Regulation of specialists and generalists by neural variability improves pattern recognition performance

Aaron Montero[a], Ramon Huerta[a,b], Francisco B. Rodriguez[a]

[a]*Grupo de Neurocomputación Biológica, Dpto. de Ingeniería Informática. Escuela Politécnica Superior. Universidad Autónoma de Madrid, 28049 Madrid, Spain.*
[b]*BioCircuits Institute, University of California, San Diego, La Jolla, CA 92093-0402, USA.*

Abstract

To analyze the impact of neural threshold variability in the mushroom body (MB) for pattern recognition, we used a computational model based on the olfactory system of insects. This model is a single-hidden-layer neural network (SLN) where the input layer represents the antennal lobe (AL). The remaining layers are in the MBs that are formed by the Kenyon cell (KC) layer and the output neurons that are responsible for odor learning. The binary code obtained for each odorant in the output layer by unsupervised learning was used to measure the classification error. This classification error allows us to identify the neural variability paradigm that achieves a better odor classification. The neural variability is provided by the neural threshold of activation. We compare two hypotheses: a unique threshold for all the neurons in the MB layer, which leads to no variability (homogeneity), and different thresholds for each MB layer (heterogeneity). The results show that, when there is threshold variability, odor classification performance improves. Neural variability induces populations of neurons that are specialists and generalists. Specialist neurons respond to fewer stimulus than the generalists. The proper combination of these two neuron types leads to performance improvement in the bioinspired classifier.

Keywords: Artificial neural networks, neural threshold, neural variability, heterogeneity, homogeneity, olfactory system, pattern recognition, generalist neuron, specialist neuron, odor learning, odor classification, odor discrimination, gain control, unsupervised learning, neural sparseness condition.

1. Introduction

The olfactory system of insects is made of a complex neural machinery made of at least four processing layers [1] capable of classifying a large number of odorants from an unlimited number of stimuli that are highly variable [2] (different gas concentrations, mixtures, etc). The main reasons to chose the olfactory system of insects are: the simplicity of the structural organization [3, 4, 5, 6, 7, 8, 9, 10], the nature of the neural coding [11, 12, 13, 14, 15, 16, 17, 2, 18, 19], the advent of the genetic manipulation techniques that isolate brain areas during the formation of memories [20, 21, 22, 23], and the extensive odor conditioning experiments that shed light into the dynamics of learning during discrimination tasks [24, 25, 26, 27, 28, 29]. Olfactory systems implement simple mechanisms to realize a quick and stable odorant discrimination [30], a goal we want to achieve through computer modeling. Our focus in this work is on neural variability. The driving question is how neural heterogeneity impacts system performance in pattern recognition.

Neural diversity is widespread in the brain, even within the same neural types there is a large heterogeneity in the intrinsic properties and the connectivity patterns, one hypothesis that explains this puzzling observation is functional differentiation within the same types [31]. Another explication is the hypothesis of homeostatic regulation of neural systems, in particular in the olfactory system [32, 33, 34, 35]. However, as we show in this paper, neural heterogeneity can be very beneficial in terms of improving performance in pattern recognition tasks.

Typical models of the olfactory system use very little variability in the excitability in the neurons, implemented by fixed neural thresholds. However, recent applied research on artificial noses determined that using heterogeneous detection thresholds for different odorants, you can improve gas discrimination [36, 37]. This is one of the motivations why we study neuron threshold variability in the information process achieved by neural olfactory system. Additionally, it has been reported that neural thresholds vary in olfactory receptor neurons (ORN) [38] and in Kenyon cells (KCs) [39]. Neural variability in the form of a broad distribution of thresholds is a generic property of neurons in the brain.

To investigate if neural threshold variability increases odorant classification performance, we use a simple model of the olfactory system [40, 41] based on McCullouch-Pitts neurons [42]. The insect olfactory pathway starts at the antenna, where a massive number of receptors encoding the odor stimulus in a high-dimensional space. This information is then sent to the AL for additional processing. The AL exhibits complex dynamics produced by the interaction of its excitatory and inhibitory neural populations [43, 44, 13]. The excitatory cells

Email address: aaron.montero@uam.es (Aaron Montero)

are called projection neurons, PNs, because they only transmit the result of AL computation to deeper regions. Moreover, recordings from the AL in the locust indicate that the activity in the projections of the excitatory neurons of the Locust remains nearly constant despite large variations of the odor concentration [45]. Therefore, a gain control mechanism [46, 47] controlling neuronal activity in the AL is likely to exist [48]. The projection neurons deliver the AL output to a very large number cells of Kennyon of the MB using a fan-out connectivity that increases the separability between different odor encodings. This fan-out phase combined with the sparse firing for these KCs [39, 49, 50] facilitates the odorant discrimination process realized in a fan-in phase by output neurons, which are involved in memory formation and storage [51, 52, 20].

We focus on the AL and MB (model in Fig.1), where the input to single-hidden-layer neural network (SLN) is the AL activity, which is connected to MB through a non-specific connectivity matrix [50]. The reason for this non-specific connectivity matrix is due to the individual connection variability of insects of the same species [53, 54]. The other layers of the SLN, hidden and output, are composed by KC and output neurons respectively. These are connected by a connectivity matrix that implements Hebbian-like learning [52, 55].

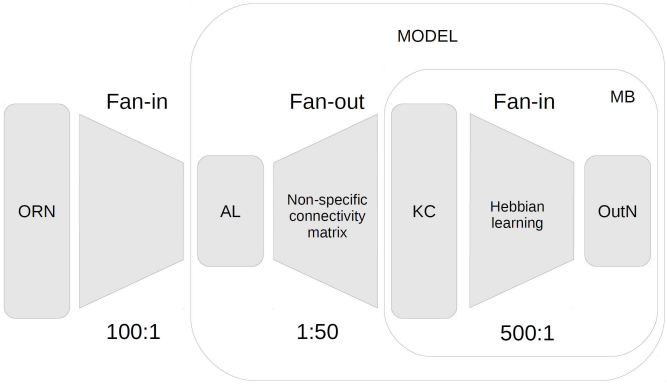

Figure 1: The structure of the model is divided into the antennal lobe (AL) and mushroom body (MB). The MB is divided into the Kenyon cell (KC) layer and output neurons (OutN). The ratios shown are taken from the locust brain size in this particular case.

