A neuromorphic model of olfactory processing and sparse coding in the Drosophila larva brain

To cite this article: Anna-Maria Jürgensen et al 2021 Neuromorph. Comput. Eng. 1 024008

View the article online for updates and enhancements.

You may also like

- Two sparsities are better than one: unlocking the performance benefits of sparse–sparse networks
  Kevin Hunter, Lawrence Spracklen and Subutai Ahmad

- Combining sparse coding and time-domain features for heart sound classification
  Bradley M Whitaker, Pradyumna B Suresha, Chengyu Liu et al.

- FRIST—flipping and rotation invariant sparsifying transform learning and applications
  Bihan Wen, Saiprasad Ravishankar and Yoram Bresler
A neuromorphic model of olfactory processing and sparse coding in the Drosophila larva brain

Anna-Maria Jürgensen, Afshin Khalili, Elisabetta Chicca, Giacomo Indiveri and Martin Paul Nawrot

Computational Systems Neuroscience, Institute of Zoology, University of Cologne, Cologne, Germany
Department Genetics of Learning and Memory, Leibniz Institute for Neurobiology, Magdeburg, Germany
Bio-Inspired Circuits and Systems Lab, Zernike Institute for Advanced Materials, University of Groningen, Groningen, The Netherlands
Groningen Cognitive Systems and Materials Center, University of Groningen, Groningen, The Netherlands
Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland

E-mail: mnawrot@uni-koeln.de

Keywords: neuromorphic computing, spiking neural network, insect olfaction, artificial intelligence, Drosophila, sparse coding, spike frequency adaptation

Abstract
Animal nervous systems are highly efficient in processing sensory input. The neuromorphic computing paradigm aims at the hardware implementation of neural network computations to support novel solutions for building brain-inspired computing systems. Here, we take inspiration from sensory processing in the nervous system of the fruit fly larva. With its strongly limited computational resources of <200 neurons and <1.000 synapses the larval olfactory pathway employs fundamental computations to transform broadly tuned receptor input at the periphery into an energy efficient sparse code in the central brain. We show how this approach allows us to achieve sparse coding and increased separability of stimulus patterns in a spiking neural network, validated with both software simulation and hardware emulation on mixed-signal real-time neuromorphic hardware. We verify that feedback inhibition is the central motif to support sparseness in the spatial domain, across the neuron population, while the combination of spike frequency adaptation and feedback inhibition determines sparseness in the temporal domain. Our experiments demonstrate that such small, biologically realistic neural networks, efficiently implemented on neuromorphic hardware, can achieve parallel processing and efficient encoding of sensory input at full temporal resolution.

1. Introduction
Neuromorphic computing [1] is a novel paradigm that aims at emulating the naturalistic, flexible structure of animal brains on an analogous physical substrate with the potential to outperform von Neumann architectures in a range of real-world tasks [2, 3]. It can inspire novel AI solutions [4–6] and may support control of autonomous agents by spiking neural networks [7–9]. A major challenge for brain-inspired neuromorphic solutions is the identification of computational principles and circuit motifs in animal nervous systems that can be utilized on neuromorphic hardware to exploit its benefits.

Drawing inspiration from neural computation in the nervous systems of insects is particularly promising for developing neuromorphic computing paradigms. With their comparatively small brains ranging from \(\approx 10000\) neurons in the fruit fly larva to \(\approx 1\) million neurons in the honeybee, insects are able to solve many formidable tasks such as the efficient recognition of relevant objects in a complex environment [10, 11], perceptual decision making [12–14], or the exploration of unknown terrain and navigation [15–19]. They also show simple cognitive abilities such as learning, or counting of objects [20–24]. At the same time, their compact nervous systems are optimized for energy efficient computation with limited numbers of neurons and synapses, making them ideally suited to meet current neuromorphic hardware limitations regarding network size and topology. Spiking neural networks modeled after the insect brain have been shown to support efficient
sensory processing [25], learning [7, 26], foraging and navigation [27–29], and counting [28]. Model studies also include earlier neuromorphic implementations of insect-inspired computation [4, 5, 9, 30–33].

Sparse coding [34, 35] is a fundamental principle of sensory processing, both in invertebrates [36–40] and vertebrates [41–45]. By transforming dense stimulus encoding at the receptor periphery into sparse representations in central brain areas, the sensory systems of animals achieve energy efficient and reliable stimulus encoding [35, 46], which increases separability of items [47–50]. Sparse coding in neural systems has two major components [39]. Population sparseness refers to the representation of a stimulus across the entire population of neurons, such that only few neurons are activated by any specific stimulus and different stimuli activate largely distinct sets of neurons. Re-coding from a dense peripheral input to a sparse code in central brain areas supports stimulus discriminability and associative memory formation by projecting stimulus features into a higher dimensional space [51–53]. Temporal sparseness indicates that an individual neuron responds with only a few spikes to a specific stimulus configuration [34, 54, 55] supporting the encoding of dynamic changes in the sensory environment [42, 56] and memory recall in dynamic input scenarios [28].

We are interested in the transformation of a densely coded input into a sparse representation within an olfactory pathway model of the Drosophila larva. As a common feature across insect species, odor information is processed across multiple network stages to generate a reliable sparse code of odor identity in the mushroom body (MB) [36, 57, 58], a central brain structure serving as a hub for multi-sensory integration, memory formation and memory recall [10, 59]. A shared characteristic of the Drosophila larva brain and the here-used real-time neuromorphic hardware system is their relatively small network size. With this limited capacity, computational efficiency and frugal use of the limited resources are a major constraint. Implementing evolutionary-derived mechanisms from the insect brain that allow for sparse, thus more efficient stimulus encoding on the chip could help to broaden the scope of its applications. In our network model we test the efficiency of cellular mechanisms and network motifs in producing population and temporal sparseness and test their implementation on the mixed-signal neuromorphic hardware DYNAP-SE [60] in comparison to a software simulation using the Python-based spiking neural network simulator ‘Brian2’ [61].

