by eKCs. This stage rests on a type of spike timing dependent plasticity and mutual inhibition between eKCs, the combination of which leads to self-organized, simple representations of odors which appear very suitable for association and memory. Experimental evidence for such self-organization has been shown in the piriform cortex of rats [53]. Direct experimental evidence for the type of representations seen in the model described here has, however, yet to be found.

We built our model with realistic spiking neurons and synapses obeying a local spike timing dependent plasticity rule. Even though the complexity of the model and its numerical cost in terms of computation time are much higher than for an artificial neural network, this step seems necessary to make the applicability of a model to real biological systems plausible. It has become clear in numerous works [15,14,27,41,3] that systems built with more realistic models of biological components can be quantitatively and qualitatively different from the more abstract connectionist approaches. In the context of this article, for example, non-disjoint representations of odors that were easily implemented in our previous more abstract work on the olfactory system of insects [19] could not be observed under any conditions in the more realistic system described here.

2 Model Description

2.1 Model Neurons

Our model is implemented with spiking neurons represented by a phenomenological, discrete time dynamical map [4,38] with fixed time steps $\Delta t = 0.5$ms. In contrast to phenomenological conductance based models, this map can be computed very quickly such that we can simulate neural ensembles with realistic population sizes, on the order of thousands of neurons. In particular, the membrane voltage $V(t+\Delta t)$ of a neuron at time $t+\Delta t$ is

$$V(t+\Delta t) = \begin{cases} V_{\text{spike}}\left(\frac{\alpha V_{\text{spike}}}{V_{\text{spike}}-V(t)-\beta I_{\text{syn}}}+\gamma\right) & V(t)\le 0 \\ V_{\text{spike}}\,(\alpha+\gamma) & V(t)\le V_{\text{spike}}\,(\alpha+\gamma)\ \&\ V(t-\Delta t)\le 0 \\ -V_{\text{spike}} & \text{otherwise,} \end{cases} \tag{1}$$

where $V_{\text{spike}} = 60$ mV, $\alpha = 3$, $\beta = 2.64$ MΩ, and $\gamma = -2.468$. β reflects the input resistance of the cell and was chosen such that match with the corresponding values for the equivalent Hodgkin-Huxley model and Rall type synapses (see below, Fig. 1).

The model neurons exhibit a simple spiking behavior in response to DC input (Fig. 1A) very much like Hodgkin-Huxley neurons (Fig. 1B). The Hodgkin-Huxley neuron shown for comparison is a standard model [45] with three conductances, I_{Na}, I_{Kd}, and I_{leak}. In particular,

$$\begin{aligned}\frac{dV(t)}{dt} = -\frac{1}{C}\Big(&g_{\text{Na}}m(t)^3h(t)\big[V(t)-E_{\text{Na}}\big]+\\ &g_{\text{K}}n(t)^4\big[V(t)-E_{\text{K}}\big]+\\ &g_{\text{leak}}\big[V(t)-E_{\text{leak}}\big]\Big)-I_{\text{syn}}.\end{aligned} \tag{2}$$

The activation and inactivation variables m, h, and n are governed by $\frac{dy}{dt} = \alpha_y(V(t))[1-y(t)] - \beta_y y(t)$, where y represents m, h, and n, respectively and

$$\begin{aligned}\alpha_m &= 0.32(-52-V)/(\exp((-52-V)/4)-1)\\ \beta_m &= 0.28(25+V)/(\exp((25+V)/5)-1)\\ \alpha_h &= 0.128\exp((-48-V)/18)\\ \beta_h &= 4/(\exp((-25-V)/5)+1)\\ \alpha_n &= 0.032(-50-V)/(\exp((-5-V)/5)-1)\\ \beta_n &= 0.5\exp((-55-V)/40)\end{aligned}$$

The conductances and reversal potentials are $g_{\text{Na}} = 7.15\mu$S, $E_{\text{Na}} = 50$ mV, $g_{\text{K}} = 1.43\mu$S, $E_{\text{K}} = -95$ mV, $g_{\text{leak}} = 0.0267\mu$S, $E_{\text{leak}} = -63.56$ mV, and the membrane capacitance is $C = 0.3$ nF.

The map model and the conductance based model have also very similar spike latencies in response to EPSPs of different size (Fig. 1C). Because all neurons in this work operate in a very sparse regime the good agreement of the two models in the shown tests implies that the conductance based model would yield qualitatively and quantitatively the same results as the map neurons.

We used the original map neurons [38] that match this generic and well-tested Hodgkin-Huxley model for spiking neurons instead of developing a custom model to match the emerging line of models for the honeybee KCs [20,34,54]. While the current study is closer to the biological systems than our previous work [19], it is not yet describing a specific insect and specializing to a neuron model designed to match honeybee KCs would be pretentious. Nevertheless, the chosen model is as a type 1 model well suited to allow low firing frequencies and the observed sparse activity in KCs [35]. The response latency as a function of the synaptic input is qualitatively similar to Wüstenberg's model of honeybee KCs [54], see inset in figure 1C.

Reaching a steady state in the learning process requires long simulation runs; The computational speed of map based neurons allowed us to simulate our model