Our goal is to analyze, first, how information is processed in the olfactory system and, second, the role of threshold variability in this system. Hence, we compare the existence of threshold variability (heterogeneous thresholds) with their absence (homogeneous threshold) to determine whether this improves odorant classification. To this end, we measure the classification error obtained in the output layer after applying unsupervised learning. A correctly classified odorant always generates the same output pattern class A' for a given input pattern class A.

We conclude that odorant classification can improve with neuron threshold variability or heterogeneity, leading us to label neurons as generalists or specialists [56, 57]. Moreover, the classification performance is closely related to sparse activity of the KC population [39, 58] which can be regulated by neural thresholds too in addition to the connectivity degrees [50].

2. Olfactory model

2.1. Neuron model

In locusts, activity patterns in the AL are practically time-discretized by a periodic feedforward inhibition onto the MB calyxes [59] with very low KC activity [39]. Thus, the information is represented by time-discrete, sparse activity patterns with the KCs locked on the 50 ms local field potential oscillation cycle. Because of these neurons are inactive most of the time, but being activated, their neuronal response is produced by the coincidence of concurrent spikes followed by a reset, we have used the McCulloch-Pitts model [42] in all neurons of the hidden and output layers, as mentioned above. This neuron model uses the threshold step function as activation function. Therefore, we have the following (see network model in Fig.2):

$$\begin{aligned} y_j &= \varphi\Big(\sum_{i=1}^{N_{AL}} c_{ji}x_i - \theta_j\Big), \quad j = 1,\ldots,N_{KC}, \\ z_l &= \varphi\Big(\sum_{j=1}^{N_{KC}} w_{lj}y_j - \varepsilon_l\Big), \quad l = 1,\ldots,N_{OutN}, \end{aligned} \tag{1}$$

where x_i, y_j and z_l are activation states for a input, hidden and output neuron respectively, c_{ji} and w_{lj} are weights linking two neurons, θ_j and ε_l are thresholds for the hidden and output neuron respectively, and φ is the Heaviside activation function. The Heaviside activation function φ is 0 when its argument is negative or 0 and 1 otherwise.

2.2. Network model

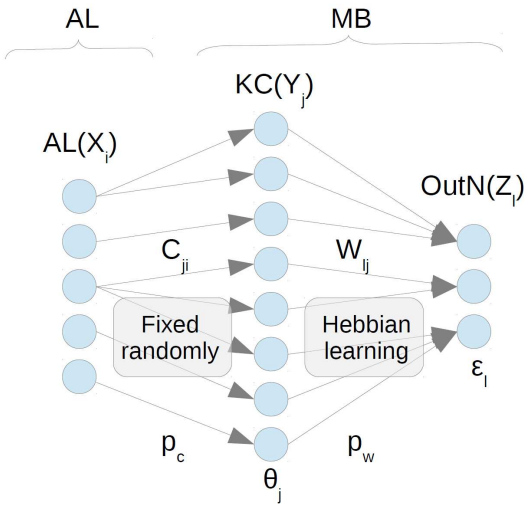

Figure 2: Network model composed of 3 layers: Antennal Lobe, Kenyon cells and output neurons. The input layer X is connected to the hidden layer Y by a random matrix C. This hidden layer is connected to the output layer Z by other random matrix W, whose weights are updated by Hebbian learning. These random matrices are created with connection probabilities p_c and p_w. The thresholds, biases, of hidden neurons and output neurons are θ and ε respectively.

The network model is a SLN (Fig.2) with an input layer of 50 neurons, a hidden layer with $2,500$ neurons (locust has a ratio of 1:50 between neurons of the input and hidden layer) and an output layer with 5 neurons [41] (Table 1). These dimensions were chosen because they ensure a high probability of classification for the input used [40] for a relatively low computational cost.

The connectivity matrices, C and W, are initialized at the beginning of each learning process. We generate a matrix with random values uniformly distributed in the range $[0, 1]$. The binary connection values in the connectivity matrix use p_c and p_w, as a threshold on the values of the random matrix such that if an entry value is equal or less than p_c or p_w, the connection is established otherwise is set to 0. The connectivity matrix C remains fixed throughout the learning process, while the connectivity matrix W is updated using Hebbian learning. The synaptic model of this network is completely binary. Therefore, activation states for a neuron and weights can only take values 0 or 1, except for the input X.

2.3. Hebbian learning

As mentioned above, the connectivity matrix W linking the KCs and output neurons undergoes associative learning, which can be simulated by using Hebbian learning [52, 55]. This learning is subjected to certain thresholds, whose selection will be detailed in an upcoming section. Hebbian learning allows the strengthening or weakening of the connections given by the following connectivity matrix [40, 41]:

$$w_{lj}(t+1) = H(z_l, y_j, w_{lj}(t)),$$

$$H(1,1,w_{lj}(t)) = \begin{cases} P(H(1,1,w_{lj}(t)) = 1) = p_+ \\ P(H(1,1,w_{lj}(t)) = w_{lj}(t)) = 1 - p_+ \end{cases}$$
$$H(1,0,w_{lj}(t)) = \begin{cases} P(H(1,0,w_{lj}(t)) = 0) = p_- \\ P(H(1,0,w_{lj}(t)) = w_{lj}(t)) = 1 - p_- \end{cases} \quad (2)$$

$$H(0,1,w_{lj}(t)) = w_{lj}(t), \quad H(0,0,w_{lj}(t)) = w_{lj}(t),$$

where the future connection state $w_{lj}(t+1)$ is determined by a function $H(z_l, y_j, w_{lj}(t))$, which depends on the output layer neuron z_l, the hidden layer neuron y_j and the current connection state $w_{lj}(t)$. If the output layer neuron has not fired, the connection state is not changed. However, if the output layer neuron has fired, the connection state depends on the hidden layer in the following ways:

- If the hidden layer neuron has fired, then the connection between these neurons is created with a probability p_+.
- If the hidden layer neuron has not fired, then the connection between these neurons is destroyed with a probability p_-.