2. Methods

2.1. Spiking neural network model

The architecture of the spiking neural network model as shown in figure 1(A) uses the exact numbers of neurons in each population and the reconstructed connectivity for one hemisphere as published in the electron-microscopic study of a single animal [62, 63]. The network consists of 21 olfactory receptor neurons (ORN) at the periphery, 21 projection neurons and 21 local interneurons (LNs) in the antennal lobe and 72 Kenyon cells (KCs). In each brain hemisphere there is exactly one anterior paired lateral (APL) neuron. We hypothesize that the APL receives input from most or all mature KCs [64] included in this network model. Due to technical limitations of the DYNAP-SE chip with a maximum in-degree of 64 synapses for one neuron we randomly chose 64 KCs that provide input to the APL. This choice was fixed for the model, both on the hardware network and in the software simulation. We further hypothesize, based on evidence in the adult species, that all ORNs and all KCs have a mechanism of cellular spike frequency adaptation (SFA).

2.2. Implementation on the DYNAP-SE neuromorphic hardware

The olfactory pathway model of the Drosophila larva was implemented using the dynamic neuromorphic asynchronous processor (DYNAP-SE) [60] (figure 1(C)). This processor is a full-custom mixed-signal analog/digital VLSI chip, which comprises analog circuits that emulate neurons and synapses with biologically plausible neural dynamics. Given the analog nature of the circuits used, the synapses and neurons exhibit parameter variability that is characteristic also of real neurons. The analog circuits used, implement multiple aspects of neural dynamics, such as spike-frequency adaptation (implemented as a shunting inhibitory synapse), refractory periods, exponentially decaying currents, voltage-gated excitation and shunting inhibition [60, 65]. The silicon neurons circuits, similar to their biological counterparts, produce spikes. In the chip, these are stereotyped digital events which are routed to target synapses by a dedicated address event representation (AER) infrastructure [66, 67]. The conductance-based synapses are current-mode circuits [65] that produce an EPSC with biologically plausible dynamics, which are then injected into the neurons leak compartment. This compartment acts as a conductance block, which decreases the input current as the membrane potential increases. One of the inhibitory synapses subtracts charge directly from the membrane capacitance and provides a shunting inhibition mechanism [65]. All other synaptic currents are in turn summed together and integrated in the post-synaptic neurons leak compartment.

The model (figure 1(A)) was initially developed in software and the neural architecture was then mapped onto the mixed-signal hardware by configuring the AER routers and programming the chip digital memories to connect the silicon neurons via their corresponding synapses. The parameters of the hardware setup were
fine-tuned using the on-chip bias generator, starting from the estimates provided by the software simulation. Computer-generated control stimuli, in the form of well defined spike trains, were provided to the chip via a custom field programmable gate array (FPGA) board. Each neuron population was implemented on a single core, using in total five cores and two chips. All the circuit biases of neurons belonging to different cores could be tuned independently. The synapses from ORN to PN, PN to KC, and KC to APL were designed as excitatory whereas the synapse from LN to PN and APL to KC were implemented as inhibitory. SFA was implemented in the ORN and KC neuron populations.

Three separate recordings were done, one for each of the three odors with 20 trial repetitions (figure 1(B)). Within each of the three experiments all conditions (different sparseness mechanisms enabled) were recorded always in the same order (LN + APL + SFA, LN + SFA, LN, APL + SFA, LN + APL, SFA).

2.3. Computer simulation of the spiking neural network

The simulations were implemented in the network simulator Brian2 [61] and run on an X86 architecture on an Ubuntu 16.04.2 Server. All neurons (figure 1(A)) were modeled as leaky integrate-and-fire neurons with conductance-based synapses. The membrane potential $v_i$ obeys a fire-and-reset rule, being reset to the resting potential whenever the spike threshold is reached. The reset is followed by an absolute refractory period of 2 ms, during which the neuron does not integrate inputs (table 1). The membrane potential of a neuron in a particular neuron population ($v_O^i, v_L^i, v_P^i, v_K^i, v_A^i$) is governed by the respective equation. The neuron parameters can be found in table 1.

\[
\frac{d}{dt} v_O^i = g_L (E_L - v_O^i) + g_{\text{Input}O} (E_E - v_O^i) - g_{\text{Input}A} (E_A - v_O^i) \\
\frac{d}{dt} v_L^i = g_L (E_L - v_L^i) + g_{\text{Input}L} (E_E - v_L^i) \\
\frac{d}{dt} v_P^i = g_L (E_L - v_P^i) + g_{\text{Input}P} (E_E - v_P^i) - g_{I_P} (E_I - v_P^i) \\
\frac{d}{dt} v_K^i = g_L (E_L - v_K^i) - g_{\text{APL}K} (E_I - v_K^i) + g_{\text{Input}K} (E_E - v_K^i) - g_{\text{Input}A} (E_A - v_K^i)
\]
Table 1. Network simulation parameters.

