Bio-inspired Encoding for a Real-time and Stable Odor Detection with a Highly-Redundant Optical Artificial Olfactory System

Gabriele Magna*[1], Eugenio Martinelli[2], Roberto Paolesse[1], Corrado Di Natale[2]

1. Department of Chemical Science and Technology, University of Rome Tor Vergata, via della Ricerca Scientifica 1, 00133 Roma, Italy
2. Department of Electronic Engineering, University of Rome Tor Vergata, Via del Politecnico 1, 00133; Roma Italy

*corresponding author: gabriele.magna@uniroma2.it, via della Ricerca Scientifica 1, 00133, Rome, Italy

Abstract

The standard model of artificial olfaction includes arrays of partially selective sensors and pattern recognition algorithms. Despite the brilliant results, such simplified platforms fail to achieve most of the natural olfaction features even if current sensors may rival olfactory receptor performances. Conversely, the olfaction architecture is specific to processing signals of millions of neurons, consisting of copious copies of a limited variety of receptors. The actual gas sensor arrays are still far from the abundance of natural olfaction receptors, so pattern recognition algorithms mainly differ from the processing in natural olfaction.

This paper tries to cover this gap by applying a biologically derived processing algorithm to a highly redundant array. A webcam recorded the RGB channel color intensities of 15 optical indicators due to exposure to different chemical vapors in real-time. Each pixel was assumed as a distinct sensor obtaining about 28000 copies representative of 45 cross-selective virtual receptors.

This outstanding redundancy enables deeply studying a model of the olfactory bulb configured as a network of inhibitory and excitatory elements. Subsequently, a Self-Organizing Map (SOM) is used to create an over-segmented reference space from the network outputs, where the exposure to different stimuli produces adsorption-desorption paths of activated neurons in real-time. Remarkably, even when a large part of the sensors, 65%, provided anomalous responses, the network provides robust odor identification. Using such a large and easy-to-make platform opens the way to a better understanding and facile implementation of algorithms that trace the architecture of the biological olfaction pathway.

Keywords: *Artificial olfaction, Optical sensors, Gas Sensors, Redundant sensor array, Sensor Fault, Olfactory bulb model*

1

1. Introduction

Olfaction is a principal sense in all animals. It is essential in fundamental life functions such as feeding, mating, and avoiding dangers [1]. The sense of olfaction is based on detecting airborne molecules and on the ability to discriminate a selected pattern of molecules in a background of an innumerable chemical variety. This task should require many highly selective sensorial organs, each tuned to a specific molecule. However, one of the principal features of olfaction is the excess of the number of detectable odors compared to the pool of available receptors [2]. This difference is explained considering that a perceived stimulus corresponds to a combination of receptors' responses [3,4].

The attempts to replicate the principles of odor encoding led to artificial olfactory systems made by combining an array of partially selective gas sensors with a pattern recognition algorithm [5]. Despite the many applications of electronic noses, these systems are very far from reaching the characteristics of natural olfaction [6]. A common explanation of the differences between natural and artificial systems involves the limited sensitivity, selectivity, and stability of gas sensors. On the other hand, little attention is paid to the fact that also natural receptors seem affected by similar restrictions such as signal fluctuation, poor reproducibility, and short lifetime [7,8]. Thus, besides the properties of individual receptors, signal processing has to play a fundamental role in the olfactory system performance. The evidence above leads to reconsidering the role of single sensor performances, displacing the focus to developing array processing architectures with dedicated firmware/software to properly organize and decode the information carried by sensor array signals [9].

In the past years, examples of olfactory processing models have been applied to the data of artificial sensors. These attempts show that, in general, sensitivity, selectivity, and robustness increase in comparison with traditional pattern recognition algorithms. For example, Diamond et al. considered a model of the insect antennal lobe, able to process noisy signals in real-time [10]. They found that the early response is sufficient to identify, with high accuracy, up to 20 individual chemicals from the time-series responses of metal oxide sensors. Borthakur et al. developed a spiking neural network (SNN) algorithm that utilizes online learning based on local synaptic rules such as spike-timing-dependent plasticity (STDP) for signal restoration and identification [11].

The physical detection of molecules in olfaction is carried out by a large number of individual olfactory receptor neurons (ORNs) each expressing a single kind of olfactory receptor proteins. A striking characteristic of olfaction is the redundancy of ORNs with the same nominal functionality. For instance, in mice, the little above 1000 kinds of different receptors are replicated in up to 5 million olfactory receptor neurons (ORNs) in each nasal cavity [12].

The processing of such a large number requires a precise organization of the olfactory pathway. It is a common opinion that the axons of ORNs carrying the same receptor converge into structures called glomeruli

2

[13]. As a consequence, the signals generated by millions of ORNs contribute to the activation of only thousands of glomeruli [14].

The glomerulus aggregation seems to imply that more than the redundancy, the important feature of olfaction is the number of receptor kinds. Nonetheless, the redundancy of sensors, at least for the obvious consequences related to the decrease of dispersion in repeated measurement [15], is considered a necessary ingredient for the full exploitation of biomimetic algorithms.

Few examples of high redundant chemical sensors have been provided so far. For instance, Marco et al. developed a large sensor array based on conducting polymers despite the difficulties of connecting a large number of resistors [16]. Conductive sensor array requires specific manufacturing processes; furthermore, the yield of processes combining microsystems technologies and chemical synthesis may not be particularly high with a non negligible amount of failures. A more simple and sustainable approach can be achieved with optical sensors. The use of computer peripherals, such as scanners and webcams, as sensor transducers was introduced several years ago [17,18]. The use of these devices is particularly appealing for electronic noses because a single camera shot can detect color or fluorescence changes of a multitude of indicators [19,20].

In this paper, this principle has been exploited for the preparation of a massively redundant sensor array made of an array of chemical color indicators. The data generated by this platform have been processed by a model, originally introduced by Riedel and Schild [21], of the encoding pathway of the glomerular units in the olfactory bulb.

The Riedel-Schild model is a multicompartmental closed-loop architecture implementing inhibitory and excitatory elements, which approximates the neural organization of the olfactory bulb limited to mitral and periglomerular cells [22]. While the mitral cells exert excitatory connections with homologous stimuli, namely sensors of the same kind, periglomerular units have an inhibitory function and receive heterologous inputs from all the sensors in the array. The original Riedel-Schild model was designed to process steady signals and, for example, was previously applied to process the steady-state signals of an array of four replicas of four different metal oxide semiconductor sensors [23]. These papers show that the olfactory bulb model can be used as an unsupervised preprocessing step. The application of a standard classification algorithm shows that the model preprocessing strongly increases the accuracy of compound identification and the stability with respect to fault sensor events.

On the other hand, several studies evidence that the analysis of the dynamic response of chemical sensors increases the discrimination power of the sensor array, and it may considerably reduce the response time of sensors [24,25]. Similar considerations were also observed in natural olfaction, for instance the latency of olfactory neuron signals is found to play a significant role in olfactory coding [26]

The importance of dynamic signals prompts the investigation of the capability of the Riedel-Schild model to process transient signals of a sensor array.

3

For this scope, 15 chemically sensitive materials that change their color under exposure to volatile compounds were selected. The materials were spotted in the same substrate and imaged by a digital camera in a single shot, recording three signals corresponding to the red, green, and blue color channels [20]. In these materials, the absorption of airborne molecules elicits a change in the absorption spectrum, which differently affects the intensity of the three color channels that might be considered individual sensor elements. Since each of the 15 sensitive materials produces three almost independent signals, the combination of real and virtual sensors leads to the definition of 45 kinds of receptors.

Concerning the possibility of producing replicas, the size of indicator spots exceeds the camera's spatial resolution; then, in the image, each spot is fractioned in many pixels. Since digital cameras show image resolutions of more than millions of pixels, this approach can easily generate arrays with sensor elements comparable with the number of ORNs in olfaction.

In this paper, since the 15 spots of receptors correspond to 9551 pixels of the image and considering three virtual sensor channels (red, green, and blue intensities), the whole array resulted in 28653 elements sorted into 45 different classes of sensors. In the simple implementation of the receptor-glomerulus pairing, the olfactory bulb model comprises 45 mitral cells, one per each class of sensors, and the same number of periglomerular cells, one per mitral cell.

The output of the olfactory bulb model is a multivariate signal whose dimension corresponds to the number of mitral cells. Data processed by the olfactory bulb model were projected onto a Self-Organizing Map (SOM) [27] to visualize and appreciate the peculiar signal of each volatile compound. The SOM is a particular neural network that provides a non-linear mapping of high dimensional space onto a regular grid of neurons. The SOM is governed by an algorithm that assigns to each neuron a representative vector in the original sensor space. In practice, the volume of the sensor space is distributed to the neurons of the SOM assigning larger volumes to regions where the density of experimental signals is smaller. Due to the similarity with the olfactory bulb projections, SOM has been frequently used to represent the sensor space of electronic noses [28–31].

The combination of the sensor array responses, the Riedel-Schild olfactory bulb model, and the SOM mapping showed the array selectivity with respect to the tested volatile compounds and the robustness versus changes of the inputs due to sensor failures. The volatile compound identification remains stable up to 65% of failure of individual sensors. These findings pave the way for the development of fully sustainable electronic noses expanding the biological similarities and fully exploiting the potentialities of sensors redundancy and biomimetic data processing.