2.4. Classification error

The classification error represents the percentage of odorants which have not been correctly classified. To calculate this percentage, we assume that a correctly classified odorant always generates the same output pattern class A' for a certain input pattern class A. Therefore, since we know how many clusters there are in the input, we expect the same number of clusters to appear in the output. This is the expected clustering used to measure the classification error which is calculated after obtaining the model output. The output is clustered and compared with the original ground truth, all of those odorants that are inconsistent with this clustering is our classification error:

$$error = \frac{\#(C_{in} \setminus C_{out})}{\#C_{in}} \times 100 \quad (3)$$

where *error* is the number of elements in the algebraic set difference (symbol "\" in the equation 3) of the input clustering C_{in} with the output clustering C_{out}, divided by the number of elements in the input clustering $\#C_{in}$.

3. Odorants

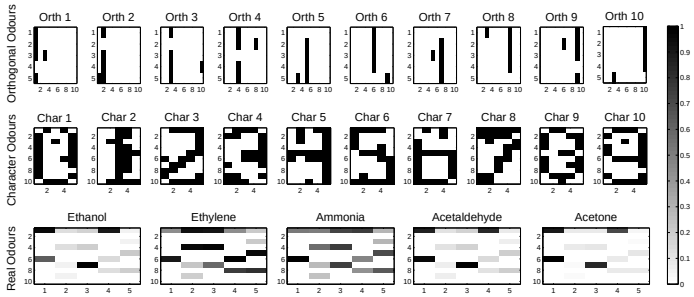

Figure 3: Examples of orthogonal, character, normalized real patterns (in order from top to bottom). Colors: black (fully active neuron), white (inactive neuron), gray (different degrees of activity).

We use odorant patterns with different complexity degrees (Fig.3). The simplest odorants are orthogonal with no overlapping activity. Character odorants have overlapping and contain binary information (input neurons are active or inactive 0,1). Finally, the more complex odorants are extracted from data provided by electronic noses. These are real numbers (input neurons have different degrees of activation) and we refer to them as real odorants. We use 100 patterns of each kind of odorant. This model is robust to noise [60]. Therefore, to get 100 patterns of orthogonal and character odorants, we replicate the 10 pattern classes of these odorants adding some noise. Real patterns are composed by recordings of 16 chemical sensors [61, 62] for 5 pattern classes: ethanol, ethylene, ammonia, acetaldehyde and acetone. Each pattern is subjected to gain control [48, 47], that keeps a constant neuronal activity for all of them. These sets of patterns are evaluated by 5-cross-validation, such that the training data set has 80 patterns and test set has 20 patterns (Table 2).

Parameter	Value	Description
N_{AL}	50	Number of neurons for the input layer
N_{KC}	2500	Number of neurons for the hidden layer
N_{OutN}	5	Number of neurons for the output layer
x_i	$\{0,1\},[0,100] \in \mathbb{R}$	Input neuron i
y_j	$\{0,1\}$	Hidden neuron j
z_l	$\{0,1\}$	Output neuron l
c_{ij}	$\{0,1\}$	Weight between a input neuron i and a hidden neuron j
w_{jl}	$\{0,1\}$	Weight between a hidden neuron j and an output neuron l
θ_j	$[0,+\infty] \in \mathbb{R}$	Threshold for a hidden neuron j
ε_l	$[0,+\infty] \in \mathbb{R}$	Threshold for an output neuron l
p_c	$[0,1] \rightarrow \{0.1,0.3,0.5\}$	Connection probability between an input neuron i and a hidden neuron j
p_w	$[0,1] \rightarrow 0.5$	Connection probability between a hidden neuron j and an output neuron l
p_+	$[0,1] \rightarrow 0.2$	Probability of strengthening of weight w_{jl}
p_-	$[0,1] \rightarrow 0.1$	Probability of weakening of weight w_{jl}

Table 1: Table of parameters and their values in the model. These parameters, in order of appearance, correspond to dimensions, neurons, weights and thresholds of the network, and probabilities used in the model.

Odorants	Orthogonal	Character	Real
Classes	10	10	5
Patterns for each class	10	10	20
Patterns	100	100	100
Training	80	80	80
Test	20	20	20

Table 2: Table of odorants. This table shows, for each kind of odorant, the number of pattern classes and number of patterns for each class. Also, the total number of patterns and those used for training and test sets.

4. Threshold selection

Every neuron is capable of firing and transmitting information to the next neuron when it exceeds a certain threshold. To select the value of these thresholds, we use the concept of limit threshold. A limit threshold is the number of stimuli received in a neuron for a given odorant. This is the minimum threshold which prevents a neuron from spiking for a given odorant.

To compare the use of a unique threshold for all neurons in the MB layer (homogeneous thresholds) with the use of different thresholds for each neuron of the MB layer (heterogeneous thresholds), we use these limit thresholds in the training stage. In the case of homogeneous thresholds, the minimum and maximum values for these limit thresholds determine the range of selected thresholds. For heterogeneous thresholds, the variance of limit thresholds establishes the populations of generalist and specialist neurons. The threshold selection method varies according to the type of neuron.

4.1. Limit thresholds

A limit threshold is the minimum threshold value which prevents a neuron from spiking for a given odorant. This value has been used as threshold in order to prove how important threshold variability is for classification tasks. This limit threshold is calculated for each neuron and each odorant as follows:

$$\theta_j^O = \sum_{i=1}^{N_{AL}} c_{ji} x_i^O, \qquad \varepsilon_l^O = \sum_{j=1}^{N_{KC}} w_{lj} y_j^O \tag{4}$$

where neuron j spikes $\forall \theta_j, 0 \leq \theta_j < \theta_j^O$, and neuron l spikes $\forall \varepsilon_l, 0 \leq \varepsilon_l < \varepsilon_l^O$. Being θ_j^O the limit threshold for a KC (j) and an odorant (O), and ε_l^O the limit threshold for an output neuron (l) and an odorant (O). These thresholds are calculated only one time in the odorant classification process before Hebbian learning is applied. Therefore, the limit threshold matrix stores all limit thresholds of a layer. In case of hidden layer, it has dimension $N_{KC} \times N_{odor}$, and $N_{OutN} \times N_{odor}$ for the KC and output layer.

$$\Theta = \begin{pmatrix} \theta_1^1 & \cdots & \theta_1^{N_{odor}} \\ \vdots & \ddots & \vdots \\ \theta_{N_{KC}}^1 & \cdots & \theta_{N_{KC}}^{N_{odor}} \end{pmatrix} E = \begin{pmatrix} \varepsilon_1^1 & \cdots & \varepsilon_1^{N_{odor}} \\ \vdots & \ddots & \vdots \\ \varepsilon_{N_{OutN}}^1 & \cdots & \varepsilon_{N_{OutN}}^{N_{odor}} \end{pmatrix} \tag{5}$$

The purpose of these matrices is to know all possible thresholds for each layer, for all neurons and odorants, and choose the one which improves odorant classification.