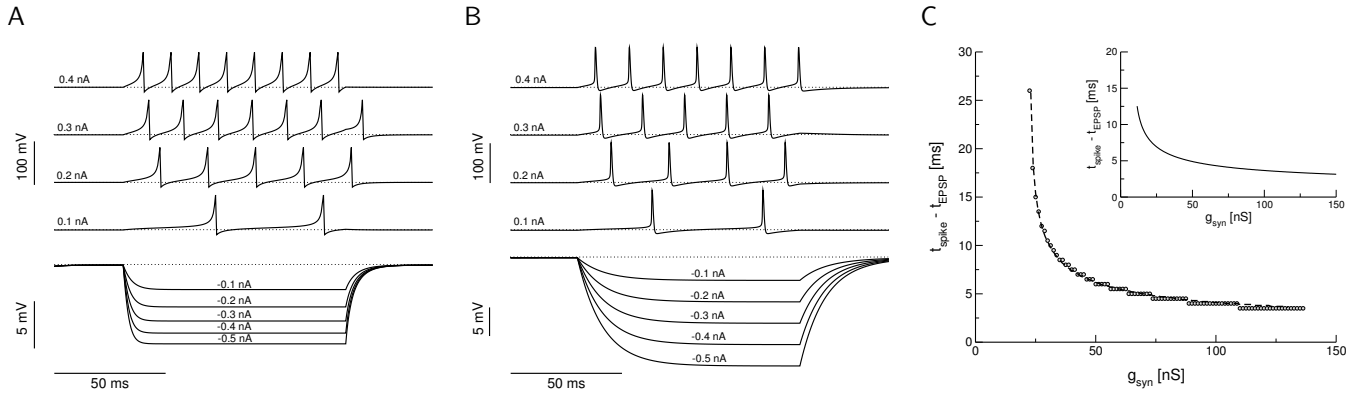

Fig. 1 A) Response of the neuron model to DC current injection of −0.5 to 0.4 nA in steps of 0.1 nA. B) Response of a standard conductance based model [45] to DC current input for comparison. C) Spike response latency for the map model (circles) and the conductance based model (dashed line) in response to EPSPs evoked by presynaptic voltage pulses from a resting potential of −60 mV to +50 mV of 1.5 ms duration ($\tau_{\rm syn} = 2$ ms). The inset shows the corresponding response latency curve for the KC model of Wüstenberg et al. [54] ($\tau_{\rm syn} = 10$ ms).

system with 2720 neurons and on the order of 150,000 synapses over 100,000 ms on standard PC hardware. The map model was used for all neurons except the PNs. PN activity was represented by short, square voltage pulses at the spike times determined by the input patterns. The resting potential was set to -60 mV and the pulse voltage to 50 mV with a duration of 1.5 ms.

2.2 Synaptic Currents

Synapses are modeled by

$$I_{\rm syn}(t+\Delta t) = g_{\rm syn} S(t)(V_{\rm rev} - V_{\rm post}(t)), \tag{3}$$

$$S(t+\Delta t) = \begin{cases} S(t)e^{-\Delta t/\tau_{\rm syn}} + \delta & \text{presynaptic spike at } t \\ S(t)e^{-\Delta t/\tau_{\rm syn}} & \text{otherwise.} \end{cases} \tag{4}$$

$S(t)$ describes the amount of neuro-transmitter active at the postsynaptic receptors. At each time step in which the presynaptic neuron is spiking, a quantile characterized by $\delta = 0.25$ is released within Δt and then decreases exponentially with rate $\Delta t/\tau_{\rm syn}$ implementing neuro-transmitter diffusion and uptake. The presynaptic neuron is considered to be spiking when its membrane potential first crosses 0 mV. For the voltage pulses representing PN activity, spiking is synonymous to the voltage being at its high value 50 mV. $V_{\rm rev}$ is the reversal potential of the synapse and is $V_{\rm rev} = 0$ mV for all excitatory and $V_{\rm rev} = -92$ mV for all inhibitory synapses. $V_{\rm post}(t)$ denotes the membrane potential of the postsynaptic neuron.

2.3 Network geometry

The model network is illustrated in Fig. 2. Even though the architecture and the coding in our classification system were mainly inspired by the findings in locust, a full size simulation of the locust olfactory system is computationally too expensive. We, therefore, built a smaller system that roughly follows the statistics of Drosophila with 100 PNs, 2500 iKCs and 100 eKCs. We performed simulations with and without feedforward gain control. Connectivity is determined by a random process and then fixed throughout the simulation. Each PN-iKC pair is connected with probability $p_{\rm PN,\,iKC} = 0.15$. The synaptic timescale was $\tau_{\rm PN,\,iKC} = 2$ ms and the mean synaptic strength was $g_{\rm PN,\,iKC} = 4.545$ nS without and $g_{\rm PN,\,iKC} = 5.25$ nS with gain control. This choice matches the average activity in the iKCs for both cases. We added a Gaussian jitter with standard deviation 1.25 nS to the mean synaptic strengths. The gain control was implemented through 20 lateral horn interneurons (LHIs) receiving input from all PN and inhibiting all iKCs. The inputs to LHIs have strengths leading to increasing inhibition onto the iKCs with increasing activity in the PNs. In particular, the synapses afferent to LHI number n have maximal conductance $\theta/(c+n)$, with $\theta = 53.75$ nS and $c = 15$, and time scale $\tau_{\rm PN,\,LHI} = 1$ ms. The inhibitory synapses from LHIs to iKCs are characterized by $g_{\rm LHI,\,iKC} = 8.75$ nS and $\tau_{\rm LHI,\,iKC} = 3$ ms.

Any iKC-eKC pair is connected with probability $p_{\rm iKC,\,eKC}$. An existing synapse is initially active (strong) with probability $p^+_{\rm iKC,\,eKC}$ and inactive (weak) with $1 - p^+_{\rm iKC,\,eKC}$. We initially use all-to-all connectivity, $p_{\rm iKC,\,eKC} = 1$ and 20% active synapses, $p^+_{\rm iKC,\,eKC} = 0.2$. The strengths for active and inactive synapses are drawn from Gaussian distributions with mean $g^+_{\rm iKC,\,eKC} = 1.25$ nS and standard deviation $\sigma^+_{\rm iKC,\,eKC} = 0.25$ nS for the active synapses and $g^-_{\rm iKC,\,eKC} = 0.125$ nS and $\sigma^-_{\rm iKC,\,eKC} = 0.025$ nS for the inactive synapses if not stated otherwise. The time scale for excitatory synapses is $\tau_{\rm iKC,\,eKC} = 10$ ms.

The eKCs in the MB lobes interact through all-to-all mutual inhibition. The mutual inhibition is implemented by inhibitory synapses governed by 3 with $g_{\rm syn} = 75$ nS