2. Materials and Methods

2.1 Chemosensitive optical indicators

4

Colorimetric indicators were made of porphyrins, acid-base indicators, and their blends. The sensitive dyes were dissolved in a gas permeable polymeric matrix to avoid molecular aggregation. In a previous work, we reported that mixtures of porphyrins and pH indicators extend the sensing mechanisms beyond the characteristics of single constituents [32]. The method of sensor preparation is fully reported in ref. [32]. Briefly, the indicators were deposited onto a transparent plastic substrate from THF solutions composed of 1% wt of dyes, 33% wt of polyvinyl chloride, and 66% wt of di(2- ethylhexyl) sebacate as plasticizer. In the case of mixtures, the dye aliquot is equally divided between the two components. Figures 1A and 1B show the spots as they appear on the substrate and the corresponding sensitive material compositions.
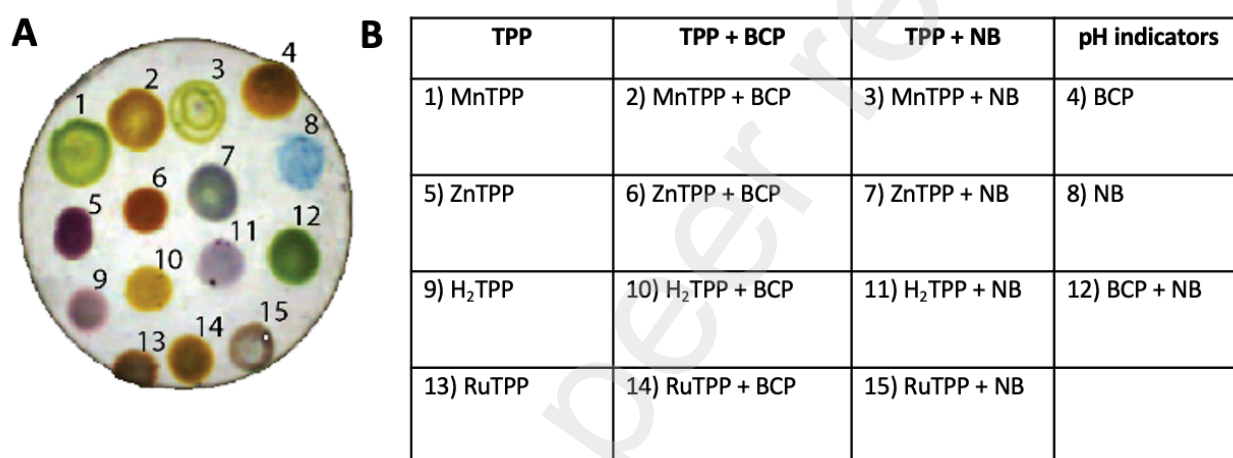

TPP	TPP + BCP	TPP + NB	pH indicators
1) MnTPP	2) MnTPP + BCP	3) MnTPP + NB	4) BCP
5) ZnTPP	6) ZnTPP + BCP	7) ZnTPP + NB	8) NB
9) H$_2$TPP	10) H$_2$TPP + BCP	11) H$_2$TPP + NB	12) BCP + NB
13) RuTPP	14) RuTPP + BCP	15) RuTPP + NB	

Figure 1. A) image of the sensitive spots on a transparent plastic substrate. B) List of indicators: MnTPP: Manganese-5,10,15,20-tetraphenyl porphyrin chloride; ZnTPP: Zinc-5,10,15,20-tetraphenyl porphyrin; H$_2$TPP: 5,10,15,20-tetraphenyl porphyrin; RuTPP: Ruthenium-5,10,15,20-tetraphenyl porphyrin; BCP: bromocresol purple; NB: nile blue.

2.2 Optical Sensor array

The chemosensitive indicators were spotted on the same glass substrate and placed in a cell with transparent walls for illumination and imaging. A computer screen illuminated the indicator scene displaying a white background, while a digital camera acquired the images in transmission mode at 1 frame per second (Philips SPC900NC/00). This experimental setup is known as Computer Screen Photo-assisted Technique [33]. The digital camera implements a color measurement technique based on the Bayer filter that correlates the colors measured in a pixel with those surrounding [34]. For simplicity, each pixel was considered an independent element, and the correlation due to Bayer filtering was included as part of the intrinsic chemical correlation of the indicators.

The pixels of each spot form a collection of homologous sensor replicas where signal differences among the pixels stem mainly from the non-uniform local distribution of dye into the membranes visible in Figure 1.

5

Additional effects like the non uniform light distribution and camera angle of view can introduce further differences in pixel signals.

The camera imaged the sensor surface at one frame per second at a resolution of 240x320 pixels. Images were processed to extract the color changes from the pixels of the spots by manually selecting circular regions of interest around each spot. The total area of 15 spots resulted in 9551 pixels that generated 28653 virtual sensors in 45 different kinds of sensors considering the triplet of data in the RGB channels.

2.3 Calibration samples

Sensors have been tested with vapors of the following six volatile organic compounds: acetic acid, butanol, hexane, triethylamine (TEA), trimethylamine (TMA), and toluene. These compounds are representative of different chemical families. Since the indicators are very sensitive to electron donor species, two similar amines have been included to test the capability to discriminate between similar compounds.

All compounds, except TMA, are liquid at standard conditions. The pure liquid was poured into a bubbler and kept at 25°C to stabilize the vapor pressure in the headspace (see Figure 2). In the case of trimethylamine, a certified bottle containing 3000 ppm of TMA in nitrogen gas was utilized.

Two mass flow controllers (MKS) regulate the concentration by diluting the vapor pressure with a nitrogen gas stream. Measurement protocol includes flushing the sensors before and after each exposure in a nitrogen flow to stabilize the baseline and restore the initial conditions. Each sequence lasts around 210 seconds, producing almost 5000 samples subdivided equally in training and test. Samples were classified individually, meaning that the algorithm aims at identifying the gas in the environment at each second. Notably, the test dataset includes a concentration for each compound not comprised in the training subset to investigate the classification performance toward unknown concentrations. The total flow rate was always kept constant at 200 sccm. Table 1 reports the measured concentrations. Finally, the overall dataset is almost 5000x28000 matrix, comprising around 140 million data points.

Table 1 *Concentrations of compounds utilized in training and test phases*

Sample	training concentrations [% Pressure vapor]	test concentrations [% Pressure vapor]
acetic acid	10, 15	15, 20
butanol	10, 20	20, 30
hexane	1, 3	3, 5
Triethylamine (TEA)	1, 2	2, 3
Trimethylamine (TMA)	600, 900 [ppm]	900, 1200 [ppm]
toluene	20, 25	25, 30
Samples (sampling time = 1s)	**2619 time samples**	**2337 time samples**

3. Multicompartment model

3.1 Processing architecture

Figure 2 depicts the overall architecture of the artificial olfaction system. The intensities of each color channel (red, green, and blue) have been considered independent virtual sensor signals. Thus, each pixel generates three virtual sensor elements. Eventually, from 15 indicators, 45 different artificial receptors are obtained. This number also defines the number of glomeruli of the multicompartment model.
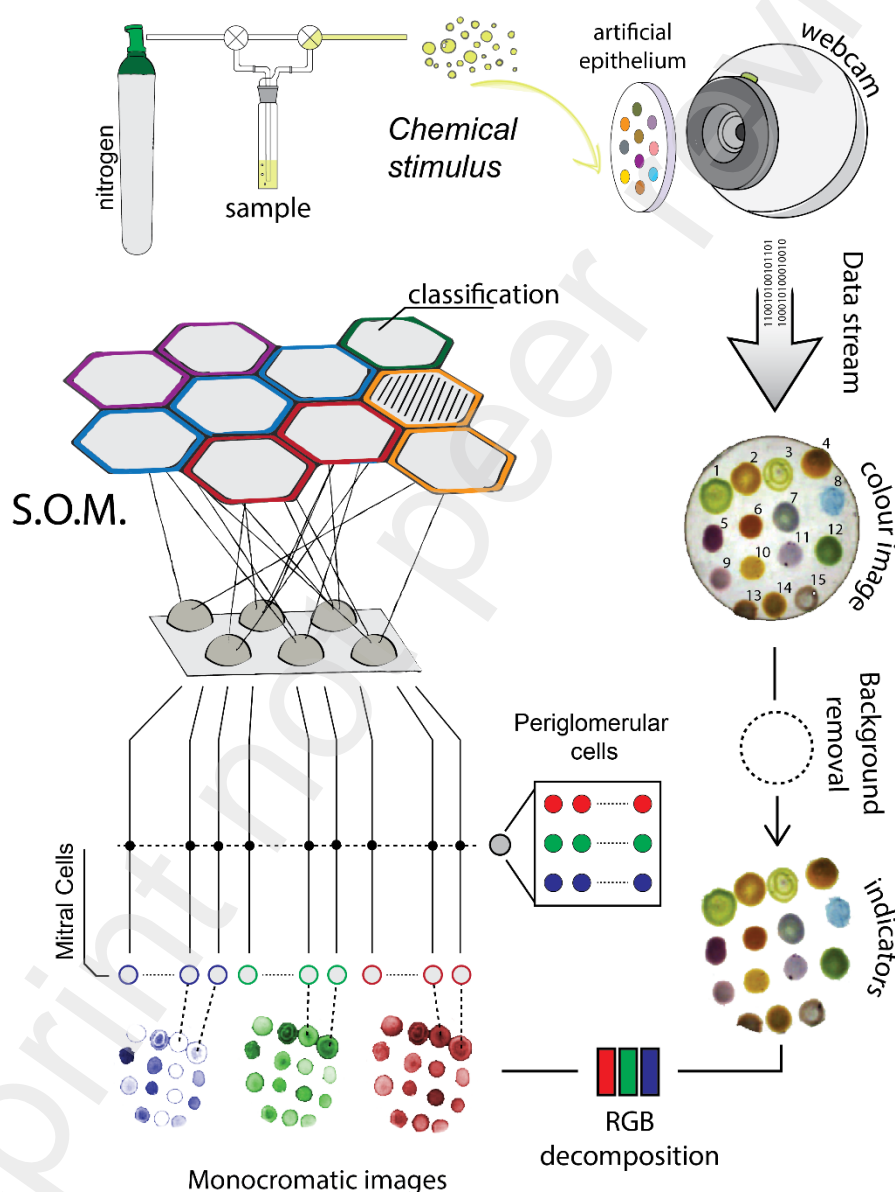

Figure 2. *Experimental architecture. A controlled vapor dispenser delivers a chemical stimulus to the sensors' surface (artificial epithelium). A camera takes continuous snapshots of the sensor's surface to record the color changes of spots. The RGB changes of pixels associated with the spots are extracted and used to feed the olfactory bulb network. Eventually, the mitral cell outputs are processed by a SOM that provides real-time visualization of the artificial olfaction system.*

7

3.2 Riedel-Schild model

The network contains 45 glomerular units (GUs), corresponding to the kinds of sensors. Each GU comprises a mitral cell (MC), a periglomerular cell (PG), the synapses that connect sensors to MC and PG cells to the sensors, and the synapses connecting the PG cells to the MCs of each glomerulus.

The mitral cells are univocally associated with corresponding PG cells and *vice versa*. PG cells provide intraglomerular feedback inhibition on the single branches of the mitral cell. The mitral cell receives inputs from sensors of the same type (homologous inputs), whereas the PG cells receive information from all sensors (heterologous inputs). Synaptic weights regulate the interaction of sensors with the afferent neurons. Figure 3 shows the connections in a single glomerulus.