4.2. Homogeneous thresholds

To calculate the threshold of each layer, we set the network weights, C and W, and obtain the limit threshold matrix of the hidden layer. We take the minimum and maximum of this matrix and use these values, and those among them, as thresholds for the hidden layer (Algorithm.1, lines 3-5). The goal is to obtain the minimum classification error for each threshold and the spike rate for the minimum value.

In order to achieve this minimum classification error, we obtain the limit threshold matrix for the output layer, of each hidden layer threshold, and we take the minimum and maximum of this matrix. Moreover we modify the weights of W by Hebbian learning. We calculate the classification error for all possible combinations and take the minimum observed. This value is

the minimum classification error for a hidden layer threshold (Algorithm.1, lines 9-18).

Algorithm 1 Homogeneous thresholds

1: Set C and W in function of p_c and p_w
2: $\Theta = C^t X$ //Limit threshold matrix for the hidden layer
3: $\theta_{min} = min(\Theta)$ //minimum matrix Θ
4: $\theta_{max} = max(\Theta)$ //maximum matrix Θ
5: $N = \theta_{max} - \theta_{min} + 1$ //number of thresholds
6: $error[N] = 1$ //vector that stores the minimum error for each θ
7: **for** $n = 0 \rightarrow N - 1$ **do**
8: $\theta = \theta_{min} + n$
9: $E = W^t Y$ //Limit threshold matrix for the output layer
10: $\varepsilon_{min} = min(E)$ //minimum matrix E
11: $\varepsilon_{max} = max(E)$ //maximum matrix E
12: $M = \varepsilon_{max} - \varepsilon_{min} + 1$
13: **for** $m = 0 \rightarrow M - 1$ **do**
14: $\varepsilon = \varepsilon_{min} + m$
15: $HebbianLearning(z, y, w)$
16: **if** $error < error(\theta)$ **then**
17: $error(\theta) = error$
18: **end if**
19: **end for**
20: **end for**

4.3. Heterogeneous thresholds

We set the C and W weights and calculate the limit threshold matrix for the hidden layer. Using this matrix, we obtain the distribution of limit thresholds for each hidden layer neuron. Then, neurons are labeled as generalists or specialists [56, 57] depending on the distribution. This labeling determines the subsequent threshold selection. Generalist neurons extract the general properties of odorants and therefore respond to a wide range of them. Specialist neurons are essential for odor discrimination and therefore respond only to a single odorant. Thus, variance of limit thresholds may be a measure that allows us neural labeling as generalist or specialist. Mainly because general neurons should have a low variance of limit threshold, since these neurons are similarly stimulated by all odorants, and specialists neurons should have a greater variance because they are more sensitive to a certain odorant (Fig.4). The reasoning being that if the variance of the threshold limit increases, the greater the distribution of limit thresholds. Then, this greater value allows that the neuron in question can be more specific.

We use different proportions of generalist and specialist neurons to select thresholds. To achieve this we first calculate the variance of limit thresholds for each neuron. Subsequently these values are sorted from lowest to highest and, as we increase the percentage of generalist neurons from 1 to 100, we take a higher percentage of neurons with low variance (Algorithm.2, lines 5-8).

Once the neurons are labeled, the neural thresholds are calculated (Algorithm.2, line 9). Since generalist neurons do not provide sufficient information to discriminate odorants, these neurons cannot fire. Thus, we assigned them to the maximum limit threshold for these neurons. For specialist neurons, we used a specialization coefficient α. This coefficient determines the percentage of odorants that the neuron does not respond to. Since the coefficient may vary from one neuron to another, it is calculated based on the variations of the distribution of limit thresholds. Therefore, we take the previous value as threshold where the largest decrease of odorants by limit threshold is located. This method has its limitations, but in many cases it allows us to achieve the neuron's appropriate specialization (Fig.4).

Algorithm 2 Heterogeneous thresholds

1: Set C and W in function of p_c and p_w
2: $\Theta = C^t X$ //Limit threshold matrix for the hidden layer
3: $N = 100$
4: $error[N] = 1$
5: $var = variance(\Theta)$ //threshold variance for each neuron
6: **for** $n = 1 \rightarrow N$ **do**
7: $p = percentile(var, n)$ //variance value for which a n percentage of neurons is below
8: $label\_neuron(p)$ //it labels using p each neuron as generalist or specialist
9: $\theta = threshold\_selection$ //each threshold is assigned depending on the maximum limit threshold for generalist neurons or α_θ for specialist neurons
10: $E = W^t Y$ //Limit threshold matrix for the output layer
11: $var = variance(E)$
12: **for** $m = 1 \rightarrow N$ **do**
13: $p = percentile(var, m)$ //variance value for which a m percentage of neurons is below
14: $label\_neuron(p)$ //it labels using p each neuron as generalist or specialist
15: $\varepsilon = threshold\_selection$ //each threshold is assigned depending on the maximum limit threshold for generalist neurons or α_ε for specialist neurons
16: $HebbianLearning(z, y, w)$
17: **if** $error < error(n)$ **then**
18: $error(n) = error$
19: **end if**
20: **end for**
21: **end for**

To achieve this minimum classification error, for each percentage used in the hidden layer we take all possible integer percentages and calculate the threshold for each neuron in the output layer. Also we update the weights of W by Hebbian learning. We calculate the classification error for all possible combinations and take the minimum observed. This value is the minimum classification error for a percentage used in the hidden layer (Algorithm.2, lines 10-20).