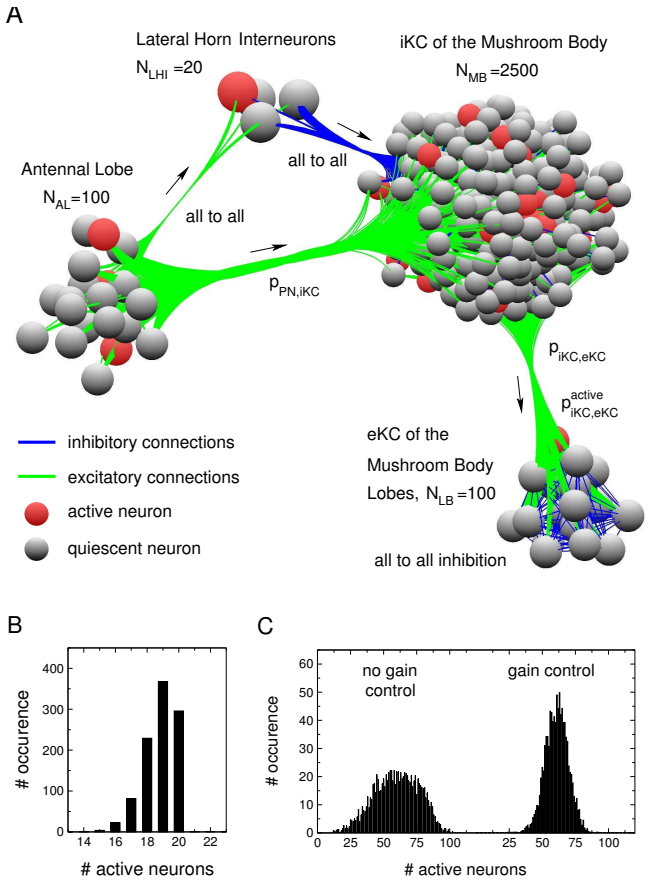

Fig. 2 A) Model system in 1 : 5 scale. Panels B) and C) show the distribution of the number of active neurons in the AL (B) and the MB (C) for a structured set of 1000 inputs ($p_{\mathrm{AL}} = 0.2$, $p_{\mathrm{perturb}} = 0.1$). The distribution of the number of active neurons in the MB lobes depends on the details of the learning rule and the history of learning and is not shown. Note the difference in the MB activity with and without gain control.

and $\tau_{\mathrm{syn}} = 5$ ms. The assumption of such a connectivity is motivated by our computational paradigm and as of now a speculation.

2.4 Synaptic Plasticity

All iKCs-eKCs synapses are modified by a spike timing dependent plasticity rule,

$$g_{\mathrm{raw}}(t+\Delta t) = \begin{cases} g_0 + (g_{\mathrm{raw}}(t) - g_0)e^{-\Delta t/\tau_{\mathrm{decay}}} & \text{no spikes at } t \\ g_0 + (g_{\mathrm{raw}}(t) - g_0)e^{-\Delta t/\tau_{\mathrm{decay}}} & \text{pre or postsynaptic} \\ \quad +A(t_{\mathrm{post}} - t_{\mathrm{pre}} - \tau_{\mathrm{shift}}) & \text{spike at } t \end{cases} \tag{5}$$

$$A(\tau) = \begin{cases} y_- & \tau < \tau_- \\ a_-\tau + y_0 & \tau_- < \tau \le 0 \\ a_+\tau + y_0 & 0 < \tau \le \tau_+ \\ y_+ & \tau_+ < \tau \end{cases} \tag{6}$$

with $\tau_- = -(1/c_{10}+1/c_{11})\tau_l c_{11}/2, \tau_+ = (1/c_{01}+1/c_{11})\tau_l c_{11}/2$, $a_- = -a_+ = 2g_{\max}/(\tau_l c_{11})$, $y_- = -g_{\max}/c_{10}$, $y_0 = g_{\max}/c_{11}$, and $y_+ = -g_{\max}/c_{01}$. While the definition might look rather complex, the learning rule is basically spike timing dependent plasticity including a time delay for the finite transmission speed of the synapse. The parameters are $c_{10} = 10^5$, $c_{01} = 20$, $c_{11} = 5$, $\tau_l = 25$ ms, $\tau_{\mathrm{shift}} = 10$ ms, $\tau_{\mathrm{decay}} = 10^5$ ms, and $g_0 = 0.125$ nS if not stated otherwise. Because of the large value of c_{01} the synaptic strength is only depressed if post-synaptic spikes are not paired with presynaptic ones, consistent with experimental observations [36,37,28,30] and biophysical models [1,52].

To avoid run-away behavior of the synaptic strength the raw synaptic strength $g_{\mathrm{raw}}(t)$ is filtered by a sigmoidal function $g_{\mathrm{syn}} = g_{\max}(\tanh((g_{\mathrm{raw}}(t)-g_{\mathrm{mid}})/g_{\mathrm{slope}})+1)/2$. This filter implements the restriction of the maximal synaptic conductance of real synapses due to limited resources in terms of neurotransmitters and the maximal number of ion channels. The maximal synaptic strength $g_{\max}$ is important in determining the average activity in the eKCs. It is matched to produce approximately the same average activity in all simulations. If not mentioned otherwise $g_{\max} = 3.75$ nS, $g_{\mathrm{mid}} = g_{\max}/2$, ans $g_{\mathrm{slope}} = g_{\mathrm{mid}}$.

2.5 Input sets and simulation

The activity patterns of the AL are chosen according to the following prescription. We first choose N_{class} basis patterns $\mathbf{x}_\mu$, $\mu = 1,\ldots,N_{\mathrm{class}}$ by randomly choosing $n_+ = p_{\mathrm{AL}}N_{\mathrm{AL}} = 20$ active neurons for each pattern. The patterns are stored as binary vectors, where 1 means a neuron is active in this input pattern and 0 means it is quiescent. The basis patterns are chosen independently. Then we use each of the basis patterns to generate a class of input patterns, $\mathbf{x}^\nu_\mu$, $\nu = 1,\ldots,n_{\mathrm{class}}$ by switching active neurons off (replacing 1 by 0) and other, randomly chosen, neurons on (replacing the previous value by 1). This perturbation happens with a probability p_{perturb} which determines the similarity of patterns within a class; the lower p_{perturb}, the more similar are the input patterns within each class. We used $p_{\mathrm{perturb}} = 0.1$ in most simulations. By this mechanism we get classes of inputs that are highly correlated within the class and uncorrelated across classes. The fluctuation in the number of active neurons is small (Fig. 2B) and stems solely from perturbations in which an active neuron is switched off and the chosen replacement was already active.

The resulting patterns resemble the observed activity in the AL of locust with respect to the average activity of PNs of about 20% per local field potential cycle and the distribution of active neurons that is close to a binomial distribution [43]. So far no measurements of the spiking activity pattern of the complete AL could be performed.

Testing our system with observed activity patterns from the AL of, e.g., locust would be of great interest but will have to be delayed until such data become available. During the initial experience phase during which the system forms its representation of odor inputs, it periodically receives randomly picked input patterns. All input patterns are perfectly synchronized assuming that there is no information contained in the spike timing within a 50 ms time window. The experienced system and the naive are then tested on a new set of inputs that was generated from the same class basis vectors as the training set. Synapses stay plastic during the testing phase except when the naive system's performance is tested. In testing, all inputs are given in a fixed order, each input once.

2.6 Computational procedures

The system is implemented in C++, compiled with the GNU g++ compiler version 2.96 (RedHat) and run on a RedHat Linux MOSIX cluster with 22 AMD Athlon processors. 100 simulated seconds execute on a single processor in about 2000 minutes.