<table>
<thead>
<tr>
<th>Neuron parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacitance ORN</td>
<td>Cm = 100 pF</td>
</tr>
<tr>
<td>Capacitance PN</td>
<td>Cm = 30 pF</td>
</tr>
<tr>
<td>Capacitance LN</td>
<td>Cm = 50 pF</td>
</tr>
<tr>
<td>Capacitance APL</td>
<td>Cm = 200 pF</td>
</tr>
<tr>
<td>Leak conductance ORN</td>
<td>gL = 5 nS</td>
</tr>
<tr>
<td>Leak conductance PN and LN</td>
<td>gL = 2.5 nS</td>
</tr>
<tr>
<td>Leak conductance APL</td>
<td>gL = 5 nS</td>
</tr>
<tr>
<td>Leak potential ORN</td>
<td>EL = 60 mV</td>
</tr>
<tr>
<td>Leak potential PN and LN</td>
<td>EL = 60 mV</td>
</tr>
<tr>
<td>Leak potential KC and APL</td>
<td>EL = 60 mV</td>
</tr>
<tr>
<td>Threshold potential ORN and KC</td>
<td>Vth = -35 mV</td>
</tr>
<tr>
<td>Threshold potential PN and LN</td>
<td>Vth = -30 mV</td>
</tr>
<tr>
<td>Threshold potential APL</td>
<td>Vth = -30 mV</td>
</tr>
<tr>
<td>Resting potential ORN and LN</td>
<td>VR = -60 mV</td>
</tr>
<tr>
<td>Resting potential PN</td>
<td>VR = -60 mV</td>
</tr>
<tr>
<td>Resting potential KC</td>
<td>VR = -55 mV</td>
</tr>
<tr>
<td>Resting potential APL</td>
<td>VR = -60 mV</td>
</tr>
<tr>
<td>Refractory time</td>
<td>τref = 2 ms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Synaptic parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excitatory potential</td>
<td>EL = 0 mV</td>
</tr>
<tr>
<td>Inhibitory potential</td>
<td>EI = -75 mV</td>
</tr>
<tr>
<td>Excitatory time constant</td>
<td>τe = 5 ms</td>
</tr>
<tr>
<td>Inhibitory time constant</td>
<td>τi = 10 ms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Synaptic weights</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight input-ORN</td>
<td>wORNinputORN = 3 nS</td>
</tr>
<tr>
<td>Weight ORN–PN</td>
<td>wORNPN = 30 nS</td>
</tr>
<tr>
<td>Weight ORN–LN</td>
<td>wORNLN = 9 nS</td>
</tr>
<tr>
<td>Weight LN–PN</td>
<td>wLNPN = 2 nS</td>
</tr>
<tr>
<td>Weight PN–KC</td>
<td>wPNKC = 1 nS</td>
</tr>
<tr>
<td>Weight KC–APL</td>
<td>wKCAPL = 50 nS</td>
</tr>
<tr>
<td>Weight APL–KC</td>
<td>wAPLKC = 100 nS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adaptation parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptation time constant</td>
<td>τla = 1000 ms</td>
</tr>
<tr>
<td>Adaptation reversal potential</td>
<td>ELa = -90 mV</td>
</tr>
</tbody>
</table>

\[
C_m \frac{d}{dt} v_A = g_L(E_L - v_A) + g_{KAPL}(E_E - v_A)
\]

where ORNs (equation (1)) and KCs (equation (4)) are equipped with an additional spike-triggered adaptation (equation (6)) where \(g_{la}\) is the adaptation conductance and \(\tau_{la}\) is the decay time constant. With every spike \(g_{la}\) is increased in ORNs and KCs by 0.1 nS and 0.05 nS, respectively.

\[
\frac{d}{dt} g_{la} = - \frac{g_{la}}{\tau_{la}}
\]

Note, that the neuron model used in our computer simulations is the widely used conductance based leaky integrate-and-fire neuron [68] with an additional adaptation conductance in ORNs and KCs. This model does not match perfectly well the silicon neuron physically implemented on the DYNAP-SE board, which can be modeled by a current-based adaptive exponential integrate-and-fire model [65] (see Discussion). All code for the software implementation is accessible via https://github.com/nawrotlab/DrosophilaOlfactorySparseCoding.

2.4. Spontaneous activity

The input to the ORNs in our network model was modeled as stochastic point process realizations. It mimics the sum of spontaneous receptor activation and odor driven activation of the ORNs. On the chip, each ORN received a Poisson input to achieve a baseline firing rate \(\approx 5\) Hz. In the simulation each ORN received excitatory synaptic input modeled as a gamma process (shape parameter \(k = 3\)) to generate a similar baseline rate. The spontaneous firing rate of larval ORNs was previously measured in the range of 0.2–7.9 Hz, depending strongly on receptor type and odor identity [69, 70]. On the chip we measured a spontaneous ORN firing
rate of $6.2 \pm 3.0$ Hz. In the simulated model the average ORN baseline activity was estimated as $6.0 \pm 1.4$ Hz. Thus, ORNs on chip and in the simulation exhibit a similar spontaneous activity in the upper range of the empirical distribution.

### 2.5. Odor stimulation protocol

On the chip and in the computer simulation we included a warm up time (1.5 s and 0.3 s, respectively), which was excluded from the analyses. On the chip this restored the baseline biases following odor application. In the computer simulation this warm up period ensured that neuronal membranes and conductances were more heterogeneous at the beginning of the experiments.

We used a set of three different odors to study the effect of odor similarity. Figure 1(B) shows the activation profile (point process intensities) and overlap of all three odors across the 21 input channels. For each odor, the profile indicates the ORN-type specific activation level, mimicking the fact that each ORN expresses a genetically different receptor type. Similarity of odors is represented in the overlapping activation where odor 1 and odor 3 are distant (zero overlap), while odor 2 is constructed to have the same amount of overlap with the two other odors. The stimulation protocol assumes a 2 s odor stimulus on top of the baseline input with an activation rate according to figure 1(B).