5. Results

The results are divided into two parts. First, we compare the paradigm that provides better classification results, that use

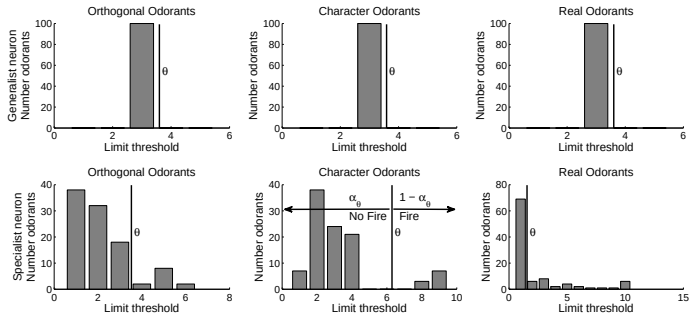

Figure 4: Examples of minimum (generalist neuron) and maximum (specialist neuron) variance of limit thresholds. θ is the neural threshold of a hidden neuron. For a generalist neuron, θ is the maximum limit threshold and, thus, this neuron cannot fire for the training patterns. For a specialist neuron, the neural threshold is taken from a degree of specialization, α_θ where neurons can fire.

homogeneous thresholds (no threshold variability) or heterogeneous ones (threshold variability). We show the results for different sets of odorants and different connection probabilities for the hidden layer, p_c. Finally, we present the results for a particular p_c, which shows the relationship between classification error and spikes rate for different odorants.

5.1. Results on different odorant sets and connection probabilities

We have run 100 simulations for each set of odorants and each p_c probability. We used connection probabilities p_c ranging from 0.1 to 0.5 based on previous studies [50, 63]. In case of probability p_w, we took a single value, $p_w = 0.5$, since W weights are subject to Hebbian learning and therefore the variability of this probability does not involve substantial changes in classification. For the Hebbian learning probabilities, we took as reference the results of another study [41] for the computational model that we used. This paper shows a clear division between successful pairings obeying roughly $p_- \geq 0.2p_+$ and unsuccessful ones for smaller p_- ($p_- < 0.1$). Thus, we decided to take $p_+ = 0.2$ and $p_- = 0.1$ as our Hebbian probabilities, because they let us to perform a good classification. We ran the Hebbian learning for 20 iterations without inserting an output target (unsupervised learning).

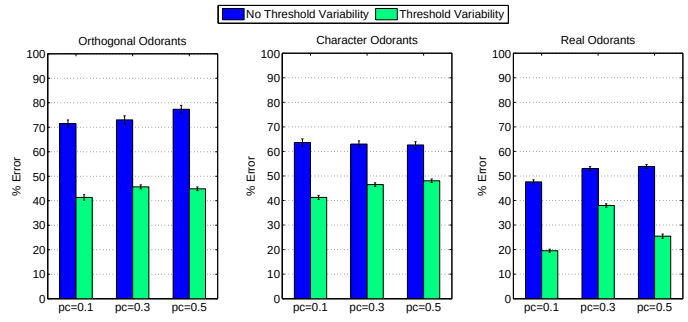

Figure 5: Comparison between the absence and presence of threshold variability for test sets of odorants, for different sets of odorants and connection probabilities. Sample means with 95% confidence intervals of standard errors (error bar calculated over 100 simulations).

These averaged results for the test set (Fig.5) show that threshold variability allows to achieve a lower classification error. Moreover, the classification success is related to low p_c probabilities, specifically $p_c = 0.1$.

The odorants with the best results are the real odorants. The success of these odorants over other simpler as orthogonal odorants may be due to the following reasons. Although a gain control was applied to all odorants, the total neuronal activity of an input pattern (n_{AL}) is different depending on the set of odorants: orthogonal $n_{AL} = 5$, character $n_{AL} = 22$ and real $n_{AL} = 100$. These differences are due to the characteristics of each odorant. Considering that, for real odorants, there are almost no inactive neurons in the input that induces greater variability of limit thresholds. Furthermore, the change of the network dimensionality regarding the previous work [60] meant that the noise introduced in the orthogonal and character odorants have greater negative impact on the results. Finally, in the case of real odorant we have 5 pattern classes instead of 10 that may improve the results.

5.2. Classification error - spike rate

We have taken the averaged results, which we have seen above, and observed the relationship between classification error and spike rate for a particular case of connectivity probability, $p_c = 0.1$.

These results show that minimum classification error in both strategies (homogeneous and heterogeneous thresholds) is related to a low spike rate (Fig.6). Because we can observe in Table 3 that all spikes values are low ($\leq 30\%$). This is consistent with the hypothesis that high population sparseness in the KC layer improves odorant discrimination. The reason for this behavior is that if spike rates are high, neurons cannot be selective when firing and therefore its discrimination ability may worsen. However, if spike rates are too low, neurons can be too selective and the output layer cannot receive enough information to properly classify the input.

	Orthogonal	Character	Real
No Threshold Variability	1%	7%	30%
Threshold Variability	22%	25%	20%

Table 3: Table of the maximum spike rate associated to the minimum classification error observed in Figure 6. These values are the spikes rate for the odorant that stimulate a higher number of neurons and not the average of all odorants.

This rationale is observable when there is no threshold variability (top panels Fig.6), since thresholds go from the lowest to highest limit threshold. However, threshold variability always try to get selective neurons. Since what varies is the percentage of generalist neurons, which do not imply higher or lower thresholds. In this case it is observed that the variation in the percentages of generalist and specialist neurons only have relevance to real odorants that offer a greater range of limit thresholds. Since orthogonal and character odorants can be classified even if the percentage of specialist neurons is reduced, only when they finally disappear the classification error increases to its maximum value. Therefore, these odorants can maintain