2.7 Distance functions

We investigated the classification performance in terms of distances of activity vectors of the total activity in 50 ms snapshots. The norm of an activity vector is $|\mathbf{z}| = \sum_{i=1}^{N_X} |z^i|$, where z^i, $i = 1, \ldots, N_X$, denotes the entries of the N_X dimensional vector $\mathbf{z}$, that represent whether there is a spike (1) or not (0) in the ith neuron. The distance between two activity patterns is $D(\mathbf{z}_\mu, \mathbf{z}_\nu) = |\mathbf{z}_\mu - \mathbf{z}_\nu|$. The average class vector of class μ is defined as $\langle \mathbf{z}_\mu \rangle = \langle \mathbf{z}_\mu^\nu \rangle_\nu = \frac{1}{n_{\text{class}}} \sum_{\nu=1}^{n_{\text{class}}} \mathbf{z}_\mu^\nu$. The inter-class distance is then measured as the distance between average class vectors,

$$D_{\text{inter}} = \frac{1}{N_{\text{class}}(N_{\text{class}} - 1)/2} \sum_{\mu=1}^{N_{\text{class}}-1} \sum_{\nu>\mu} D(\langle \mathbf{z}_\mu \rangle, \langle \mathbf{z}_\nu \rangle), \tag{7}$$

whereas the intra-class distance is measured as the average distance of class members to the average class vector,

$$D_{\text{intra}} = \frac{1}{N_{\text{class}}} \sum_{\mu=1}^{N_{\text{class}}} \frac{1}{n_{\text{class}}} \frac{\nu = 1}{n_{\text{class}}} D(\mathbf{z}_\mu^\nu, \langle \mathbf{z}_\mu \rangle). \tag{8}$$

For a fair comparison between distances of activity patterns in layers of different sizes and average activity levels, we normalized distances to the number of active neurons. In particular,

$$D_{\text{act}}(\mathbf{z}_\mu, \mathbf{z}_\nu) = D(\mathbf{z}_\mu, \mathbf{z}_\nu)/(|\mathbf{z}_\mu| + |\mathbf{z}_\nu|). \tag{9}$$

This normalized distance is bounded between zero and one, $0 \leq D_{\text{act}} \leq 1$ and was used in the calculation of D_{inter} and D_{intra} bounding them in the same way.

3 Results

3.1 General classification performance

Our main tool for analyzing classification performance is the comparison of the total activity within one snapshot in the different layers of the olfactory system in response to different input patterns. For this analysis we represent the activity in the three layers of the system as binary vectors of dimension N_X, X representing AL, MB or LB. Each entry in these vectors is one, if the corresponding neuron spiked within a 50 ms time window, and zero, if it did not. Because of the sparseness of activity in the system, there are never two spikes within the 50 ms time window corresponding to one snapshot. We measure the separation of responses to inputs of different classes and the identification of inputs of the same class using distance functions of activity vectors. In particular, we compare the distance between the average activity vectors of classes of inputs, D_{inter}, to the average distance of the activity vectors of inputs within each class to the average activity vector, D_{intra}. To compare between representations in the different layers of the system that have different size, all distances are normalized by the total activity. Details are given in the description of the distance functions (section 2.7). Fig. 3 shows an example of the normalized distances between the activity patterns in response to individual inputs in the different layers for sets of inputs with 10, 20, 50 and 100 input classes of 10 inputs each, obtained without gain control in the MB. For details about the choice of input classes confer to the model description (section 2.5). The structure of the input is apparent from the distances in the AL. While being rather large between inputs of different classes, the distances within the classes are rather small. In the iKCs both types of distances are enlarged, even in terms of the normalized distance per active neuron. In the MB lobes the differences between distances within classes and between classes are smeared out in the naive system. After learning, however, the distances are enlarged between classes and diminished within the classes. The system successfully self-organizes and chooses an efficient representation for the encountered structured input. The quantitative analysis in terms of inter- and intra- distances is shown in Fig. 4. For 10, 20 and 50 input classes the system successfully classifies all inputs after 100 s experience with one input per 50 ms. Class representations in the eKCs of the MB lobes are formed by disjoint sets of representing neurons for each class, i.e., each neuron is active for only one class. Depending on the number of input classes the representation of the classes is more or less sparse. If the number of input classes is small the system has sufficiently many eKCs to represent classes by the activity of several neurons for each class without overlaps. Such redundancy in representing few input classes is indeed observed. If the number of different input classes is larger, less and less neurons are available

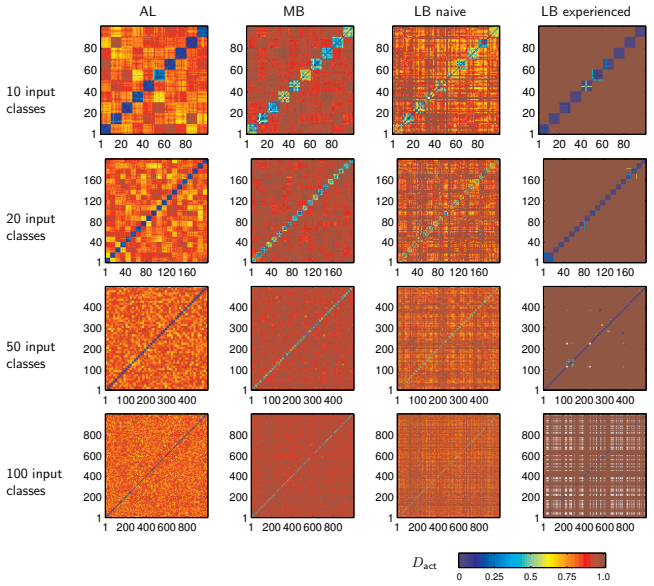

Fig. 3 Pairwise distances of activity patterns in the AL, MB and LB in response to inputs structured in input classes. the color of the square at (x, y) shows the normalized distance D_{act} between input number x and y. The inputs are ordered by classes, i.e., the first 10 inputs are from class 1, the next from class 2, and so on. The color white corresponds to quiescence in response to one (or both) of the inputs. The system successfully classifies for 10, 20, and 50 input classes whereas cases of no response in the LB layer appear for 100 input classes.