### 2.6. Data analysis

#### 2.6.1. Sparseness measure

Sparseness was quantified by the widely used modified version \cite{71} of the Treves–Rolls measure \cite{72}.

$$S = 1 - \left( \frac{1}{N} \sum_{i=1}^{N} a_i \right)^2$$

where $a_i$ indicates either the spike count of neuron $i$ (population sparseness, $S_{\text{pop}}$), or the binned ($\Delta t = 20$ ms) population spike count (temporal sparseness, $S_{\text{pop}}$) for the 2 s with odor stimulation. $S$ assumes values between zero and one, with high values indicating sparse responses. This measure has been repeatedly used to quantify sparseness in insect olfactory processing \cite{36, 47, 52, 54, 73–77}. We report the average and standard deviation across the three odors. We then tested the effect of excluding specific sparseness mechanisms. To test for significance of the effects of lateral inhibition and SFA, the condition with only lateral inhibition enabled was compared with the condition with only SFA present (LN vs SFA) using a t-test for related samples. To test the effect of feedback inhibition via the APL, the condition including all mechanisms (LN + APL + SFA) was compared with LN + SFA. Tests were performed independently for temporal and population sparseness.

#### 2.6.2. Activation measure

We define the additional measure of activation as

$$A = \frac{1}{N \cdot k} \sum_{i=1}^{N} \sum_{j=1}^{k} \Theta(a_{ij})$$

where $a_{ij}$ indicates the spike count of neuron $i$ in the time bin $j$ and $\Theta$ is the Heaviside step function. Thus, $\Theta(a_{ij})$ indicates the binary response of neuron $i$ in time bin $j$. To assess population activation $A_{\text{pop}}$ we apply a single time bin for the complete 2 s odor stimulation time. Then $A_{\text{pop}}$ measures the fraction across all $N$ neurons that are odor-activated by at least one single spike. We quantify temporal activation $A_{\text{tmp}}$ by binning the stimulus time into $k = 20$ bins of $w = 100$ ms. Thus, $A_{\text{tmp}}$ measures the binary response probability across time bins for each neuron. Our definition of activation is related to the complementary measure of ‘activity sparseness’ defined in \cite{71}. Both measures were then averaged over all 20 trials. As results we report the average and standard deviation across the three odors.

#### 2.6.3. Distance measure

To assess the differences in odor distance between sparse and dense KC odor code we used the cosine distance (equation (9)). Vectors $a$ and $b$ each represent the average number of spikes evoked by all 72 KCs during the second odor presentation across 20 independent model instances. Cosine distance between $a$ and $b$ was calculated as:

$$D_{\text{cos}} = 1 - \frac{\sum_{i=1}^{n} a_i \cdot b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} \cdot \sqrt{\sum_{i=1}^{n} b_i^2}}$$

#### 2.6.4. Correlation across sparseness conditions

To test for qualitatively similar effects of the different sparseness conditions on the chip and simulation we correlated the results across the six data points (sparseness conditions) between the chip and the simulation.
For significance testing we generated 100 random unique permutations of the means from the simulation and correlated these 100 data series with that of the chip (LN + APL + SFA, LN + APL, APL + SFA, LN + SFA, LN, SFA; figure 3). The average of these 100 correlations was \( 0.07 (sd = 0.42) \) for \( S_{\text{pop}} \) and \( 0.01 (sd = 0.52) \) for \( S_{\text{imp}} \). In both cases the distribution of correlations was normal, established using D’Agostino-Pearson test for normality. The average of these 100 correlations each was used to evaluate the similarity of the effects on the chip and in the simulation.

3. Results

The larval nervous system with its limited neural network size and low complexity lends itself to the emulation on neuromorphic hardware. We analyzed a single hemisphere olfactory network model of the first instar *Drosophila* larva with <200 neurons and <1000 synapses comparing an implementation on the neuromorphic hardware DYNAP-SE [60] with a computer simulation of the same network. We were particularly interested in the contribution of different cellular and circuit mechanisms to the transformation of a dense input pattern at the periphery into a sparse odor representation in the MB.

3.1. Olfactory pathway model

Our spiking neural network model comprises four computational layers (figure 1(A)). Its structure, the size of the neuron populations and their connectivity are based on the connectome of a single right hemisphere as reconstructed from electron-microscopic data of one individual *Drosophila* larva MB by Eichler and colleagues [62]. Peripheral processing is carried out by 21 ORNs, each expressing a different olfactory receptor type [63, 78]. ORNs make one-to-one excitatory connections with 21 PNs and with 21 LNs that together constitute the antennal lobe. Each LN forms inhibitory synapses onto all PNs, establishing lateral inhibition. The PNs make divergent random connections with a total of 72 KCs, the primary cells of the MB, where each KC receives excitatory input from 1–6 PNs. The APL receives input from all of the matured KCs [64]. All KCs with a well-developed dendrite [62] fall into this category and those are the only ones included in our circuit model. We therefore assume a dense convergent connectivity with essentially all presynaptic KCs (in our case 64 out of 72 due to technical limitations on the chip, see Methods). We further implemented inhibitory feedback from the APL onto all KCs [64]. Overall, this blueprint of the olfactory network is highly similar to that in the adult fly except for the smaller neuron numbers and reduced anatomical complexity (see Discussion). Each model instance implemented here utilizes the exact same connectivities. We thus simulate a single individual rather than an average animal.