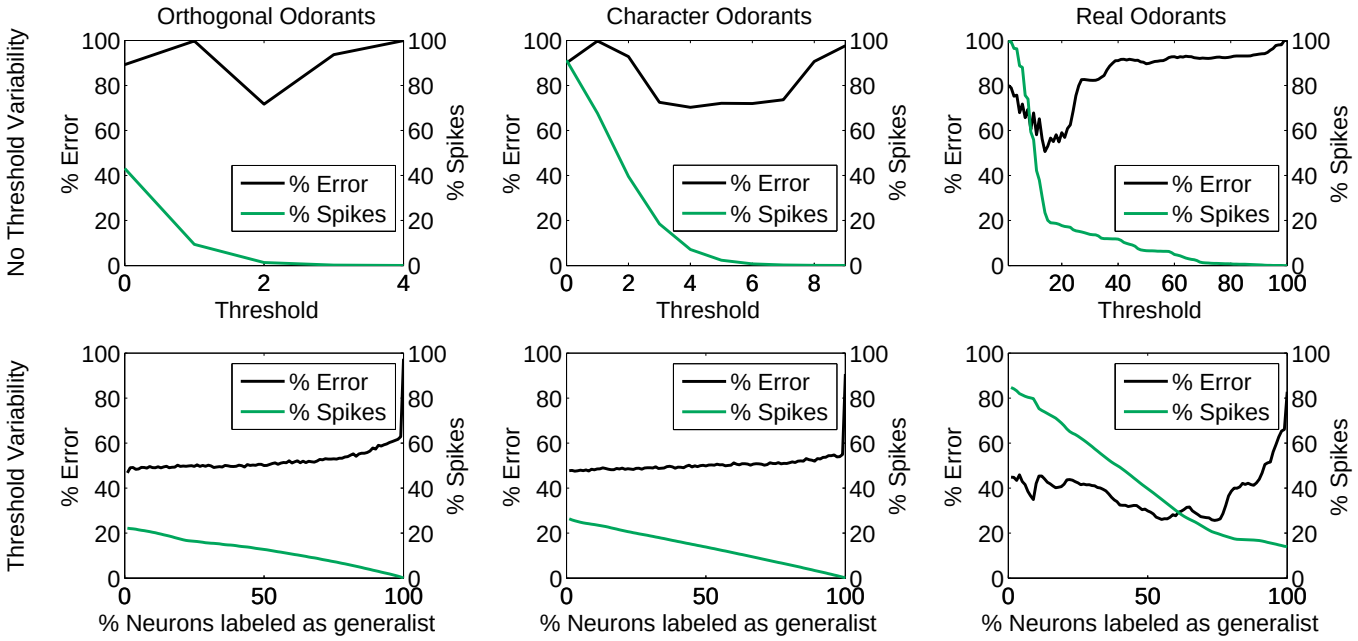

Figure 6: Classification error and maximum spike rate for different kind of odorants ($p_c = 0.1$). The neurons labeled as generalists are those with a lower variance of limit thresholds. These results are the average of 100 simulations for different C and W weights. These results show, for all pattern classes, that classification error is always lower for threshold variability.

uniform classification success even with a 1% of specialist neurons (25 Kenyon cells).

The final threshold distributions that achieve the minimum classification error of the figure (Fig.6), selected by our model, can be seen in figure (Fig.7) for a particular network configuration. These distributions show that generalist neurons have a fixed threshold. These neurons, as we mentioned above, do not respond to the input.

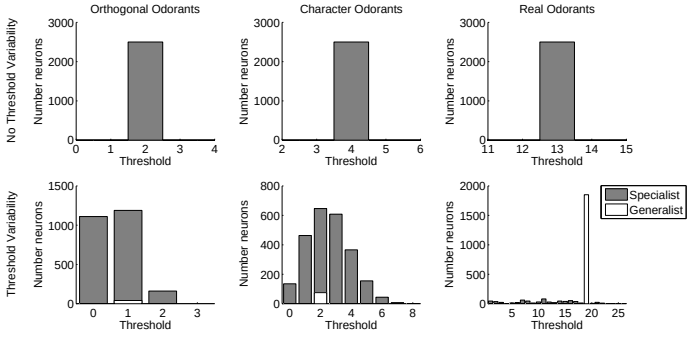

Figure 7: Example of final threshold distribution that minimizes classification error for a particular network configuration (C and W weights) and a connectivity probability $p_c = 0.1$. For threshold variability, percentages of generalist neurons are 1% (orthogonal), 3% (character) and 74% (real) in Fig. 6. Since the error does not minimize for a certain value in case of orthogonal and character odorants, these percentages vary significantly for other network configurations. However, for real odorants where there is a clear minimization, the percentage varies slightly. Generalist neurons have a fixed threshold which corresponds to the maximum limit threshold. We established that generalist neurons do not fire in this problem for the training patterns.

6. Conclusions and discussions

The main objective of this work is to analyze how information is processed by investigating the role of threshold variability. To accomplish this, we used a simple model of the olfactory system based on a connectionist model that uses random connections. Thus the model focuses on the AL and MB, where the input to single-hidden-layer neural network (SLN) is the AL activity, which is connected to MB through a non-specific connectivity matrix. The other layers of the SLN, hidden and output, are composed by KCs and output neurons respectively. These are connected by a connectivity matrix that implements Hebbian learning. Using different kind of odorants: orthogonal, character and artificial sensor array data; we compared threshold variability (heterogeneous thresholds) against the absence (homogeneous threshold) to investigate whether odorant classification is improved. Classification errors measured in the output layer after applying unsupervised learning.

We show that neural threshold variability can indeed improve classification performance as shown in Fig.5 for real odorants from a comprehensive dataset of artificial chemical sensors. The success for these odorants can be related to their greater variability of limit thresholds. Furthermore, classification errors are lower when the connection probability between AL and MB, p_c is low as well, which is consistent with information maximization criteria provided in [50].

Neural threshold variability also allows us to label neurons as generalist or specialist [56, 57] (section 4). Generalist neurons extract the general properties of odorants and therefore respond to a wider range. Specialist neurons are essential for odor discrimination and therefore respond only to a single odorant. To

label neurons as generalists or specialists we use the variance of limit thresholds for the training patterns. Limit threshold is the minimum threshold value which prevents a neuron from spiking for a given odorant. Neurons with low variance are labeled as generalists and the others as specialists. Generalist neurons take the maximum limit threshold as their threshold, therefore these neurons do not fire at least for the training patterns. Specialist neurons select the threshold by a specialization coefficient α (bottom panels Fig.4). In data classification from electronic noses, we can observe a solution with the proper combination of generalist and specialist neurons (Fig.7). The percentage of neurons labeled as generalists for this solution is 74% for connectivity probability $p_c = 0.1$ (Fig.6).

Finally, we proved that success in discrimination is related to the sparse activity documented in the KC layer [39, 58] (Fig.6). The maximum spike rate associated to the minimum classification error is 30% in the worst case (Table 3). If spike rates are high, neurons cannot be selective when firing and therefore its discrimination ability worsens. However, if spike rates are too low, neurons can be too selective and the output layer cannot receive enough information to properly generalize the input.