Considering the i-th GU, a k-index is associated with PG branches, and a j-index is associated with MC afferent neurons. Thus, each r_{ik} sensor is connected to the k-th axon of i-th PG by the synaptic strength represented by d_{ik} parameters. In the i-th MC, the j-th microbranch of MC is connected to the r_{ij} sensor signal by the synaptic strength parameter c_{ij}. The output signals of mitral micro-branches depend on both $r_{ij} \times c_{ij}$ quantities and the inhibitory contribution coming from PG connections. In particular, in the i-th GU, the inhibitory connection between the k-th periglomerular output and the j-th mitral micro-branch is determined by the logical parameter f_{ijk} (see Figure 4). f_{ijk} is a matrix containing logic parameters that state the connection in the i-th GU between the j-th mitral axon and k-th periglomerular element. The network initialization assigns these matrix elements randomly to either 0 (not connected) or 1 (connected). Finally, the output of the i-th mitral cell is the sum of all m_{ij} contributions. The equations ruling the GU responses at each instant of time are reported in the following set of equations.

$$\begin{cases} p_{ik}(t) = d_{ik}(t) * |r_{ik}(t)| \\ m_{ij}(t) = c_{ij}(t) * r_{ij}(t) \prod_{k=1}^{N}(1 - f_{ijk} * p_{ik}(t)), \forall\, i,j \\ m_i(t) = \sum_{j=1}^{M_i} m_{ij} \end{cases} \tag{1}$$

The dynamic adaptation of the synaptic strength is governed by another set of equations modulating the synaptic strength using excitatory and inhibitory coefficients. This local adaptation makes the most stimulated sensors more influential than the others:

$$\begin{cases} c_{ij}(t+1) = c_{ij}(t) + \left(\gamma_a * m_{ij}(t) * |r_{ij}(t)| - \gamma_b * c_{ij}^3(t)\right) \\ d_{ik}(t+1) = d_{ik}(t) + \left(\delta_a * p_{ik}(t) * |r_{ik}(t)| - \delta_b * d_{ik}^3(t)\right) \end{cases} \tag{2}$$

The synaptic learning speed is defined by the four parameters: γ_a, γ_b, δ_a, and δ_b.

The model implements a 'local' and nonlinear version of the Hebbian rule. The variation of the synaptic weights c_{ij} and d_{ij} depends on the local dendritic potentials of MCs (m_{ij}) and PG cells (p_{ij}) micro-branches

8

rather than on the cell outputs m_i or p_i. It is worth noting that the four parameters tune both the local excitatory and inhibitory strength. Two parameters, γ_a and γ_b, regulate the adaptation of mitral cells, and the other two, δ_a and δ_b, the excitatory and inhibitory power of PG cells. The initial value of the synaptic weights between sensors and MCs (c_{ij}) and sensors and PG cells (d_{ik}) are randomly settled as 0 or 1.

A peculiar feature of the colorimetric sensor is that the variation of the intensity of the color can occur either positively or negatively. Since the sign might confound the inhibitory and excitatory synaptic actions, the absolute value of color changes has been properly considered to preserve the standard set of equations. As a result, the network can deal with both positive and negative signals and provides signed outputs.

The model itself does not require separated training and test phase, and the two sets of equations 1 and 3 can always be maintained active. Here, for the sake of simplicity, the adaptation has been limited to the model's training phase, and only the first set of equations was kept active during test data processing.

Figure 3. Multicompartment model applied to colorimetric indicator array as epithelium. Colour information is divided into RGB components. Thus, the spot is represented by pixels, considered independent sensors. The resulting sensors were univocally connected to a mitral cell. On the other hand, all the virtual sensors participate in the inhibition by connecting with the corresponding PG cell.

3.3 Self-Organizing Map

The self-organizing map (SOM) is an unsupervised neural network that represents events in a multidimensional space onto a grid of neurons. [27]. Here, the SOM has been used to represent the real-time progression of the mitral cell outputs that occur in a 45-dimensions vector space. The dynamic behaviors of signals produce paths onto the SOM, and the activated neurons can be assigned to a specific VOC. Furthermore, deviations from the path of repeated exposures can signal faults, drift, or general malfunctions. The sensor signals are expected to cluster into six classes (one per VOC). Since the sensors experience an adsorption and desorption phase for each exposure, each VOC class can be split into two subclasses describing the adsorption and desorption phases. Additionally, a 7th class corresponding to the baseline is also considered. The baseline is the chemical background (nitrogen gas) and the path of responses to VOCs originate and return to the baseline region.

The SOM was trained with the standard algorithm implemented in Matlab. After the network convergence, the neurons of the map were labeled with the class name corresponding to the VOC that produced the highest number of activation events. The never activated neurons were assigned to the class of the nearest labeled neuron chosen with a k-NN procedure [35] based on the neuron weights. In the test phase, MC outputs were projected on the map and classified according to the previously assigned neurons label.

3.4 Classification and robustness to fault sensors events

Figure 4 shows the rationale for the evaluation of the model performance. The tests include the correct identification of VOCs and the robustness of identification in the presence of fault sensor events.

The classification properties have been evaluated by comparing the data processed by the multicompartmental model with those obtained by simply averaging each glomerulus input (see Figure 4A). The average is a primitive approximation of the glomerulus functions, and it has been used in redundant sensor arrays to improve the signal-to-noise ratio [20,31]

Sensor malfunctions were simulated in a random number of sensors, abruptly setting the signals to zero (Figure 4B). Since fault events are generated at the pixel level, each fault affects the three virtual sensors. This harsh condition is critical for most sensor systems and algorithms dealing with faults. For example, most adaptive models maybe not be sufficiently fast to counteract these kinds of occurrences. Finally, the proposed system showed the same performance in the less challenging case, where some of the sensor responses progressively tended to zero. For the sake of synthesis, this case has not been reported.
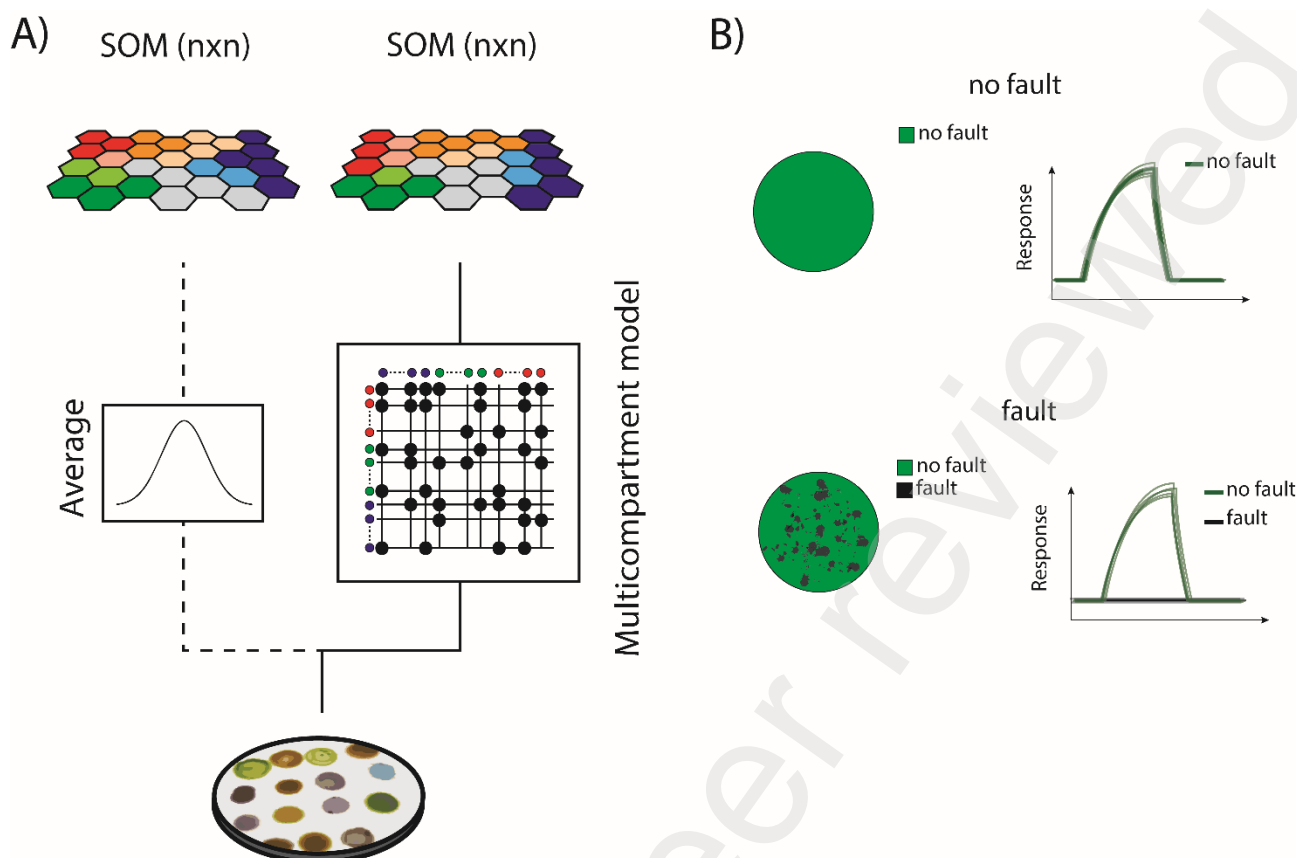
10

Figure 4. A) The benefits of the multicompartmental olfactory bulb model were evaluated by comparing the SOM processing of the mitral cell outputs with those achieved by a primitive olfactory bulb model that produces the average of the input signals. B) The resilience of sensor faults was evaluated by inducing faults on a random portion of the pixels of the spots up to 65% of sensing elements.

4. Results and Discussions

4.1 Sensors data and model training

The model's parameters were adapted through a recursive presentation of a training dataset made of around 2600 samples. Figure 5 reports the average signals of the training dataset. The network has been trained several times, changing the initial values of parameters. In most cases, the network parameters converged to a final value after ten iterations. The convergence of the model parameters is weakly dependent on the data sequence presentation (see Figure S1A and S1B), so the data sequence was maintained unaltered to minimize the parameters to check.

11

Figure 5. *Average signals of homologous indicators training signals. Each dye spot produces three virtual sensors considering the responses recorded in the red, green, and blue channels. The sequence of volatile compounds is shown on top.*

Before training, sensor data have been inspected to ensure equal treatment of the network of signals related to different VOCs. Indeed, the sensors showed a non-equal sensitivity for the test compounds, even if they are cross-selective. The more manifest case is the response to amines, which is about ten times larger than the response to alkanes. This outcome is chemically plausible considering that the dyes have high sensitivity to acids and basis and amines are strong Lewis bases [18]. The differences in the intensities of responses may bias the updating of weights during the model's training. Indeed, the adaptation of the parameters follows nonlinear relationships (equations 2), and different ranges of intensities in the mitral input signals might force the model to operate at too distant working points, making difficult the balance of inhibition and excitation (see Figure S2 as an example). A viable solution consists of normalizing data at every instant of time t using

12

linear normalization. This latter procedure divides the vector containing the sensor response at the time t by its norm, ensuring a constant magnitude of inputs over time. Data are normalized as reported following:

$$r(t) = \hat{r}(t)/norm(\hat{r}(t)) \tag{3}$$

where $r_{ij}(t)$ and $\hat{r}_{ij}(t)$ are the signals after and before the normalization.

To correctly interpret the outputs of the model, the signals of mitral cells are scaled back by the same normalization factor $norm(\hat{r}(t))$. A renormalized value $\hat{m}_i(t)$ can be calculated from mitral outputs using the following transformation:

$$\hat{m}_i(t) = m_i(t) * norm(\hat{r}(t)) \tag{4}$$

4.2 Riedel-Schild output model output

Figure 6 shows the input-output signals of mitral cells during exposure to 2% of the vapor pressure of triethylamine. Comparing the averages of the input signals of mitral cells (Figure 6A) with the corresponding outputs (Figure 6B), the evident increase in the signal-to-noise ratio is the first straightforward advantage of the convergence of signals into the GUs. However, the comparison of Figures 6A and 6B shows that the output signal-to-noise ratio is larger than the average of the input signals.