3.2. Circuit motifs and cellular adaptation

Our network model utilizes different cellular and circuit mechanisms that have been suggested to support a sparse code in the insect MB. To this end, the network topology includes three relevant motifs. First, the LN connectivity in the antennal lobe constitutes lateral inhibition as a motif that generally enhances neural contrast [34, 79] and that is implemented in the olfactory system of virtually all insects [36, 80–85], as well as in computational models thereof [25, 28, 52, 86]. Second, the random connectivity from PNs to a larger number of KCs is net divergent and sparse, expanding the dimensionality of the coding space [51, 87, 88]. Third, our model includes inhibitory feedback from the APL neuron onto all KCs. This has been shown to directly affect KC populations sparseness in the adult fly [47, 89] (see Discussion).

At the cellular level, all neurons in the network are modeled as leaky integrate-and-fire neurons. ORNs and KCs are equipped with a cellular SFA mechanism, a fundamental and ubiquitous mechanism in spiking neurons [34, 90]. ORNs have been shown to adapt during ongoing stimulation *in vivo*, both in larval [91] and adult [92, 93] *Drosophila*. The exact nature of the adaptation mechanism in the ORNs is still under investigation [92, 94, 95]. In KCs, a strong SFA conductance has conclusively been demonstrated in the cockroach [96] and the bee [97].

3.3. Dynamics of network response to odor stimulation

The response dynamics across all network stages to a single constant odor stimulation (figure 1(B)) with odor 1 is shown in figure 2(A) (chip) and figure 2(B) (simulation). At stimulus onset, a subset of all ORNs is activated according to the corresponding receptor response profile (figure 1(B), top). The ORN responses are phasic-tonic as a result of SFA with a higher firing rate at odor onset. The spike count histogram averaged across the 21 neurons of the ORN population fits the typical experimentally observed response profile observed in adult *Drosophila* [74, 92]. In the larva, little is known about stimulus adaptation in the ORNs [70]. The physical realization of SFA on chip is different from the simulation, which may partly explain the delayed response to odor onset- and offset of some neurons and the initially slower increase of the phasic response on the chip (figure 2(A), see Discussion). The off-response expressed in a prolonged silence of the odor-activated ORNs...
Figure 2. Dynamic network response. Network response to a stimulation with odor 1 for the chip (A) and the simulation (B). The odor was presented for 2 s, preceded by a 2 s baseline and followed by 2 s again without odor. Warmup times are excluded and only the time window between 1 and 3 s is shown here. Odor onset is at time = 0 and odor presence in the stimulation protocol is indicated by the shaded area throughout. Each dot denotes a single spike event of the respective neuron during an individual exemplary experiment. The lower panels (A) and (B) for each neuron population displays the averaged population spike count (across 20 trials) with a bin width of 100 ms.

in the simulation is an effect of SFA: the integrated adaptation current that has reached a steady state during the odor stimulation period now decays only slowly, acting in a hyperpolarizing fashion and thus reducing spiking probability [52] of the ORNs. This effect is barely visible and delayed on the chip (see Discussion). At the level of the antennal lobe both PNs (dark blue) and LNs (light blue) are excited only by the ORNs and thus follow their phasic-tonic response behavior and exhibit an inhibited off-response (figure 2), although neither neuron type is adaptive itself. The spatio-temporal response pattern of the PNs and LNs resembles the typical response pattern measured in vivo in adult flies and bees [81, 98, 99], including an inhibitory off-response in many neurons [92, 100, 101].

The KCs show very little spiking during spontaneous activity on the chip and in simulation. Only very few KCs do respond to odor stimulation (population sparse response) with only a single or few spikes (temporally sparse response). Spontaneous activity and response properties match well the in vivo situation as observed in various species [36, 54, 58]. The population spike count indicates a very brief population response within the first 100 ms, while the tonic KC response remains only slightly above the spontaneous activity level (cf [54]. Finally, the single APL driven by the excitatory KC population follows the brief phasic and weak tonic response of the KCs.

3.4. Analysis of sparsening factors in space and time
We investigate the translation from the peripheral dense code in the ORN and PN population into a central sparse code in the KC population, disentangling the contribution of the three fundamental biological mechanisms: cellular adaptation (SFA), lateral inhibition in the AL, and feedback inhibition in the MB. We
Figure 3. Mechanisms underlying population and temporal sparseness in KCs. (A) Comparison of KC population sparseness $S_{\text{pop}}$ between the chip and the simulation during 2 s with odor stimulation. The data is averaged over 20 experiments and three odors (error bars denote standard deviations across odors). Six (eight) experimental conditions were tested with a different set of sparseness mechanisms enabled. The respective mechanisms are listed below with LN (lateral inhibition via LNs), SFA, APL (feedback inhibition via the APL) and none (neither of the three). E.g., ‘LN + SFA’ denotes the presence of SFA and lateral inhibition. Data for the conditions ‘APL’ and ‘none’ only exists for the simulation (grey bars). (B) Temporal sparseness $S_{\text{tmp}}$ was computed for the same set of conditions. (C) Comparison of KC population activation $A_{\text{pop}}$ between the chip and the simulation during the 2 s odor stimulation, averaged over 20 trials and three odors (error bars denote standard deviation across odors). (D) Temporal activation $A_{\text{tmp}}$ was computed by averaging the number of 100 ms time windows in which each KC was active.

systematically varied the composition of the three mechanisms in our network, yielding five different conditions (figure 3) in which either one or two mechanisms were deactivated. SFA was only deactivated at the KC level and still present in ORNs. We did not vary the PN–KC connectivity pattern as this is identical to the anatomical pattern reported for the individual animal that we used as a reference.