Acknowledgments

This work was supported by the Spanish Government project TIN2010-19607 and predoctoral research grant BES-2011-049274. R.H. acknowledges partial support by NIDCD-R01DC011422-01.

References

[1] M. Heisenberg, Mushroom body memoir: from maps to models., Nat Rev Neurosci 4 (2003) 266–275.
[2] C. G. Galizia, R. Menzel, Probing the olfactory code, Nature Neurosci 3 (2000) 853–854.
[3] C. G. Galizia, S. L. McIlwrath, R. Menzel, A digital 3D atlas of the honeybee antennal lobe based on optical sections acquired using confocal micoscropy, Cell Tissue Res 295 (1999) 383–394.
[4] M. Mizunami, J. M. Weibrecht, N. J. Strausfeld, Mushroom bodies of the cockroach: their participation in place memory, J Comp Neurol 402 (1998) 520–537.
[5] K. Ito, K. Suzuki, P. Estes, M. Ramaswami, D. Yamamoto, N. J. Strausfeld, The organization of extrinsic neurons and their implications in the functional roles of the mushroom bodies in *Drosophila melanogaster Meigen*, Learn Mem 5 (1998) 52–77.
[6] S. M. Farris, N. J. Strausfeld, Development of laminar organization in the mushroom bodies of the cockroach: Kenyon cell proliferation, outgrowth, and maturation, J Comp Neurol 439 (2001) 331–351.
[7] J. D. Armstrong, K. Kaiser, A. Müller, K.-F. Fischbach, N. Merchant, N. J. Strausfeld, Flybrain, an on-line atlas and database for the drosophila nervous system, Neuron 15 (1995) 17–20.
[8] N. J. Strausfeld, Organization of the honey bee mushroom body: Representation of the calyx within the vertical and gamma lobes, J Comp Neurol 450 (2002) 4–33.
[9] C. G. Galizia, B. Kimmerle, Physiological and morphological characterization of honeybee olfactory neurons combining electrophysiology, calcium imaging and confocal microscopy, J Comp Physiol A 190 (2004) 21–38.
[10] J. G. Hildebrand, G. M. Shepherd, Mechanisms of olfactory discrimination: Converging evidence for common principles across phyla, Annu Rev Neurosci 20 (1997) 595–631.
[11] R. F. Galan, M. Weidert, R. Menzel, A. V. M. Herz, C. G. Galizia, Sensory memory for odors is encoded in spontaneous correlated activity between olfactory glomeruli, Neural Comput 18 (2006) 10–25.
[12] G. Laurent, A systems perspective on early olfactory coding, Science 286 (1999) 723–728.
[13] G. Laurent, Olfactory network dynamics and the coding of multidimensional signals, Nature Rev Neurosci 3 (2002) 884–895.
[14] G. Laurent, Dynamical representation of odors by oscillating and evolving neural assemblies, Trends Neurosci 19 (1996) 489–496.
[15] T. Faber, J. Joerges, R. Menzel, Associative learning modifies neural representations of odors in the insect brain, Nature Neurosci 2 (1999) 74–78.
[16] R. Abel, J. Rybak, R. Menzel, Structure and response patterns of olfactory interneurons in the honeybee *Apis mellifera*, J Comp Neurol 437 (2001) 363–383.
[17] C. G. Galizia, J. Joerges, A. Küttner, T. Faber, R. Menzel, A semi-in-vivo preparation for optical recording of the insect brain, J Neurosci Meth 76 (1997) 61–69.
[18] N. J. Vickers, T. A. Christensen, T. C. Baker, J. G. Hildebrand, Odour-plume dynamics influence the brain's olfactory code., Nature 410 (2001) 466–470.
[19] K. C. Daly, R. F. Galán, O. J. Peters, E. M. Staudacher, Detailed characterization of local field potential oscillations and their relationship to spike timing in the antennal lobe of the moth manduca sexta., Frontiers in Neuroengineering 4 (2011).
[20] T. Zars, M. Fischer, R. Schulz, M. Heisenberg, Localization of a short-term memory in drosophila., Science 288 (2000) 672–675.
[21] T. Zars, R. Wolf, R. Davis, M. Heisenberg, Tissue-specific expression of a type I adenylyl cyclase rescues the *rutabaga* mutant memory defect: In search of the engram, Learn Mem 7 (2000) 18–31.
[22] T. Zars, Behavioral functions of the insect mushroom bodies., Curr Opin Neurobiol 10 (2000) 790–795.
[23] Y. Wang, N. J. D. Wright, H. F. Guo, Z. Xie, K. Svoboda, R. Malinow, D. P. Smith, Y. Zhong, Genetic manipulation of the odor-evoked distributed neural activity in the *Drosophila* mushroom body, Neuron 29 (2001) 267–276.
[24] M. E. Bitterman, R. Menzel, A. Fietz, S. Schäfer, Classical conditioning of proboscis extension in honeybees (apis mellifera)., J Comp Psychol 97 (1983) 107–119.
[25] B. H. Smith, G. A. Wright, K. C. Daly, Learning-based recognition and discrimination of floral odors, in: N. Dudareva, E. Pichersky (Eds.), Biology of Floral Scent, CRC Press, 2005, pp. 263–295.
[26] J. S. Hosler, K. L. Buxton, B. H. Smith, Impairment of olfactory discrimination by blockade of GABA and nitric oxide activity in the honeybee antennal lobes, Behav Neurosci 114 (2000) 514–525.
[27] B. H. Smith, An analysis of blocking in binary odorant mixtures: An increase but not a decrease in intensity of reinforcement produces unblocking, Behav Neurosci 111 (1997) 1–13.
[28] B. H. Smith, C. I. Abramson, T. R. Tobin, Conditional withholding of proboscis extension in honeybees (apis mellifera) during discriminative punishment., J Comp Psychol 105 (1991) 345–356.
[29] C. R. Gallistel, S. Fairhurst, P. Balsam, The learning curve: implications of a quantitative analysis., Proc Natl Acad Sci U S A 101 (2004) 13124–13131.
[30] R. Huerta, Learning pattern recognition and decision making in the insect brain, AIP Conference Proceedings 1510 (2013) 101.
[31] E. Krook-Magnuson, C. Varga, S. Lee, I. Soltesz, New dimensions of interneuronal specialization unmasked by principal cell heterogeneity, Trends Neurosci. 35 (2012) 175–184.
[32] G. W. Davis, I. Bezprozvanny, Maintaining the stability of neural function: a homeostatic hypothesis, Annu Rev Physiol. 63 (2001) 847–69.
[33] G. W. Davis, Homeostatic signaling and the stabilization of neural function, Neuron 8 (2013) 718–728.
[34] G. Turrigiano, S. B. Nelson, Homeostatic plasticity in the developing nervous system, Nature Reviews Neuroscience 5 (2004) 97–107.
[35] F. B. Rodríguez, R. Huerta, M. de la Luz Aylwin, Neural sensitivity to odorants in deprived and normal olfactory bulbs, PLoS ONE 8 (2013).
[36] B. Doleman, N. Lewis, Comparison of odour detection thresholds and odour discriminablities of a conducting polymer composite electronic nose versus mammalian olfaction, Sensors and Actuators B 72 (2001) 41–50.
[37] J. Fonollosa, A. Vergara, R. Huerta, Algorithmic mitigation of sensor failure: Is sensor replacement really necessary?, Sensors and Actuators B: Chemical (2013).