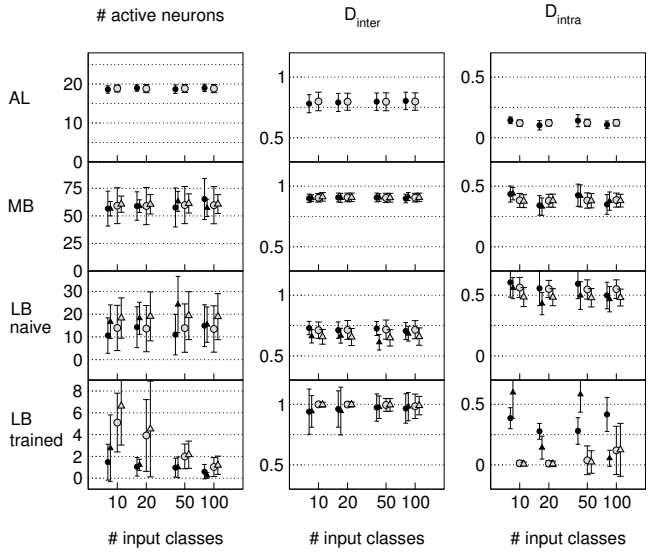

Fig. 4 The gray symbols show the system performance with (triangles) and without (circles) gain control. All data are averages over 4 independently chosen sets of inputs. The error bars mark the averages of the standard deviations within the individual trials. The effect of gain control can be seen in the reduced standard deviation of the average number of active iKC. Apart from a slight increase in overall eKC activity no significant changes occur in the MB lobes. The black symbols show the response to inputs of a new, unknown class after the system experienced a set of input classes for 100 s. The clearly reduced activity and the large intra-distance in the MB lobes for the new inputs show that the system classifies the new input as unknown rather than as a member of a known class.

to represent a class and the system adjusts accordingly keeping the representations disjoint. Strictly disjoint representations eventually lead to a partial failure in the form of no response to some inputs if the number of input classes reaches the number of eKC, see the results for 100 input classes in Fig. 4. As the MB is not the only olfactory pathway and has been shown to not be essential for basic odor perceptions by ablation experiments [16,7] this mode of failure should be interpreted as a failure to remember all classes rather than as a failure to perceive input odors. In contrast to an n-winner-take-all representation [33], the self-organized representation of odors in our system is solely determined by the learning rule, connectivity and type of activity in the system. Other, e.g., non-disjoint, representations were not observed even when changing the learning rule parameters or the maximal conductances of iKC-eKC synapses. Overly weak synapses lead to complete quiescence in the MB lobes and overly strong synapses, caused by a high upper limit of the synapse strength or too efficient reinforcement, drove the system to a complete activation for any input.

It is natural to ask what happens if the experienced system encounters inputs of a new, unknown input class. As the black symbols in figure 3 show, the inputs of the new class mostly lead to quiescence of the output neurons. We interpret this as the system not recognizing these inputs as being inputs of any of the known classes. The reinforcement of synaptic strength decays slowly over time which allows the system in principle to re-organize its representation if the new input class is consistently encountered. The detailed investigation of this aspect is, however, beyond the scope of this paper.

3.2 Gain control

The general statistical properties of non-specific connectivities restrict the activity levels for successful classification to a small set of allowed levels [19], and the olfactory system, at least in locust, indeed seems to have a gain control mechanisms to regulate activity levels in the in the AL and the MB [43]. The gain control within the AL is mediated by inhibitory interneurons whereas the iKCs of the MB are subject to a feed-forward periodic inhibition. The feedforward inhibition has two functions. It separates the activity of the AL into discrete snapshots as discussed in the introduction and because it is driven by the average activity of the PNs in AL it also provides a feedforward gain control on the iKCs of the MB. From the prevalence of gain control mechanisms in the system one expects that gain control is an important factor. It

turns out, however, that for the input classes with small overlaps used here, gain control for the activity of iKCs does not seem to be of a major advantage (Fig. 4). We show below that gain control in the iKCs becomes important when input classes overlap more significantly. The irrelevance of gain control for a low degree of overlap in the input classes indicates inherent robustness of the structure of the olfactory system. The use of input patterns of nearly constant activity levels in this work implies, however, the assumption of efficient gain control in the AL.

3.3 MB-LB connectivity

Learning in a statistical framework consists of removing or introducing synaptic connections. In biological systems, however, connectivity is much less plastic, at least in the adult animal. Here, learning consists of changing synaptic conductances of existing synapses rather than forming new synaptic connections. To test whether the ability to form any connection between iKCs and eKCs is important for the classification task, we introduced randomly chosen fixed connectivities of varying connectivity degree and compared the learning success. Fig. 5A shows the classification performance for different connectivity degrees, $p_{\mathrm{iKC,\ eKC}} = 1, 0.75, 0.5$, and 0.2, while the probability of existing synapses to be initially active is kept the same, $p^{+}_{\mathrm{iKC,\ eKC}} = 0.2$. For an unbiased comparison we adjusted the synaptic strengths of the iKC-eKC synapses, such that the expectation value of the total synaptic strength of afferent synapses to each eKC is the same for all connectivities used. The resulting parameter sets are shown in table 1. Fig. 5A shows that the performance of the system is not significantly reduced if the number of available synapses is decreased. Only the intra-distances seem to be somewhat larger than for the full connectivity. Even though the robustness the system against removal of possible connections might seem surprising, closer analysis reveals that the learned input classes are stored in just a few of the total number of synapses (Fig. 5B,C). Our classification system thus is extremely robust as it finds its way to work quite efficiently regardless of the initial conditions of connections from iKCs to eKCs.

3.4 Structure of input classes

It is clear from the Fig. 4 and 5 that system performance depends on the number of trained input classes. The other limiting factor is, obviously, the structure of the input. For the analysis presented thus far we always used input sets with very clearly separated classes characterized by the low value $p_{\mathrm{perturb}} = 0.1$. Fig. 6 shows the classification performance for two less structured sets of inputs with $p_{\mathrm{perturb}} = 0.2$ and $p_{\mathrm{perturb}} = 0.3$. Here, one can clearly see a failure of the system for the less