We quantified the population activation by measuring the fraction of stimulus-activated KCs across the different conditions (see Methods) and find that it depends on the sparseness mechanisms. It is lowest in the control condition with 20.6(28.6%) responding neurons on the chip and 16.7(22.9%) in the simulation (figure 3(C)). Our results show that APL is the single crucial mechanism necessary for establishing a high population sparseness in our model. All conditions that lack feedback inhibition show strongly reduced values of $S_{\text{pop}}$. Lateral inhibition can to some degree recover sparseness on the chip and in the simulation.

We now consider temporal sparseness, which again reached high values in the control condition on the order of $S_{\text{tmp}} \approx 0.8$ (hatched bars in figure 3(B)). Comparing the different conditions we find that APL feedback inhibition and SFA in the KCs have a strong supporting effect for temporal sparseness. Any condition that involves the APL reached similar high values for $S_{\text{tmp}}$. Without the APL, SFA can partially ensure temporal sparseness on the chip and in the simulation. This is also reflected in the temporal activation measure that computes the fraction of active time bins (of 100 ms duration) for the complete 2 s stimulation time (see Methods). The results shown in figure 3(D) mirror our results in figure 3(B). In the sparse control condition KCs are active in on average only 2.3% and 3.4% of the response bins for the chip and the simulation, respectively.

Overall, we observed the same mechanistic effects on chip and in the simulation for the different combinations of activated and inactivated mechanisms (figure 3). The pattern of sparseness values across all six conditions is highly and significantly correlated between the chip and the simulation results, both for $S_{\text{pop}}$ and $S_{\text{tmp}}$ with maximum correlations of 0.91 and 0.87 for $S_{\text{pop}}$ and $S_{\text{tmp}}$ across 100 permutations, respectively.

3.5. Sparse representation supports stimulus separation

How does the encoding of different odors at the KC level compare between the sparse control condition and a non-sparse condition? Feedback and lateral inhibition supported population sparseness in the KC population. We thus compared the control condition to the network in which both inhibitory mechanisms were disabled by quantifying the pairwise distance between KC stimulus response patterns for any two different odors. Figure 4 shows the response rates averaged over the 2 s stimulus duration for the three different stimuli for both chip (figures 4(A)–(C)) and simulation figures 4(D)–(F). Only a fraction of the KCs responded to any odor ($S_{\text{pop}} > 0.8$, figures 3(A) and (E)) in the sparse condition. However, when feedback and lateral inhibition are disabled, essentially all KCs showed an odor response to any of the three odors (figures 4(C) and (F)).
Figure 4. Response pattern overlap. Average spike frequency (over 20 trials) for Chip ((A), (B) and (C)) and simulation ((D), (E) and (F)) in response to three different odors. The odors were presented for 2 s (for information on the experimental protocol please refer to figure 1(B)). All panels display the overlap between the different odor representations either on PN level ((A) and (D)), KC ((B) and (E)) and non-population-sparse KC in a condition with only SFA enabled ((C) and (F)). Overlap indicates a low ability to differentiate between odors.

A similar result is obtained when looking at cosine distances between KC odor representations. Independently of the odor identities, average pairwise cosine distance was considerably larger in the sparse condition (chip: 0.39(0.2); simulation: 0.85(0.06)) than in the non-sparse (SFA only) condition (chip: 0.07(0.02); simulation: 0.31(0.09)), indicating a similar effect of population sparseness on odor discriminability on the chip and in the simulation.

4. Discussion

In the present manuscript we addressed two major questions. First, we asked whether the re-coding from a dense peripheral olfactory code into a sparse central brain representation of odors can be achieved in the small spiking neural network model of *Drosophila* larva. To this end we tested the relevance of three fundamental mechanisms in establishing population and temporal sparseness:

- cellular adaptation
- lateral inhibition
- feedback inhibition

Second, we explored the feasibility of applying this coding scheme on real-time analogue neuromorphic hardware by comparing hardware implementation with software simulation at the relevant levels of stimulus encoding and processing.

4.1. Neuromorphic implementation versus computer simulation

Our results show that the on-chip network implementation achieved the transformation from dense to sparse coding in space and time. We obtained the same general results on the chip and in simulation albeit small differences. What are possible factors contributing to these differences?

First, while the software simulation used identical parameters for all neurons and synapses in a given population, there is considerable heterogeneity across the physical hardware implementation due to device mismatch, which particularly affects currents and conductances [4, 102, 103]. This heterogeneity is manifest e.g. in spiking thresholds, postsynaptic current amplitudes and membrane time constants. The neuromorphic hardware heterogeneity generally matches the biological heterogeneity that is typically ignored in computer-based simulations.

Second, setting the neuron and synapse parameters is straightforward and exact in the computer simulation. On the chip, however, this requires the adjustment of various biasing currents. As a result, real parameters will differ from theoretical target parameters and across circuits, as well as after re-adjustment in the same circuit.
Third, we have used different neuron models in the hardware emulation and in the computer simulation. Thus there is no one-to-one correspondence of the biophysical neuron parameters in the software (table 1) and the set points of the electronic circuits. In an effort to validate the robustness of the architecture to the model details, we deliberately did not minimize this difference, for example by employing hardware-matching neuron models designed to mimic the electronic circuit of the DYNAP-SE (https://code.ini.uzh.ch/yigit/dynapse-simulator.git).

Our research goal in this study was to assess the robustness of function in a small neural network architecture that is supported by three specific cellular and circuit mechanisms. To this end we tested its implementation on the DYNAP-SE neuromorphic hardware in light and despite of the various differences between the exact computer simulation of homogeneous elements and the real-time processing on electronic hardware with inhomogeneous devices. Taking this perspective, the differences between hardware and software implementation strengthen the conclusion that the suggested mechanisms are robust in supporting population sparse and temporally sparse stimulus encoding.