Figure 6A shows that the inputs of mitral cells have different response times. Remarkably, some signals start in advance of the others. Such differences result from the local and momentary non-uniform gas distribution inside the measurement chamber and the non-homogeneous distributions of indicator molecules inside the spots. Indeed, since indicators are dispersed in the polymeric matrix, VOCs have to diffuse through the polymer before reacting with dye molecules. Thus, thicker spots have a longer response time.

In Figure 6B, the mitral cells that receive inputs from faster sensors show more prominent outputs despite other mitral cells receiving larger steady-state signals.

This behavior evidences the effect of inhibition carried out by the periglomerular cells. Indeed, the first sensor that responds also triggers the inhibition of the other mitral cells decreasing their output. The network results strongly depend on the different sensor's response times.
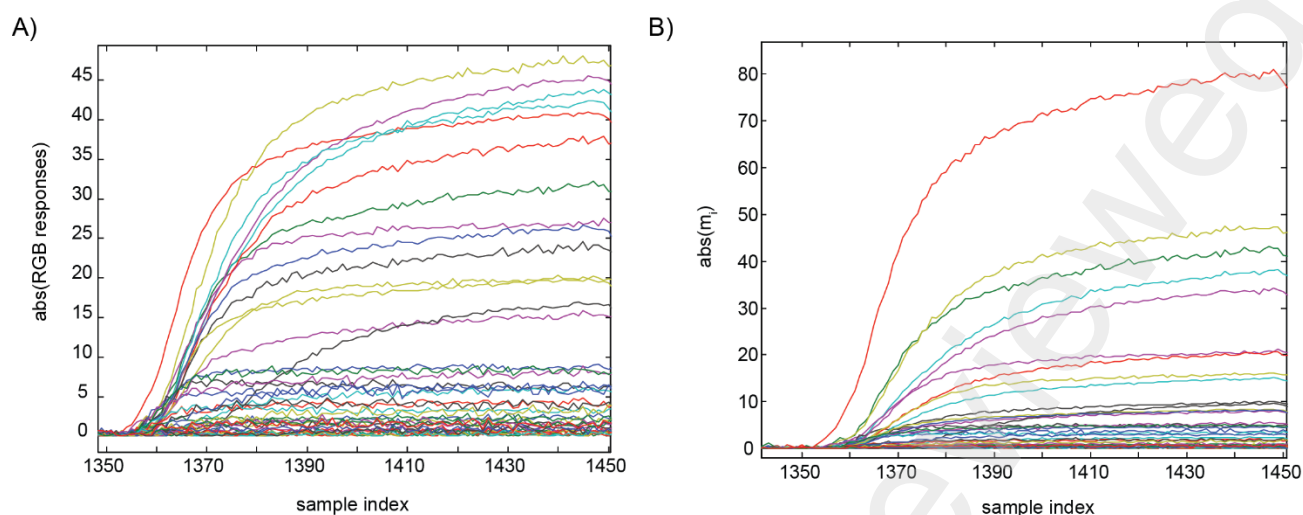
13

Figure 6. Dynamic signals in case of triethylamine exposure. A) The averages of mitral cells inputs and B) corresponding mitral cells outputs.

At the end of the training, the SOM processed mitral cell outputs to visualize signals' similarities and differences. The multicompartment model responses were utilized to build a 15×15 SOM with hexagonal topology and a Euclidean distance weight function.

Figure 7A shows the SOM mapping of the training dataset. In order to produce a labeled map, each SOM neuron was attributed to the VOC whose signals had mostly activated the neuron. It is interesting to note that, even if each VOC activates different neurons during the adsorption and desorption phases, the neurons pertinent to each VOC form closed regions. The baseline occupies a large portion of the map and directly communicates with all the other regions.

The distribution of volatile compounds on the SOM segregates the amines (electron donor species) from alcohol and acid, which are projected on the opposite side of the map. Thus the distribution of VOC classes reflects the chemical properties of the volatile compounds as captured by the indicators. Similar arrangements have been reported to occur in the olfactory bulb of mice [38].

The dynamics processing enables tracking onto the SOM the progression of the olfactory bulb signals. For example, Figure 7B shows the progression baseline-adsorption-desorption-baseline of six volatile compounds. Each path corresponds to the sequence of the neurons activated as the signal advances over time. Except for amines, the signals move from the baseline region to the corresponding VOC regions without crossing any area belonging to other VOCs. Thus, the VOC class is correctly identified just once tracks leave the baseline.

On the contrary, in the case of amines, before reaching the correct VOC regions, the signal crossed the region pertinent to hexane. In practice, amines are confused with hexane immediately after exposure and are correctly identified only later. Such a behavior is not completely surprising, considering that in triethylamine

14

and triethylamine, the amino group is complemented by alkyl groups, which confer some of the properties of alkanes.
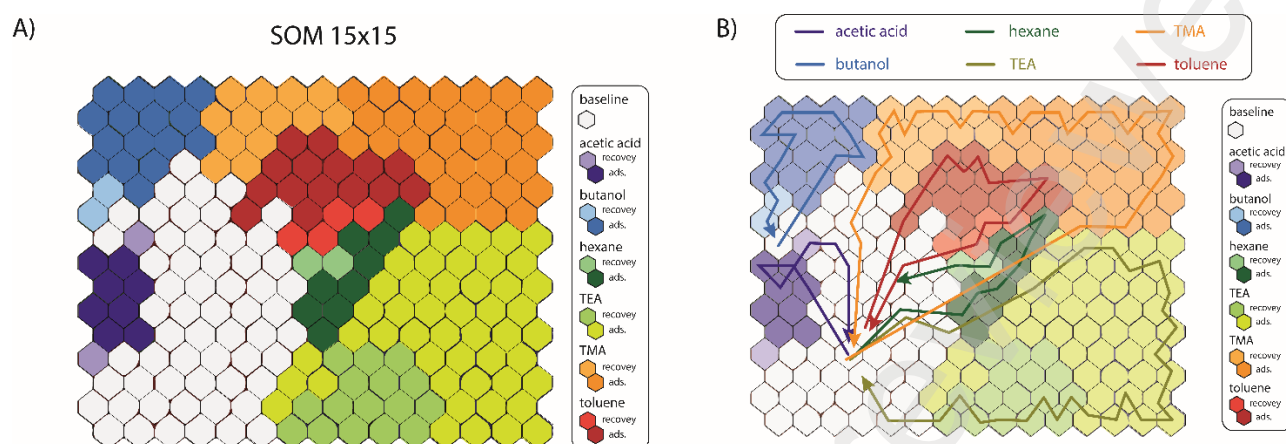

Figure 7. *A) The mitral cells' outputs are projected onto a 15x15 SOM. The signals of each VOC are mapped in a distinct region enabling the identification of the volatile compounds. Signals in adsorption and desorption span distinct sub-regions. B) path of mitral cells' outputs corresponding to the whole sequence of adsorption and desorption of the six volatile compounds.*

4.3 Tolerance to fault sensors events

The influence of faulting events on mitral cells' was evaluated by simulating a variable percentage of fault sensor events (see Figure 5B). In detail, several models were built on the training dataset and then validated using the training dataset with different percentages of faulting elements. Finally, classification was performed using a 15x15 SOM.

The effect of faults on the average sensor values is clearly visible, in particular with 65% of sensor faulted, as pictured in Figure S3. As expected, without the network processing, the intensity of the output decreases almost proportional to the percentage of fault sensors, and the classification error rapidly increases with the number of fault sensors (see figure 8). It is worth reminding that a classification error occurs when the MC signal activates a SOM neuron assigned to another class.

The effect of the olfactory bulb model to mitigate the effects of fault sensors is shown in figure 8. The dependence of correct classification with respect to the percentage of fault sensors depends on the model parameters and δ_a that rule the adaptation of the excitatory and inhibitory components of the model during training. However, the tolerance to fault events is mostly sensitive to the inhibition parameters (δ_a), and it is substantially independent of the excitatory strength (γ_a). In the parameters range shown in Figure 8, the multicompartment model always improves the classification compared to the non-processed signals. The optimal tolerance to fault is achieved with δ_a in the range $0.01 - 0.05$.
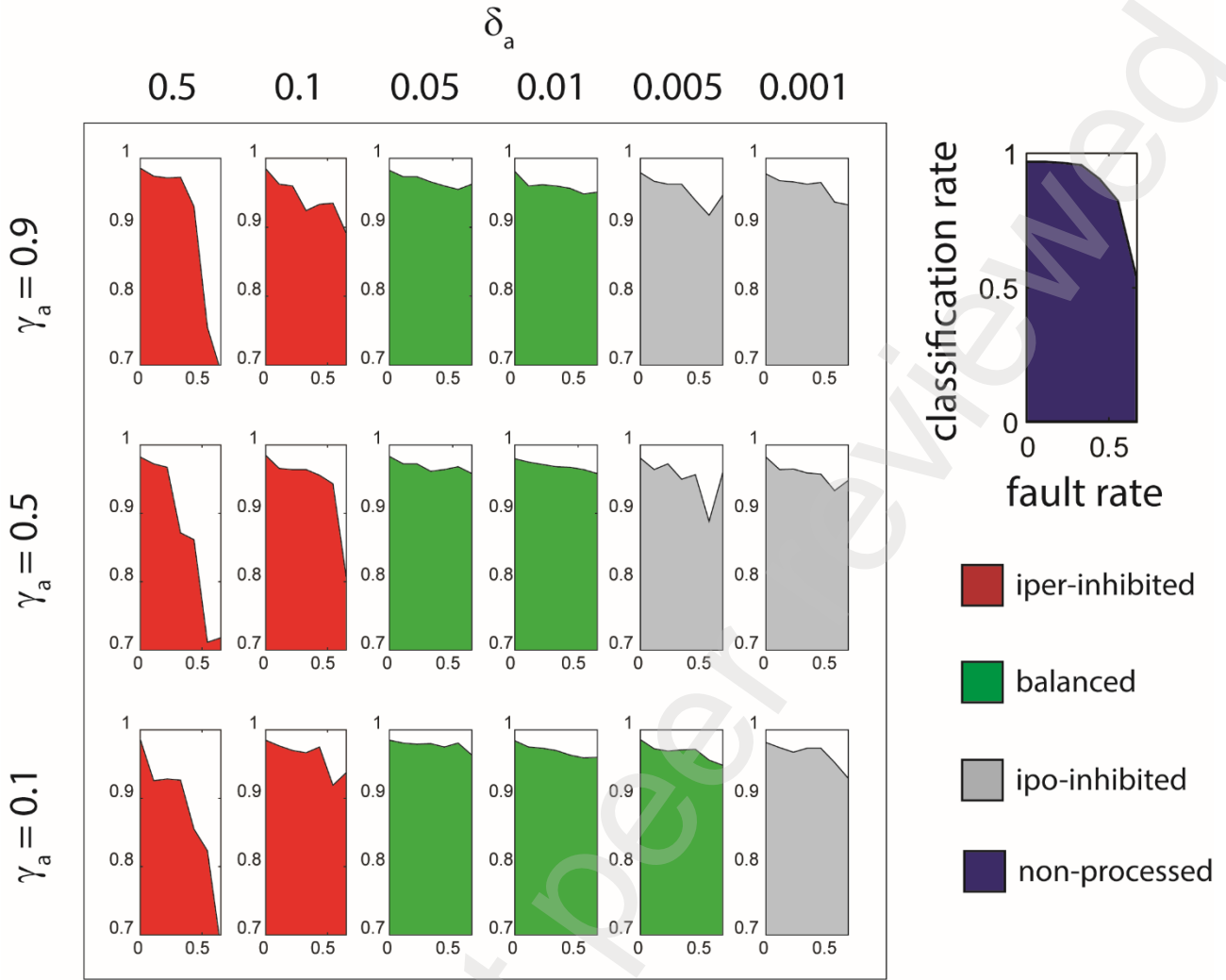
15