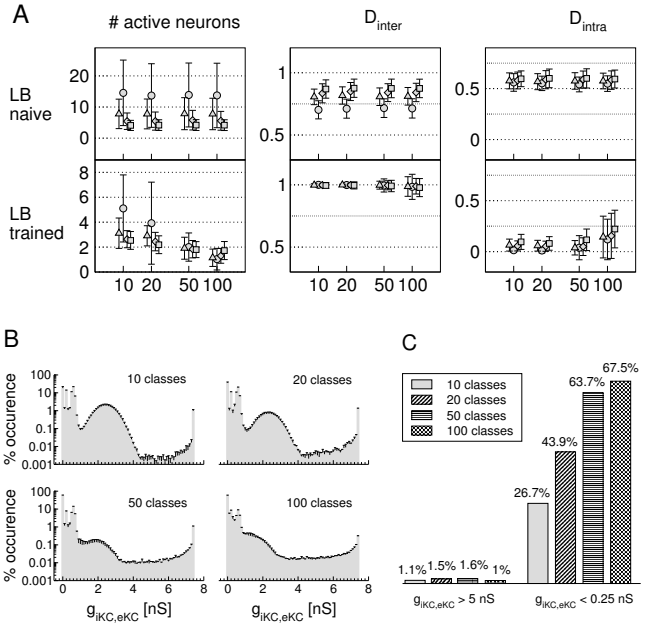

Fig. 5 A) System performance depending on the connectivity between iKCs and eKCs. Data points and standard deviations are again averages over 4 input sets. The circles correspond to full, triangles to 75%, diamonds to 50% and squares to 20% connectivity between iKCs and eKCs. The main changes with decreasing connectivity degree are decreased activity in the lobes in the naive state as well as in the experienced state and slightly higher intra-distances. B) Final distribution of synapse strengths after 100s experience. C) Percentile of strengthened (with final strength > 1 nS) and weakened (with final strength < 0.05 nS) synapses. Only a small number of synapses carries the learned information.

structured input and large numbers of input classes. The failure appears in form of quiescence of the lobe neurons in response to all input (not shown explicitly in Fig. 6, see Supplement). Fig. 6 also shows that for large overlaps between classes, gain control plays a critical role to increase the number of classifiable inputs.

Finally we can predict that, according to this analysis, the maximum number of statistically independent input classes of the type investigated here that a system of Drosophila size can classify lies between 20 and 50.

In order to compare to behavioral experiments a few more steps are necessary:

1. To observe the representation of odors in the AL experimentally and calculate the distances of the representations of as many odors/odor classes as possible. The distances have to be measured on the projection neuron response patterns. This would allow to estimate the structure of the input space.
2. Use this input space for our learning machine to automatically separate classes.
3. Determine in our simulations how many input classes of the observed type the model system can learn.

Table 1 Parameter sets used to compare performance with different connectivities between iKCs and eKCs. The parameters have been chosen to keep the total synaptic strength of all afferent synapses to each eKC constant on average.

$p_{\mathrm{iKC,\ eKC}}$	$g^{-}_{\mathrm{iKC,\ eKC}}$[nS]	σ^{-}[nS]	$g^{+}_{\mathrm{iKC,\ eKC}}$[nS]	σ^{+}[nS]	$g_{\max}$[nS]
1	0.25	0.05	2.5	0.5	7.5
0.75	0.333333	0.0666667	3.333333	0.666667	10
0.5	0.5	0.1	5	1	15
0.2	1.25	0.25	12.5	2.5	37.5

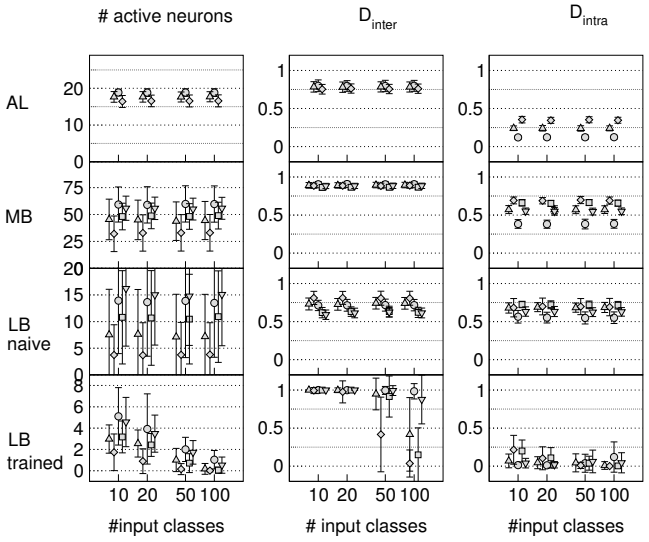

Fig. 6 System performance depending on the structure of the input set. Sets with a different degree of variability in the individual input classes were tested with the usual protocol. p_{perturb} was chosen as $p_{\mathrm{perturb}} = 0.1$ (circles), 0.2 (upward triangles without and downward triangles with gain control), and 0.3 (diamonds without and squares with gain control). The performance decreases with increasing variability in the input classes. While this decrease is negligible for few (10, 20) input classes it becomes quite considerable for 50 and leads to failure to classify for 100. Gain control in the MB improves the results considerably. The shown data are averages over 4 independent input sets each.

4. Compare this prediction to behavioral experiments with the insect, e.g., in an appetitive associative learning paradigm.

4 Discussion

Several attempts have been made to understand odor discrimination in models of the olfactory system [17, 51,13,26,9,3,32,19]. The model discussed here is, however, fundamentally different in providing a classification scheme that solely relies on the fan-in, fan-out properties of synaptic connections, the known locus of learning in the MB and otherwise entirely non-specific connectivity. At the same time it has been implemented with realistic enough model components to make it a plausible candidate for describing the olfactory system of insects. We have shown that our model system can accomplish successful classification of odor input patterns. The system self-organizes by spike timing dependent plasticity and mutual inhibition to classify combinatorially encoded input classes with simple, disjoint representations. We proved the robustness of the scheme in two ways. First, broadly modifying the connectivity between iKCs and eKCs does not prevent the system from good performance. Secondly, the odor classifications system also works well for a range of gains in the iKCs. The existing gain control only improves classification when the input classes overlap significantly. Based on our successful and robust classification scheme, we make a prediction on the number of uncorrelated input odor classes a system of Drosophila size can discriminate. Finally, the self-organized disjoint representation is suitable for associative learning which corresponds with the widely accepted hypothesis that odor conditioning has its neural correlate in the MB structure.

Our model was constructed with an unspecific connectivity for the projections from the AL to the MB as well as for the connections between iKCs and eKCs. We were able to show that there is no need for an explicitly or algorithmically specified special connectivity. The classification works on a purely statistical basis. In recent work in Drosophila [29,23,55,44] the specificity of PN projections to the MB and the protocerebrum has been investigated. Marin et al. [29], e.g., state: "inspection of axon collateral projections of different classes of PNs did not reveal obvious stereotype as compared to the striking stereotype of lateral horn axon branching pattern and terminal fields ." Other authors interpret the data differently such that the situation remains as yet unclear. With respect to the model presented here, while specificity is, as discussed, not a requirement for the system, a suitable refinement of completely unspecified connectivity might even improve the representations in the MB. In this sense the classification performance of the model presented in this paper would be a lower limit to the possible performance of systems with more specific connectivity.