There are a number of advantages and disadvantages in using the specific hardware solution tested here. The fact that the DYNAP-SE [60] operates in real-time makes it suitable for the spiking control of autonomous robots [104, 105] and renders computational speed independent of network size. Even for the small larval network considered here (exactly 136 neurons and 833 synapses) simulations were several times slower than real-time with 3.8 s simulation time per 1 s biological time at a resolution of 0.1 ns (single core CPU, 64 bit PC, Ubuntu 18.04.5). Simulation time can be sped up to meet real-time demand even for large network sizes on specialized systems [106, 107].

A challenge with the mixed-signal neuromorphic hardware was the sensitivity of the circuit bias currents to noise and temperature changes, and the real-time nature of the experiment emulation. As each experiment would require the real time evolution of the input patterns and of the network dynamics to produce its response, this led to complex and lengthy experiments. As there were different experimental conditions, with three different odor stimuli, each with 20 trials, a particular challenge was the time-consuming adjustment of the SFA time constant on the chip since it required post hoc estimation of the effective time constant during repeated spike recordings. We therefore made the a priori choice to restrict the hardware emulation experiments to only six out of eight experimental conditions (figure 3). Two more conditions in which none of the mechanisms or only the feedback inhibition via APL was active were tested in the computer simulation only (grey bars in figure 3). Still, the variability across model instances was only slightly larger on the chip than in the simulation (figure 3). In addition, new neuromorphic circuit designs will be able to compensate for these drifts by using appropriate temperature compensated bias generator circuits [108].

4.2. Mechanisms and function of population sparseness
Population sparseness at the KC level has been demonstrated for a number of species in the adult stage (see Introduction). Our model suggests that given the current knowledge of anatomical structures within the *Drosophila* larval olfactory pathway it might already be implemented at this stage with similar benefits. Different mechanisms have been suggested for the generation of population sparseness. A fundamental anatomical basis for a sparse code is the sparse and divergent connectivity between PNs and a much larger population of KCs [37, 52]. Each KC receives input from only a few PNs and thus establishes a projection from a lower into a higher dimensional space, ideally suited to generate distinct activity patterns that foster associative memory formation. Additionally, there is evidence for a low excitability of the individual KCs that require collective input from several PNs to be activated [36, 37, 96]. Connectivity in our model is based on the exact numbers from electron microscopic reconstructions of neurons and synapses in the right hemisphere of one individual brain [62]. We did not attempt to adjust excitability of KCs or PN::KC connection strength for optimal population sparseness.

Feedback inhibition has repeatedly been suggested to underlie population sparseness in several animals, including the fly larva [109]. Empirical evidence has been provided in particular in bees [100, 110] and adult flies [47, 89]. Several modeling studies have used feedback inhibition to support a sparse KC population code in larger adult KC populations [25, 27, 28, 111]. Indeed, our study shows that inhibitory feedback from the single APL neuron effectively implements a sparse code in the small population of 72 KCs (figure 3(A)). We chose to model the APL as a spiking neuron that receives input solely from KCs and inhibits KCs in a closed loop. This decision was based on experimental evidence indicating a clear polarity of the APL with input in the MB lobes and pre-synaptic densities in the calyx, presumably onto the KC dendrites [64]. Whether the APL neuron generates sodium action potentials, however, is not clear in the larva [109] and has been challenged in the adult [47]. In addition, inhibitory feedback connections within the MB have been implicated in learning through inhibitory plasticity in bees and flies, thereby modulating the sparse KC population code [47, 110, 111].
As a third factor, lateral inhibition within the Drosophila antennal lobe has been shown to increase population sparseness at the KC level [74, 112] and in a model thereof [52]. This model study showed a strong effect of lateral inhibition on population sparseness in a network tuned to the anatomy of the adult fly. In the present larval model we found a supportive effect. With lateral inhibition alone the model reached \( S_{\text{pop}} \approx 0.6 \). The interplay of feedback inhibition and lateral inhibition boosted population sparseness to \( S_{\text{pop}} \approx 0.8 \) (figure 3(A)). This observation is different from our previous results in a network simulation modeled after the adult fly [28] where lateral inhibition in the AL was sufficient to implement a high population sparseness and APL feedback inhibition had a mainly supporting effect. The fact that lateral inhibition is less effective in the larval than in the adult Drosophila model [52] is thus likely due to the one-to-one connectivity between the 21 ORNs and 21 PNs in the larva, which requires very strong excitatory synapses. This specific configuration establishes a dominant feed-forward component in the larval olfactory pathway (figure 1(A)).

Sparse stimulus representation across the neuronal population supports minimal overlap of and correlation across stimulus-specific spatial response patterns [34, 52, 89, 113, 114], which in turn benefits associative memory formation and increases memory capacity [47, 53, 72, 115]. We confirmed an increased inter-stimulus distance in the KC coding space on the chip and in the simulation when all sparseness mechanisms take effect.

4.3. Mechanisms and function of temporal sparseness

Temporal sparseness in the insect MB has been physiologically described in various species. It is expressed in a highly phasic stimulus response that typically consists of only a single or very few spikes and that is temporally locked to stimulus onset or to a fast transient increase in stimulus amplitude while the tonic stimulus response is almost absent [36, 54, 58]. In our model we implemented two mechanisms that can support temporal sparseness, inhibitory feedback via the spiking APL neuron and SFA. Our analysis in figure 3(B) showed that inhibitory feedback has the strongest effect, confirming experimental [77, 116] and modeling results [25, 27, 28, 117]. Cellular adaptation (SFA) showed a smaller but supporting effect in our network, which is partially in line with our previous models of the adult fly [28, 52, 77] in which we showed that SFA alone can suffice to generate high temporal sparseness.