Figure 8. *Parameters influence over classification rates of SOMs $\gamma_a \in [0.1\ 0.9]\ \delta_a \in [0.001\ 0.5])$ over the rate of sensor faults (from 0 to 65%). The green color indicates the optimal range of parameters. Generally, outside this range, performances are still higher than those obtained without multicompartment data processing (blue box on the right upper corner of the figure).*

The dependence on the inhibition parameters (δ_a) can be explained by considering the balance between the excitatory and inhibitory components that regulate the feedback loop. In the case of large values of δ_a, the strength of inhibition prevails over the excitatory components. The simulated fault consists of setting the signal of the faulted sensors to zero. If the inhibition prevails, the reduction of input signals to the olfactory bulb elicits a reduction of the inhibition and then an increase in the network output (Figure S4).

On the other hand, when the excitation prevails, the network cannot compensate for the reduction of input signals. The effects of the olfactory bulb model are lost, and the outputs of GU mitral cells decrease in intensity due to fault (Figure S5). When excitation and inhibition are balanced, the olfactory bulb model is almost unaffected by fault events even when they reach 65% of the totality of sensors (Figure 9).
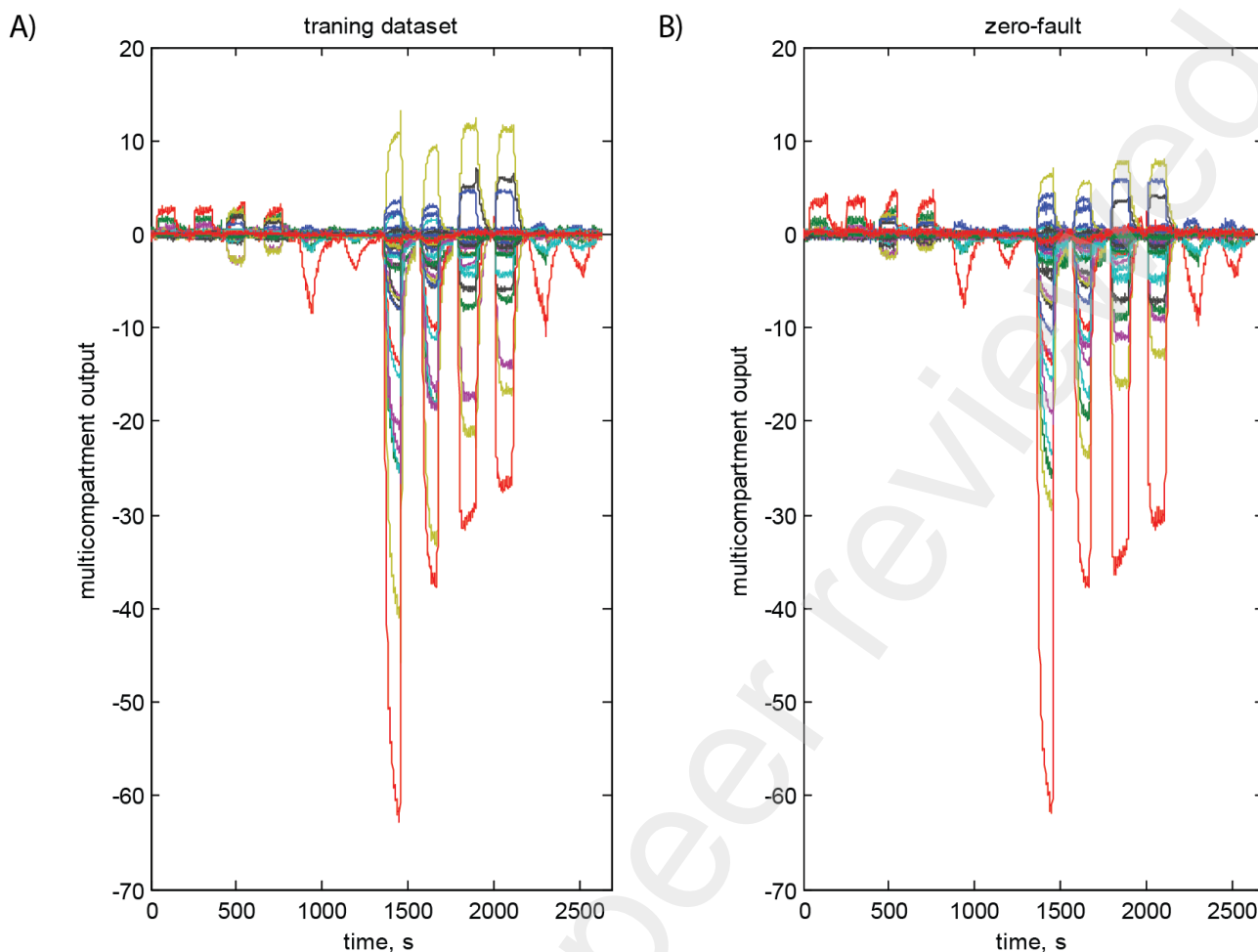
16

Figure 9. *Effect of zero-faults on multicompartment outputs in case of optimal parameter conditions. A) shows the response to the training dataset, whereas B) the output of the multicompartment model when the responses of 65% of sensors are set to zero.*

Figure 10 shows as an example the signal paths corresponding to the response to triethylamine of a full sensor array and with 65% of sensors with the output signal set to zero. The two paths activate neurons of the same region. It is interesting to note that even if the classes of samples are correctly predicted, the map is sensitive to the minor variations induced by the faults, as shown by the two paths. The invariance of paths during faulting events produces a very variation in the classification rates both in training and test (see Figures S6 and S7). On the contrary, the performance without the processing network starts to deteriorate at very small faulting percentages (Figure S8).

Figure 10. Progression of mitral cells outputs onto the SOM grid in case of no-fault (black trace) and faulting sensors (65%, red path) in the case of triethylamine exposure. B) Progress of mitral cells output during one of the exposures to triethylamine (TEA). The path with intact sensor array is compared with the path produced with 65% of faulted sensors. The two paths are largely overlapped, and more importantly in both cases only the SOM neurons pertinent to TEA are activated.

Conclusions

The implementation of olfactory processing paradigms is expected to improve the performance of artificial olfaction systems. Some of these benefits have been episodically shown by applying olfactory-derived neural network models to electronic noses. However, current electronic noses are made of few sensor units, and, except for specific cases, the replica of sensors is not a common feature of standard electronic noses. The lack of redundancy still hinders the full exploitation of the advantages of olfactory processing.

Among the sensor technologies, optical indicators are most prone to easily give rise to a multitude of replicated sensors. In particular, these sensors can use readily available devices, such as high-quality digital cameras, as transducers. Images of a set of spots of chemically sensitive dyes can be easily turned into a massive sensor array characterized by large redundancy of sensors. In such an architecture, the pixels of the image correspond to individual sensors; thus, the collection of pixels related to the same dye provides the set of replicated sensors. The sensor signals correspond to the intensity of the color channels.

In this paper, this concept has been used to develop a sensor array made of 15 sensitive materials. The camera pixels and the assumed independence of red, green, and blue channels intensity obtained a total of about 28000 sensors divided into 45 classes of real and virtual sensors.

Sensors' signals have been collected during the absorption and desorption of pure volatile compounds.e c The signals have been processed by a multicompartmental model of the olfactory bulb made of excitatory

18

and inhibitory mitral and periglomerular cells. The model was trained and tested with the dynamic response of the sensors. The output of the model has been mapped onto a SOM.

Eventually, for each exposure to a volatile compound, the signals of the artificial olfaction describe a path onto the SOM grid moving, in adsorption, from the baseline to the region of the pertinence of the volatile compound and, in desorption, back to the reference. The entrance of the signal in the region of the pertinence defines the times at which the volatile compounds are identified.

The model showed a wide range of optimal parameters in which the outputs proved to be robust versus input variations or defections. Remarkably, mitral outputs displayed analogies with biological olfaction, such as the real-time robustness versus input sensitivity loss or removals. Even removing up to 65% of the receptor inputs, the VOC identification properties of the model are almost unchanged. The investigation of dynamic responses of networks further evidence that early activated mitral cells are strongly enhanced while the late responding glomeruli are inhibited. This outcome highlights the importance of spatial distributions of receptors in the array as well as the direction of odorant stimulation. Further experiments will be necessary to ascertain the capability to detect complex odors even in variable backgrounds.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT author statement

Gabriele Magna: Formal analysis, Investigation, Measurements, Writing – original draft, visualization, software; **Eugenio Martinelli:** Conceptualization, Methodology, Writing — review and editing, Supervision, Project administration; **Roberto Paolesse:** Conceptualization, Methodology, Writing review and editing, Supervision, Project administration; **Corrado di Natale:** Conceptualization, Methodology, writing – original draft, Writing review and editing, Supervision, Project administration

References

[1]   B.W. Ache, J.M. Young, Olfaction: Diverse species, conserved principles, Neuron. 48 (2005). https://doi.org/10.1016/j.neuron.2005.10.022.

[2]   M. Meister, On the dimensionality of odor space, Elife. 4 (2015). https://doi.org/10.7554/eLife.07865.

[3]   B. Malnic, J. Hirono, T. Sato, L.B. Buck, Combinatorial receptor codes for odors, cell. 96 (1999) 713–723. https://doi.org/10.1016/S0092-8674(00)80581-4.

[4]   J.D. Mainland, Y.R. Li, T. Zhou, W.L. Liu, H. Matsunami, Human olfactory receptor responses to odorants, Sci. Data. 2 (2015) 150002. https://doi.org/10.1038/sdata.2015.2.

[5]   J.W. Gardner, P.N. Bartlett, A brief history of electronic noses, Sensors Actuators B. Chem. 18 (1994)

19

210–211. https://doi.org/10.1016/0925-4005(94)87085-3.