The main limitations of our olfactory classification system are the total number of input classes and the amount of structure in the input set. If the number of input classes exceeds the number of eKCs in the MB lobes, disjoint representation becomes impossible and the system fails to respond to some of the inputs. It is noteworthy that even in this condition the system still successfully separates input classes. It is therefore not driven into complete failure by overwhelmingly large numbers

of different inputs. Rather than misclassifying inputs, it is quiescent in response to some inputs while classifying the remaining inputs correctly. The failure to respond to some of the inputs can be interpreted as a failure to remember all input classes. The odor inputs that lead to such a response failure can still be perceived through parallel olfactory pathways consistent with the observation that MBs are not essential for basic odor perception. If the structure of the input, i.e., the ratio of intra- and inter-distances, becomes less and less pronounced, the system either starts to resolve sub-classes for small numbers of input classes or is driven into quiescence earlier than for more structured input. Again, system failure does not result in misclassification but in a failure to respond to some of the inputs. We have not addressed the role of entrainment or experience time in detail here. It is, however, quite amazing that the system classifies inputs correctly even if, as in the example of 100 input classes, on average only 20 inputs from each class were presented to the system.

Noise in the olfactory receptors and AL dynamics is encoded in the variations of inputs within input classes, whereas internal noise in the iKCs and eKCs has not been considered. Experimental data shows [35] that at least the iKCs seem to suffer very little from internal noise. We suggest that the remaining noise will be neutralized through the redundancy of multiple snapshots, i.e., the temporal aspects of the olfactory code in the AL [42]. Thus, processing single temporal snapshots is sufficient for clearly distinguishable odorants where noise does not play a major role whereas more complex stimuli, including mixtures of odors, require processing of multiple snapshots [10,53].

There is some evidence that the MB is a multi-modal integration region combining, apart from olfactory inputs, also visual [2,8] and possibly other information. The system investigated in this paper does not depend on the type and source of information projected to the MB. Without modification, it classifies events of coincident multimodal input in the MB in the same way as purely olfactory input. In this sense the system is a universal classifier.

Acknowledgements We are grateful to Gilles Laurent, Rafael Levi, and Julie Haas for helpful remarks. This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Engineering and Geosciences, under grants DE-FG03-90ER14138 and DE-FG03-96ER14592, by grants from the National Science Foundation, NSF PHY0097134 and NSF EIA0130708, and by a grant from the Army Research Office, DAAD19-01-1-0026.

References

1. H. D. I. Abarbanel, R. Huerta, and M. I. Rabinovich. Dynamical model of longt-term synaptic plasticity. *P Natl Acad Sci USA*, 99:10132–10136, 2002.
2. M. Barth and M. Heisenberg. Vision affects mushroom bodies of *Drosophila melanogaster*. *Learn Mem*, 4:219–229, 1997.
3. C. D. Brody and J. J. Hopfield. Simple networks for spike-timing-based computation, with application to olfactory processing. *Neuron*, 37:843–852, 2003.
4. B. Cazelles, M. Courbage, and M. Rabinovich. Anti-phase regularization of coupled chaotic maps modelling bursting neurons. *Europhys Lett*, 56:504–509, 2001.
5. C. Cortes and V. Vapnik. Support vector networks. *Mach Learn*, 20:273–297, 1995.
6. T. Cover. Geometric and statistical properties of systems of linear in-equalities with applications in pattern recognition. *IEEE T Electron Comput*, 14:326, 1965.
7. J. S. de Belle and M. Heisenberg. Associative odor learning in Drosophila abolished by chemical ablation of mushroom bodies. *Science*, 263:692–695, 1994.
8. B. Ehmer and W.Gronenberg. Segregation of visual input to the mushroom bodies in the honeybee (*Apis mellifera*). *J Comp Neurol*, 451:362–373, 2002.
9. B. Ermentrout, J. W. Wang, J. Flores, and A. Gelperin. Model for olfactory discrimination and learning in limax procerebrum incorporating oscillatory dynamics and wave propagation. *J Neurophysiol*, 85:1444–1452, 2001.
10. R. W. Friedrich and G. Laurent. Dynamic optimization of odor representations by slow temporal patterning of mitral cell activity. *Science*, 291:889–894, 2001.
11. C. G. Galizia, A. Küttner, J. Joerges, and R. Menzel. Odour representation in honeybee olfactory glomeruli shows slow temporal dynamics: an optical recording study using voltage-sensitive dyes. *J Insect Physiol*, 46:877–886, 2000.
12. M. Garcia-Sanchez and R. Huerta. Design parameters of the fan-out phase of sensory systems. *J Comput Neurosci*, 15:5–17, 2003.
13. A. Gelperin. Oscillatory dynamics and information processing in olfactory systems. *Exp Biol*, 202:1855–1864, 1999.
14. W. Gerstner and W. M. Kistler. *Spiking Neuron Models*. Cambridge University Press, 2002.
15. W. Gerstner, R. R. Ritz, and J. L. van Hemmen. Why spikes? Hebbian learning and retrieval of time-resolved excitation patterns. *Biol Cyber*, 69:503–515, 1993.
16. M. Heisenberg, A. Borst, S. Wagner, and D. Byers. *Drosophila* mushroom body mutants are deficient in olfactory learning. *J Neurogenet*, 2:1–30, 1985.
17. O. Hendin, D. Horn, and M. V. Tsodyks. Associative memory and segmentation in an oscillatory neural model of the olfactory bulb. *J Comput Neurosci*, 5:157–69, 1998.
18. J. S. Hosler, K. L. Buxton, and B. H. Smith. Impairment of olfactory discrimination by blockade of GABA and nitric oxide activity in the honeybee antennal lobes. *Behav Neurosci*, 114:514–525, 2000.
19. R. Huerta, T. Nowotny, M. Garcia-Sanchez, H. D. I. Abarbanel, and M. I. Rabinovich. Learning classification in the olfactory system of insects. *Neural Comput*, 16:1601–1640, 2004.
20. H. Ikeno and S. Usui. Mathematical description of ionic currents of the kenyon cell in the mushroom body of honeybee. *Neurocomputing*, 26-27:177–184, 1999.