Importantly, cellular adaptation has additional effects on stimulus coding that are not analyzed here. Being a self-inhibiting mechanism it reduces overall spiking activity, contributing to the low spontaneous and response rates in the KC population that has been repeatedly documented in various insect species [36, 54, 58]. Moreover, SFA leads to a regularization of the neuron’s spike output and a reduction of the trial-to-trial variability, effectively improving response reliability [77, 118]. Finally, SFA introduces a short-term stimulus memory expressed in the conductance state of the excited neuron population, which decays with the SFA time constant [52].

Temporal sparseness was influenced strongly by SFA in the KCs and by recurrent feedback inhibition. It usually shows as longer inter-spike-intervals both in physiological data [36, 54, 100] as well as in modeling results [28, 52, 77]. Besides the prolongation of the inter-spike-intervals over the entire duration of the experiment, SFA also caused the commonly observed odor onset effect [36, 54, 58, 100, 116] in ORNs and KCs. In our data this effect was somewhat concealed in the KCs by the overall small number of spike responses. This is a tribute to the biological plausibility with respect to data collected from adult Drosophila, where the KC rarely show spikes at baseline [58] and a very sparse odor response pattern [58, 119]. Due to the SFA in the ORN population that was active in all experimental conditions there was a good degree of temporal sparseness in the LN-only condition as well (especially on the chip). Again we chose to accept this effect as a baseline level of sparseness to compare other conditions against. In both implementations the expected effects of SFA in the KCs could be observed.

We have previously argued that the major functional role of temporal sparseness is the rapid and reliable stimulus encoding in a temporally dynamic environment [28, 77]. Indeed, temporal dynamics is high in the natural olfactory environment and depends on air movement and on animal speed, the latter being particularly high in flying insects. As a result, adult insects during flight or locomotion may encounter a rapid on-off stimulus scenario when passing through a thin odor filament [120–124]. It remains an open question whether the SFA mechanism is at all present in the KCs during larval stages and electrophysiological approaches to neural coding in the larva is scarce. Representation of high temporal stimulus dynamics is likely of minor importance for the larva as its locomotion is slow and the natural environment suitable for larval development such as e.g. a rotten fruit likely provides little olfactory dynamics. However, larva do perform chemotaxis and thus are able to sample olfactory gradients.

4.4. Outlook

Our current research extends the present model towards a plastic spiking network model of the larva that can perform associative learning and reward prediction [19, 125] inspired by recent modeling approaches in the adult [126, 127]. Together with biologically realistic modeling of individual larva locomotion and chemotactic
behavior [16] this will allow to reproduce behavioral [128–132] and optophysiological observations [64, 133, 134] and to generate testable hypothesis at the physiological and behavioral level. In the future this may inspire modeling virtual larvae exploring and adapting to their virtual environment in a closed loop scenario and the implementation of such mini brains on compact and low-power neuromorphic hardware for the spiking control of autonomous robots [7, 28, 135, 136].

Acknowledgments

This projects is funded by the German Research Foundation (DFG) within the Research Unit 'Structure, Plasticity and Behavioral Function of the Drosophila mushroom body' (DFG-FOR 2705, Grant No. 403329959, https://www.uni-goettingen.de/en/601524.html), and in part by the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Program Grant agreement No. 724295 (NeuroAgents). AMJ received additional travel support from the Research Training Group ‘Neural Circuit Analysis’ (DFG-RTG 1960, Grant No. 233866668). The authors would like to acknowledge the financial support of the CogniGron research center and the Ubbo Emmius Funds (Univ. of Groningen). We thank Sören Rüttger for support with the neuromorphic hardware setup, Hannes Rapp, Panagiotis Sakagiannis and Bertram Gerber for valuable discussions, and Albert Cardona for comments on the earlier bioRxiv version of this manuscript.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Anna-Maria Jürgensen https://orcid.org/0000-0002-7871-1887
Afshin Khalili https://orcid.org/0000-0001-5947-5551
Elisabetta Chicca https://orcid.org/0000-0002-5518-8990
Giacomo Indiveri https://orcid.org/0000-0002-7109-1689
Martin Paul Nawrot https://orcid.org/0000-0003-4133-6419

References

[16] Antoine W and Graham P 2012 What can we learn from studies of insect navigation? Anim. Behav. 84 13–20
A-M Jürgensen et al

[64] Saumweber T et al 2018 Functional architecture of reward learning in mushroom body extrinsic neurons of larval Drosophila Nat. Commun. 9 1104


[84] Letich B and Laurent G 1996 GABAergic synapses in the antennal lobe and mushroom body of the locust olfactory system J. Comp. Neurol. 372 487–514


[86] Linster C and Smith B H 1997 A computational model of the response of honey bee antennal lobe circuitry to odor mixtures: overshadowing, blocking and unblocking can arise from lateral inhibition Behav. Brain Res. 87 1–14


[93] Gourab Shandilya S, Demir M, Long I, Clark D A and Emonet T 2017 Olfactory receptor neurons use gain control and complementary kinetics to encode intermittent odorant stimuli eLife 6 e27670


[99] Meyer A, Galizia C G and Nawrot M P 2013 Local interneurons and projection neurons in the antennal lobe from a spiking point of view J. Neurophysiol. 110 2465–74


[101] Nawrot M P 2012 Dynamics of sensory processing in the dual olfactory pathway of the honeybee Apidologie 43 269–91

[102] Indiveri G et al 2011 Neuromorphic silicon neuron circuits Front. Neurosci. 5 73


A-M Jürgensen et al