[6]   I. Manzini, D. Schild, Di Natale C., Principles of odor coding in vertebrates and artificial chemosensory systems, Physiol. Rev. 102 (2022) 61–154. https://doi.org/10.1152/physrev.00036.2020.

[7]   J.E. Schwob, Neural regeneration and the peripheral olfactory system, Anat. Rec. 269 (2002) 33–49. https://doi.org/10.1002/ar.10047.

[8]   G. Lowe, G.H. Gold, Olfactory transduction is intrinsically noisy, Proc. Natl. Acad. Sci. U. S. A. 92 (1995) 7864–7868. https://doi.org/10.1073/pnas.92.17.7864.

[9]   S. Raman, B.; Stopfer, M.; Semancik, Mimicking Biological Design and Computing Principles in Artificial Olfaction, ACS Chem. Neurosci. 2 (2011) 487–499.

[10]   A. Diamond, M. Schmuker, A.Z. Berna, S. Trowell, T. Nowotny, Classifying continuous, real-time e-nose sensor data using a bio-inspired spiking network modelled on the insect olfactory system, Bioinspiration and Biomimetics. 11 (2016). https://doi.org/10.1088/1748-3190/11/2/026002.

[11]   A. Borthakur, T.A. Cleland, A spike time-dependent online learning algorithm derived from biological olfaction, Front. Neurosci. 13 (2019). https://doi.org/10.3389/fnins.2019.00656.

[12]   K. Kawagishi, M. Ando, K. Yokouchi, N. Sumitomo, M. Karasawa, N. Fukushima, T. Moriizumi, Stereological quantification of olfactory receptor neurons in mice, Neuroscience. 272 (2014) 29–33. https://doi.org/10.1016/j.neuroscience.2014.04.050.

[13]   S. Korsching, Olfactory maps and odor images, Curr. Opin. Neurobiol. 12 (2002) 387–392.

[14]   B.A. Johnson, M. Leon, Chemotopic odorant coding in a mammalian olfactory system, J. Comp. Neurol. 503 (2007) 1–34. https://doi.org/10.1002/cne.21396.

[15]   C. Di Natale, A. D'Amico, F.A.M. Davide, Redundancy in sensor arrays, Sensors Actuators A. Phys. 37–38 (1993) 612–617. https://doi.org/10.1016/0924-4247(93)80105-P.

[16]   S. Marco, A. Gutiérrez-Gálvez, A. Lansner, D. Martinez, J.P. Rospars, R. Beccherelli, A. Perera, T.C. Pearce, P.F.M.J. Verschure, K. Persaud, A biomimetic approach to machine olfaction, featuring a very large-scale chemical sensor array and embedded neuro-bio-inspired computation, Microsyst. Technol. 20 (2014) 729–742. https://doi.org/10.1007/s00542-013-2020-8.

[17]   N.A. Rakow, K.S. Suslick, A colorimetric sensor array for odour visualization, Nature. 406 (2000) 710–713. https://doi.org/10.1038/35021028.

[18]   D. Filippini, A. Alimelli, C. Di Natale, R. Paolesse, A. D'Amico, I. Lundström, Chemical sensing with familiar devices., Angew. Chem. Int. Ed. Engl. 45 (2006) 3800–3. https://doi.org/10.1002/anie.200600050.

[19]   D.R. Walt, Convergent, self-encoded bead sensor arrays in the design of an artificial nose, Anal. Chem. 71 (1999) 2192–2198. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0033153220&partnerID=40&md5=402ed942f5d90b66b8443965992119b0.

20

[20]   C. Di Natale, E. Martinelli, R. Paolesse, A. D'Amico, D. Filippini, I. Lundström, An experimental biomimetic platform for artificial olfaction, PLoS One. 3 (2008) e3139. https://doi.org/10.1371/journal.pone.0003139.

[21]   D. Schild, H. Riedel, Significance of glomerular compartmentalization for olfactory coding, Biophys. J. 61 (1992) 704–715. https://doi.org/10.1016/S0006-3495(92)81875-1.

[22]   T. Kosaka, K. Kosaka, Neuronal organization of the main olfactory bulb revisited, Anat. Sci. Int. 91 (2016). https://doi.org/10.1007/s12565-015-0309-7.

[23]   E. Martinelli, G. Magna, D. Polese, A. Vergara, D. Schild, C. Di Natale, Stable odor recognition by a neuro-adaptive electronic nose, Sci. Rep. 5 (2015). https://doi.org/10.1038/srep10960.

[24]   M.K. Muezzinoglu, A. Vergara, R. Huerta, N. Rulkov, M.I. Rabinovich, A. Selverston, H.D.I. Abarbanel, Acceleration of chemo-sensory information processing using transient features, Sensors Actuators, B Chem. 137 (2009) 507–512. https://doi.org/10.1016/j.snb.2008.10.065.

[25]   J. Burgués, S. Marco, Feature Extraction for Transient Chemical Sensor Signals in Response to Turbulent Plumes: Application to Chemical Source Distance Prediction, Sensors Actuators, B Chem. 320 (2020). https://doi.org/10.1016/j.snb.2020.128235.

[26]   S. Junek, E. Kludt, F. Wolf, D. Schild, Olfactory coding with patterns of response latencies, Neuroforum. 16 (2010) 277–280. https://www.scopus.com/inward/record.uri?eid=2-s2.0-78650159554&partnerID=40&md5=8afb18fcfdf407d82499482d5e7c6c98.

[27]   T. Kohonen, The Self-Organizing Map, Proc. IEEE. 78 (1990) 1464–1480. https://doi.org/10.1109/5.58325.

[28]   S. Licen, A. Di Gilio, J. Palmisani, S. Petraccone, G. de Gennaro, P. Barbieri, Pattern recognition and anomaly detection by self-organizing maps in a multi month e-nose survey at an industrial site, Sensors (Switzerland). 20 (2020). https://doi.org/10.3390/s20071887.

[29]   S. Licen, G. Barbieri, A. Fabbris, S.C. Briguglio, A. Pillon, F. Stel, P. Barbieri, Odor control map: Self organizing map built from electronic nose signals and integrated by different instrumental and sensorial data to obtain an assessment tool for real environmental scenarios, Sensors Actuators, B Chem. 263 (2018) 476–485. https://doi.org/10.1016/j.snb.2018.02.144.

[30]   M. Zuppa, C. Distante, P. Siciliano, K.C. Persaud, Drift counteraction with multiple self-organising maps for an electronic nose, Sensors Actuators, B Chem. 98 (2004) 305–317. https://doi.org/10.1016/j.snb.2003.10.029.

[31]   C. Di Natale, A. Macagnano, A. D'Amico, F. Davide, Electronic-nose modelling and data analysis using a self-organizing map, Meas. Sci. Technol. 8 (1997) 1236–1243. https://doi.org/10.1088/0957-0233/8/11/004.

[32]   F. Dini, G. Magna, E. Martinelli, G. Pomarico, C. Di Natale, R. Paolesse, I. Lundström, Combining porphyrins and pH indicators for analyte detection Direct Optical Detection, Anal. Bioanal. Chem.

21

407 (2015) 3975–3984. https://doi.org/10.1007/s00216-014-8445-9.

[33]   D. Filippini, S.P.S. Svensson, I. Lundström, Computer screen as a programmable light source for visible absorption characterization of (bio)chemical assays, Chem. Commun. (2003) 240–241. https://doi.org/10.1039/b210677a.

[34]   N. Lukac, R.; Plataniotis, Color Filter Arrays: Design and Performance Analysis, IEEE Trans. Consum. Electron. 51 (2005) 1260–1267.

[35]   A. Rauber, LabelSOM: On the labeling of self-organizing maps, in: Proc. Int. Jt. Conf. Neural Networks, 1999: pp. 3527–3532.

[36]   L. Fernandez, S. Marco, A. Gutierrez-Galvez, Robustness to sensor damage of a highly redundant gas sensor array, Sensors Actuators, B Chem. 218 (2015) 296–302. https://doi.org/10.1016/j.snb.2015.04.096.

[37]   B. Raman, T. Kotseroglou, L. Clark, M. Lebl, R. Gutierrez-Osuna, Neuromorphic processing for optical microbead arrays: Dimensionality reduction and contrast enhancement, IEEE Sens. J. 7 (2007) 506–514. https://doi.org/10.1109/JSEN.2007.891935.

[38]   I. Manzini, S. Korsching, The peripheral olfactory system of vertebrates: Molecular, structural and functional basics of the sense of smell | Das periphere olfaktorische system von vertebraten: Molekulare, strukturelle und funktionelle grundlagen des riechens, Neuroforum. 17 (2011) 110–118.
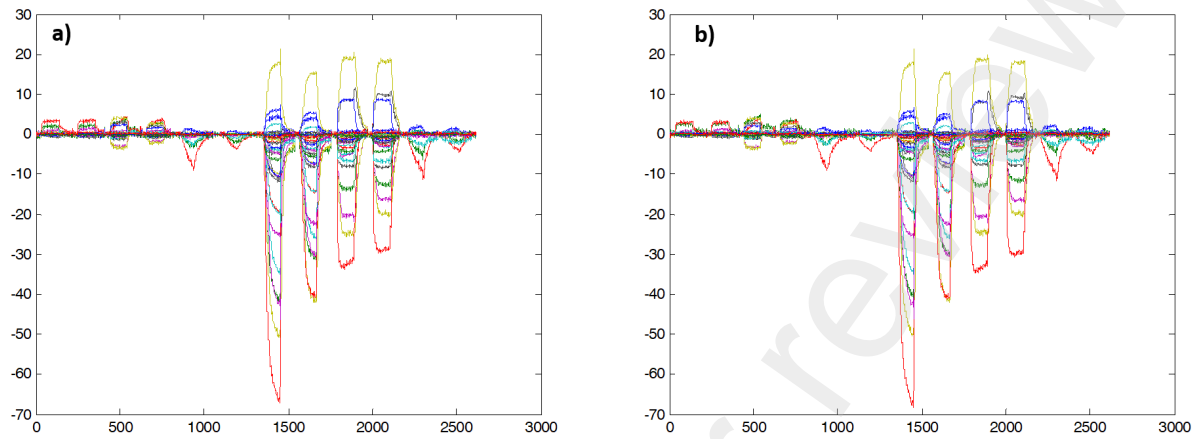
Figure S1. *Outputs of multicompartment model trained with (a) adopted temporal sequence and (b) a random permutate sample sequence.*
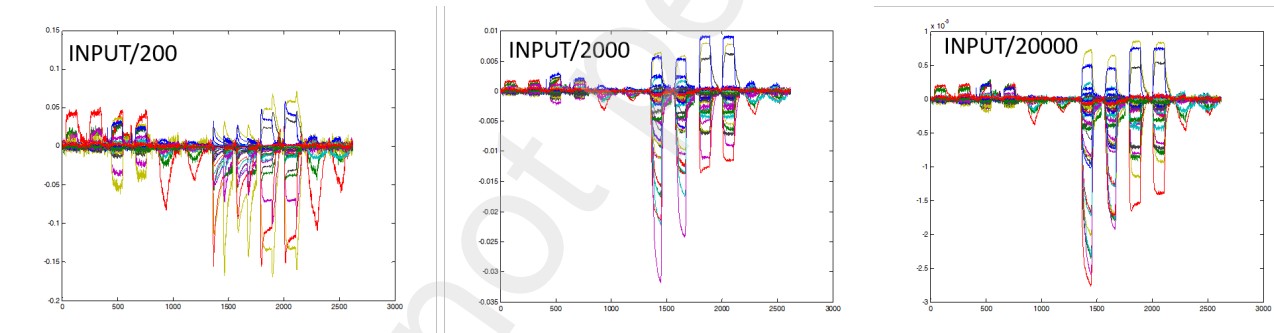