[38] A. M. Angioy, A. Desogus, I. T. Barbarossa, P. Anderson, B. S. Hansson, Extreme sensitivity in an olfactory system., Chem Senses 28 (2003) 279–284.
[39] J. Perez-Orive, O. Mazor, G. C. Turner, S. Cassenaer, R. I. Wilson, G. Laurent, Oscillations and sparsening of odor representations in the mushroom body., Science 297 (2002) 359–365.
[40] R. Huerta, T. Nowotny, M. Garcia-Sanchez, H. D. I. Abarbanel, M. I. Rabinovich, Learning classification in the olfactory system of insects, Neural Comput 16 (2004) 1601–1640.
[41] R. Huerta, T. Nowotny, Fast and robust learning by reinforcement signals: Explorations in the insect brain, Neural Comput. 21 (2009) 2123–2151.
[42] W. S. McCulloch, W. Pitts, Logical calculus of ideas immanent in nervous activity, B Math Biophys 5 (1943) 115–133.
[43] K. Daly, G. Wright, B. Smith, Molecular features of odorants systematically influence slow temporal responses across clusters of coordinated antennal lobe units in the moth manduca sexta, J Neurophysiol 28 (2004).
[44] R. F. Galán, S. Sachse, C. G. Galizia, A. V. M. Herz, Odor-driven attractor dynamics in the antennal lobe allow for simple and rapid olfactory pattern classification, Neural Comput 16 (2004) 999–1012.
[45] M. Stopfer, V. Jayaraman, G. Laurent, Intensity versus identity coding in an olfactory system, Neuron 39 (2003) 991–1004.
[46] E. Serrano, T. Nowotny, R. Levi, B. H. Smith, R. Huerta, Gain control network conditions in early sensory coding, PLoS computational biology 9 (2013).
[47] E. Salinas, P. Thier, Gain modulation: A major computational principle of the central nervous system, Neuron. 27 (2000) 15–21.
[48] S. Olsen, R. Wilson, Lateral presynaptic inhibition mediates gain control in an olfactory circuit, Nature 452 (2008) 956–960.
[49] G. Turner, M. Bazhenov, G. Laurent, Olfactory representations by drosophila mushroom body neurons, J Neurophysiol 99 (2008) 734–746.
[50] M. Garcia-Sanchez, R. Huerta, Design parameters of the fan-out phase of sensory systems, J Comput Neurosci 15 (2003) 5–17.
[51] S. E. McGuire, P. T. Le, R. L. Davis, The role of drosophila mushroom body signaling in olfactory memory., Science 293 (2001) 1330–1333.
[52] J. Dubnau, L. Grady, T. Kitamoto, T. Tully, Disruption of neurotransmission in drosophila mushroom body blocks retrieval but not acquisition of memory., Nature 411 (2001) 476–480.
[53] E. C. Marin, G. S. Jefferis, T. Komiyama, H. Zhu, L. Luo, Representation of the glomerular olfactory map in the *Drosophila* brain, Cell 109 (2002) 243–255.
[54] N. K. Tanaka, T. Awasaki, T. Shimada, K. Ito, Integration of chemosensory pathways in the *Drosophila* second-order olfactory centers, Curr Biol 14 (2004) 449–457.
[55] M. Bazhenov, R. Huerta, B. H. Smith, A computational framework for understanding decision making through integration of basic learning rules, The Journal of Neuroscience 33 (2013) 5686–5697.
[56] F. B. Rodríguez, R. Huerta, Techniques for temporal detection of neural sensitivity to external stimulation., Biol Cybern 100 (2009) 289–297.
[57] R. I. Wilson, G. C. Turner, G. Laurent, Transformation of olfactory representations in the drosophila antennal lobe., Science 303 (2004) 366–370.
[58] P. Szyszka, M. Ditzen, A. Galkin, C. G. Galizia, R. Menzel, Sparsening and temporal sharpening of olfactory representations in the honeybee mushroom bodies, J Neurophysiol 94 (2005) 3303–3313.
[59] T. Nowotny, R. Huerta, H. D. Abarbanel, M. I. Rabinovich, Self-organization in the olfactory system: one shot odor recognition in insects, Biological cybernetics 93 (2005) 436–446.
[60] A. Montero, R. Huerta, F. B. Rodriguez, Neuron threshold variability in an olfactory model improves odorant discrimination, LNCS 7930 (2013) 16–25.
[61] A. Vergara, S. Vembu, T. Ayhan, M. Ryan, M. Homer, R. Huerta, Chemical gas sensor drift compensation using classifier ensembles, Sensors and Actuators B: Chemical 166 (2012) 320–329.
[62] A. V. M. H. R. H. Irene Rodriguez-Lujan, Jordi Fonollosa, On the calibration of sensor arrays for pattern recognition using the minimal number of experiments, Chemometrics and Intelligent Laboratory Systems (2013).
[63] R. A. Jortner, S. S. Farivar, G. Laurent, A simple connectivity scheme for sparse coding in an olfactory system., J Neurosci 27 (2007) 1659–1669.