21. J. Joerges, A. Küttner, C. G. Galizia, and R. Menzel. Representations of odour mixtures visualized in the honeybee brain. *Nature*, 387:285–288, 1997.
22. J. Kauer. Response patterns of amphibian olfactory bulb neurons to odor stimulation. *J Physiol Lond*, 243:695–715, 1974.
23. T. Komiyama, W. A. Johnson, L. Luo, and G. S. X. E. Jefferis. From lineage to wiring specificity: POU domain transcription factors control precise connections of *drosophila* olfactory projection neurons. *Cell*, 112:157–167, 2003.
24. G. Laurent, M. Stopfer, R. W. Friedrich, M. I. Rabinovich, and H. D. I. Abarbanel. Odor encoding as an active, dynamical process: Experiments, computation, and theory. *Annu Rev Neurosci*, 24:263–297, 2001.
25. G. Laurent, M. Wehr, and H. Davidowitz. Temporal representations of odors in an olfactory network. *J Neurosci*, 16:3837–3847, 1996.
26. Z. Li and J. Hertz. Odour recognition and segmentation by a model olfactory bulb and cortex. *Network: Comput Nural Syst*, 11:83–102, 2000.
27. W. Maas and C. Bishop. *Pulsed Neural Networks*. MIT Press, 1999.
28. R. Malinov and J. P. Miller. Postsynaptic hyperpolarization during conditioning reversibly blocks induction of long-term potentiation. *Nature*, 320:529–530, 1986.
29. E. C. Marin, G. S. Jefferis, T. Komiyama, H. Zhu, and L. Luo. Representation of the glomerular olfactory map in the *Drosophila* brain. *Cell*, 109:243–255, 2002.
30. H. Markram, J. Lübke, M. Frotscher, and B. Sakmann. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275:213–215, 1997.
31. P. Mombaerts. How smell develops. *Nat Neurosci*, 4:1192–1198, 2001.
32. T. Nowotny, M. I. Rabinovich, R. Huerta, and H. D. I. Abarbanel. Decoding temporal information through slow lateral excitation in the olfactory system of insects. *J Comput Neurosci*, 15:271–281, 2003.
33. R. C. O'Reilly and J. L. McClelland. Hippocampal conjunctive encoding, storage, and recall, avoiding a tradeoff. *Hippocampus*, 4:661–682, 1994.
34. C. Pelz, J. Jander, H. Rosenboom, M. Hammer, and R. Menzel. I_A in Kenyon cells of the mushroom body of honeybees resembles shaker currents: Kinetics, modulation by K^+, and simulation. *J Neurophysiol*, 81:1749–1759, 1999.
35. J. Perez-Orive, O. Mazor, G. C. Turner, S. Cassenaer, R. I. Wilson, and G. Laurent. Oscillations and sparsening of odor representations in the mushroom body. *Science*, 297:359–365, 2002.
36. G. q. Bi and M. m. Poo. Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J Neurosci*, 18:10464–10472, 1998.
37. G. q. Bi and M. m. Poo. Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu Rev Neurosci*, 24:139–166, 2001.
38. N. F. Rulkov. Modeling of spiking-bursting behavior using two-dimensional map. *Phys Rev E*, 65:041922, 2002.
39. S. Sachse and C. G. Galizia. Role of inhibition for temporal and spatial odor representation in olfactory output neurons: A calcium imaging study. *J Neurophysiol*, 87:1106–1117, 2002.
40. S. Sachse, A. Rappert, and C. G. Galizia. The spatial representation of chemical structures in the antennal lobes of honeybees: steps toward the olfactory code. *Eur J Neurosci*, 11:3970–3982, 1999.
41. F. T. Sommer and T. Wennekers. Associative memory in networks of spiking neurons. *Neural Networks*, 14:825–834, 2001.
42. M. Stopfer, S. Bhagavan, B. Smith, and G. Laurent. Impaired odor discrimination on desynchronization of odor-encoding neural assemblies. *Nature*, 390:70–74, 1997.
43. M. Stopfer, V. Jayaraman, and G. Laurent. Intensity versus identity coding in an olfactory system. *Neuron*, 39:991–1004, 2003.
44. N. K. Tanaka, T. Awasaki, T. Shimada, and K. Ito. Integration of chemosensory pathways in the *Drosophila* second-order olfactory centers. *Curr Biol*, 14:449–457, 2004.
45. R. D. Traub and R. Miles. *Neural Networks of the Hippocampus*. Cambridge University Press, New York, 1991.
46. H. B. Treloar, P. Feinstein, P. Mombaerts, and C. A. Greer. Specificity of glomerular targeting by olfactory sensory axons. *J Neurosci*, 22:2469–2477, 2002.
47. N. Uchida and Z. F. Mainen. Speed and accuracy of olfactory discrimination in the rat. *Nature Neurosci*, 6:1224–1229, 2003.
48. L. B. Vosshall. The molecular logic of olfaction in *Drosophila*. *Chem Senses*, 26:207–213, 2001.
49. Y. Wang, N. J. D. Wright, H. F. Guo, Z. Xie, K. Svoboda, R. Malinow, D. P. Smith, and Y. Zhong. Genetic manipulation of the odor-evoked distributed neural activity in the *Drosophila* mushroom body. *Neuron*, 29:267–276, 2001.
50. M. Wehr and G. Laurent. Odor encoding by temporal sequences of firing in oscillating neural assemblies. *Nature*, 384:162–166, 1996.
51. J. White, T. A. Dickinson, D. R. Walt, and J. S. Kauer. An olfactory neuronal network for vapor recognition in an artificial nose. *Biol Cyber*, 78:245–251, 1998.
52. A. Whitehead, M. I. Rabinovich, R. Huerta, V. P. Zhigulin, and H. D. I.Abarbanel. Dynamical synaptic plasticity: A model and connection to some experiments. *Biol Cyber*, 88:229–235, 2003.
53. D. A. Wilson. Rapid, experience-induced enhancement in odorant discrimination by anterior piriform cortex neurons. *J Neurophysiol*, 90:65–72, 2003.
54. D. G. Wüstenberg, M. Boytcheva, b. Grünewald, J. H. Byrne, R. Menzel, and D. A. Baxter. Current- and voltage-clamp recordings and computer simulations of Kenyon cells in the honeybee. *J Neurophysiol*, 92:2589–2603, 2004.
55. H. Zhu and L. Luo. Diverse functions of n-cadherin in dendritic and axonal terminal arborization of olfactory projection neurons. *Neuron*, 42:63–75, 2004.