Figure S2. *Network response to same patterns of input with different ranges of amplification.*
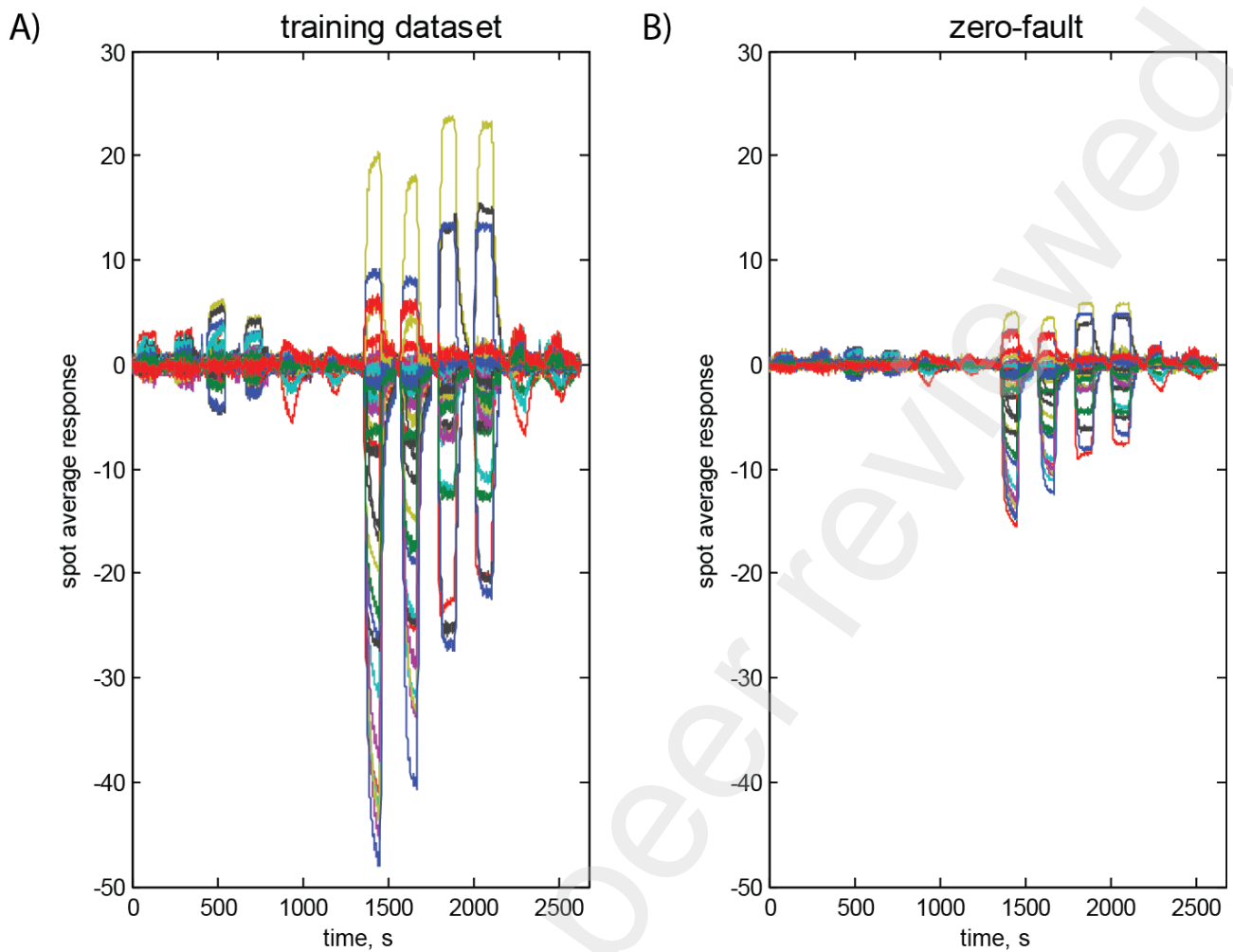
23

Figure S3. *Effect of zero-faults on average responses of single spots. A) shows the response to the training dataset, whereas B) the output of the multicompartment model when the responses of 65% of sensors are set to zero.*
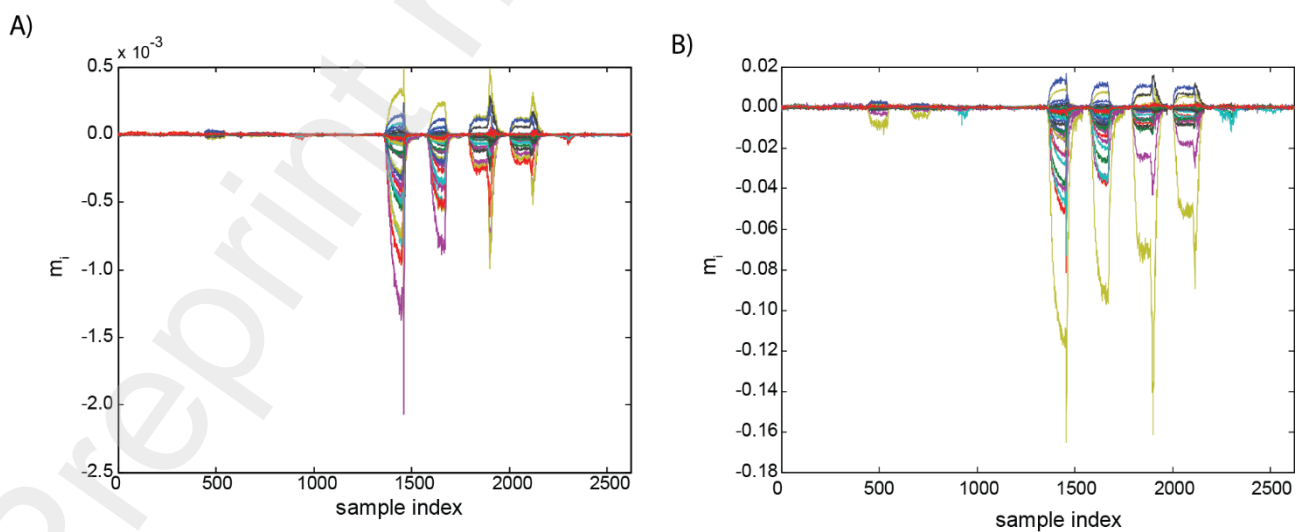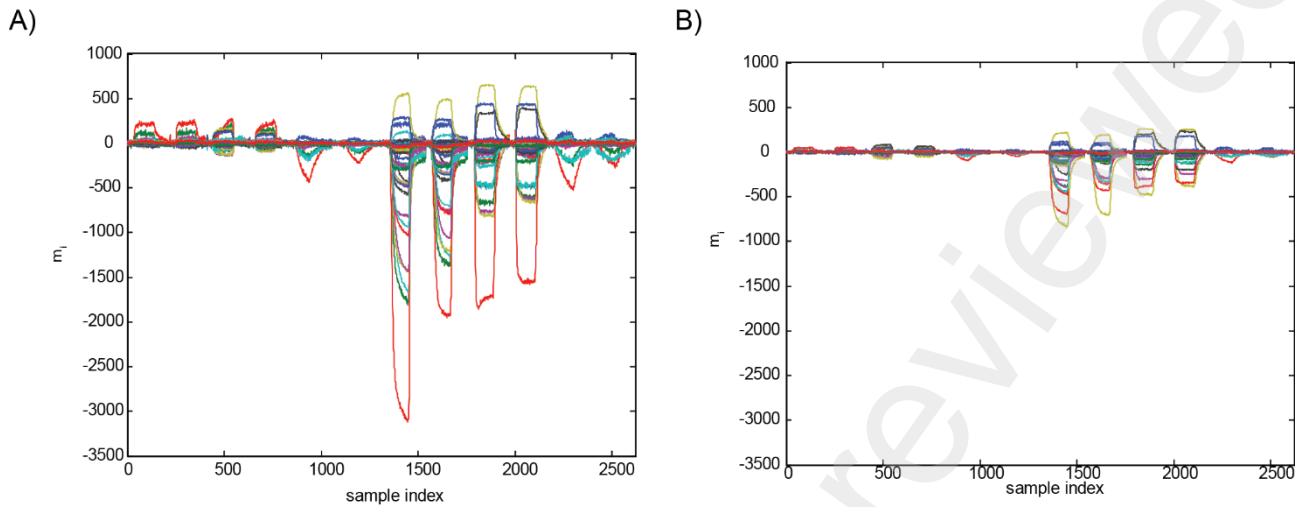

Figure S4. *Effect of zero-faults on multicompartment outputs in case of the hyper-inhibited network. A) shows the response to the training dataset, whereas B) the output of the multicompartment model when*
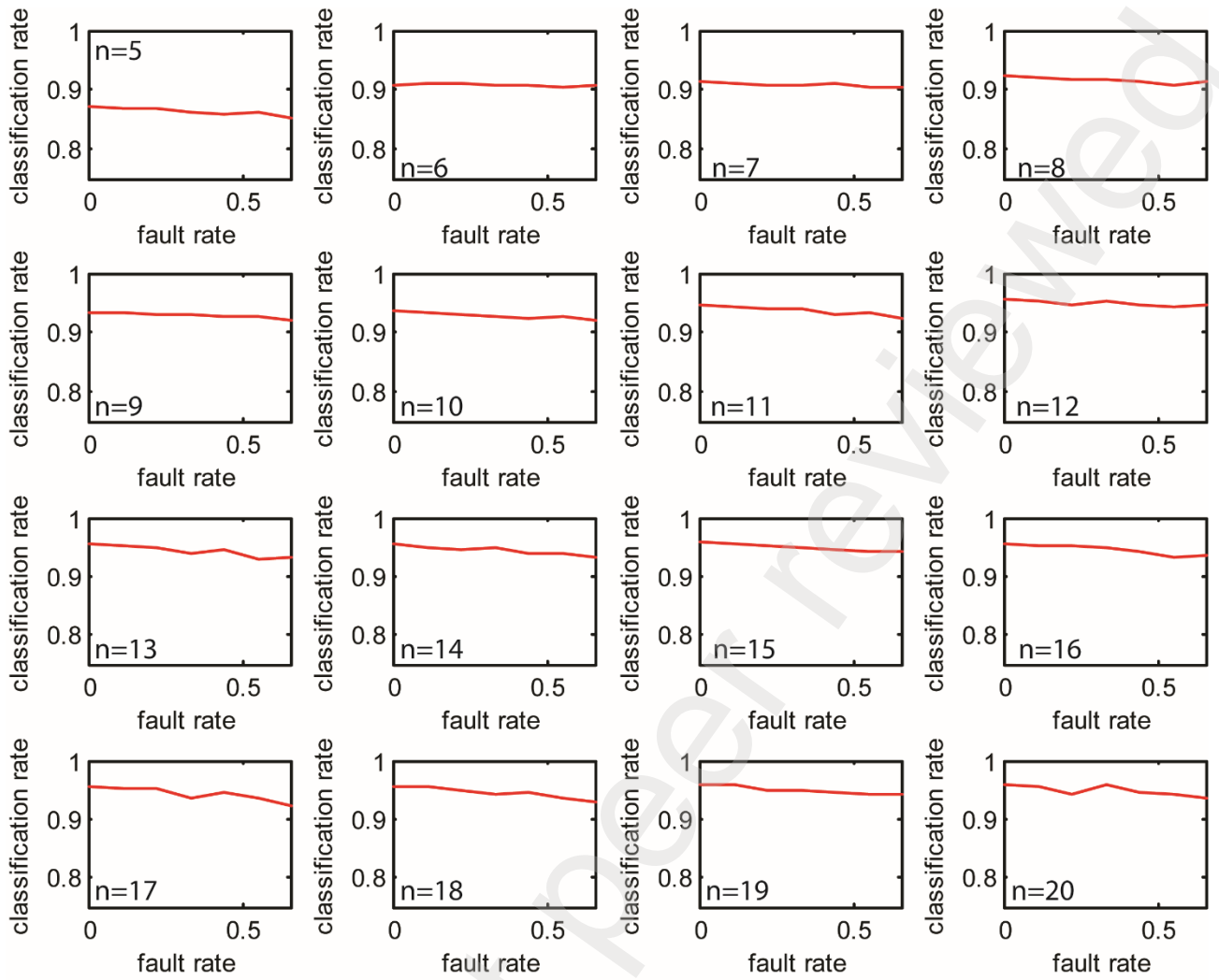
the responses of 65% of sensors are set to zero (NB test outputs are almost one order of magnitude higher than the training case).

A)

B)

Figure S5. *Effect of zero-faults on multicompartment outputs in case of the hypo-inhibited network. A) shows the response to the training dataset, whereas B) the output of the multicompartment model when the responses of 65% of sensors are set to zero.*

Figure S6. *Classification rates of training dataset in case of MC network preprocessing considering different fault rates ([0,0.65]) and SOM dimension (SOM dimension= n x n).*
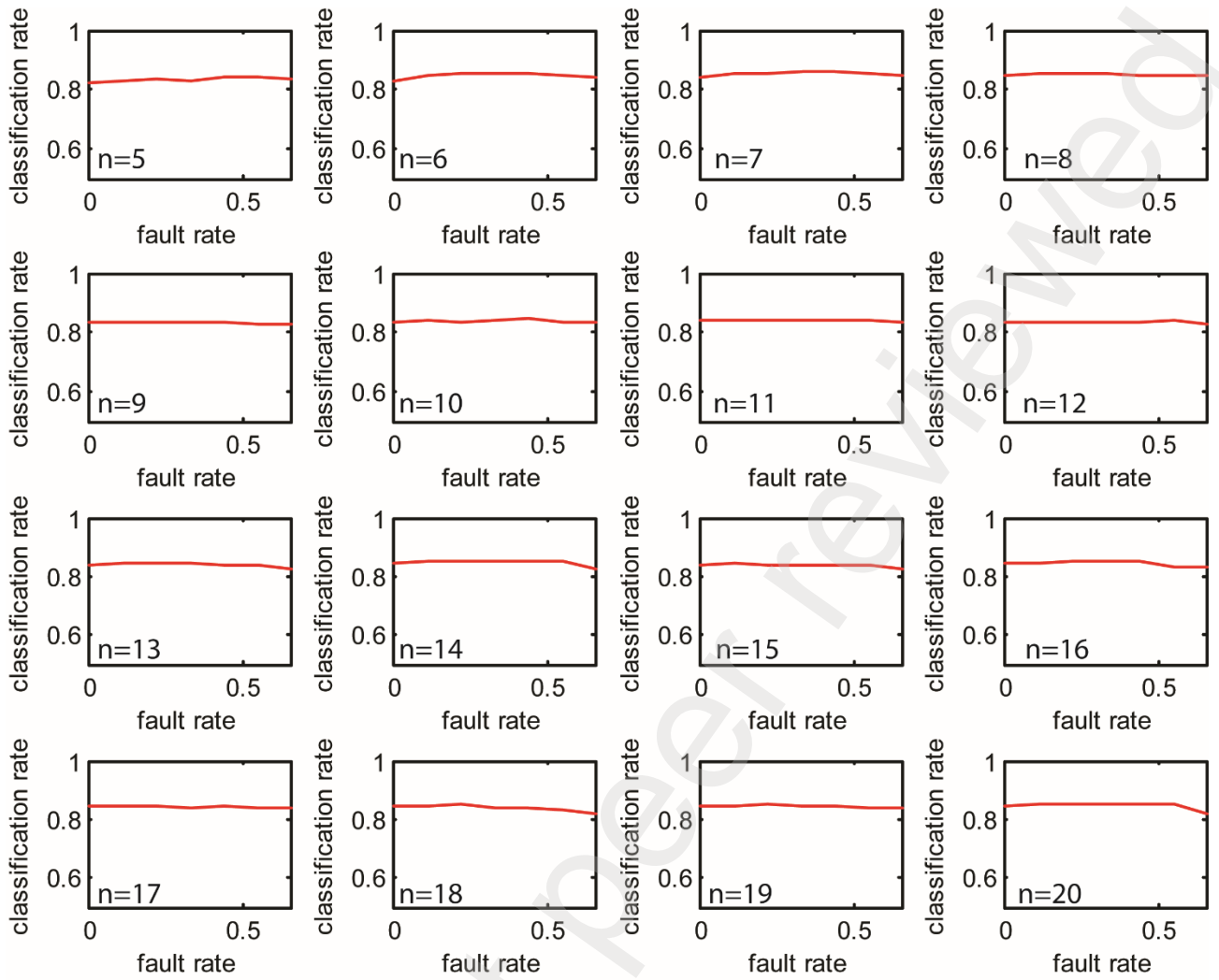
Figure S7. *Classification rates of test dataset in case of MC network preprocessing considering different fault rate ([0,0.65]) and SOM dimension (SOM dimension= n x n).*
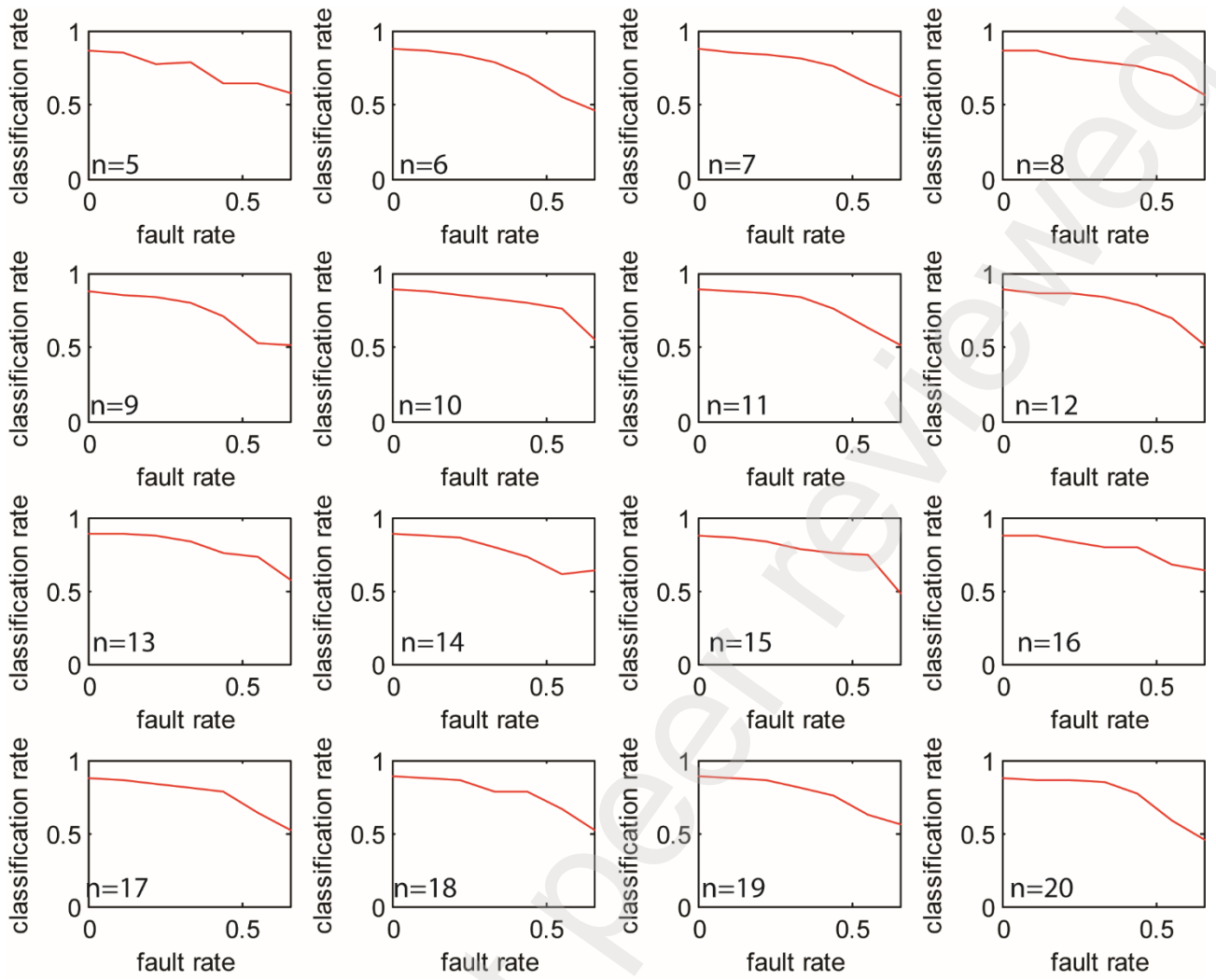
Figure S8. *Classification rates of training dataset in case of absence of MC network considering different fault rate ([0,0.65]) and SOM dimension (SOM dimension= n x n). Please note the different scales of classification rate with respect to the Figure S6 and S7 cases.*