Unsupervised Learning of a Dictionary of Neural Impulse Responses from Spiking Data*

Bahareh Tolooshams[1], Hao Wu[2], Paul Masset[2], Venkatesh N. Murthy[2], and Demba Ba[1]

[1]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA
[2]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA

Summary: The majority of approaches for studying neuronal activity use an experimental setup where a stimulus is repeatedly applied over a series of trials with time-locked and non-overlapping events. Then, the spike trains are averaged over trials and smoothed out. These approaches fail in naturalistic environments and experiments in which the stimulus comprises discrete events occurring at random times, which may elicit overlapping responses. To analyze neuronal activity patterns in such experiments, we utilize a model of the spiking rate of a neuron as the convolution of an unknown impulse response and a sparse code, representing the time when the stimulus elicits an activity pattern in the neurons response, and the response's amplitude. We fit the model to single-unit spiking data by solving a Poisson dictionary learning problem that lets us estimate a neurons impulse response, and the amplitude of the response to each stimulus, directly from the spiking data. To solve the problem, we construct an autoencoder. We used neural spiking data acquired from piriform cortex in response to odor pulses to estimate the impulse responses (dictionary) of ∼200 neurons along with the strength of the response associated with each pulse for each neuron (sparse code). The Kolmogorov-Smirnov (KS) test shows that the model fits the data well. Our analysis shows that, at the level of a single neuron, the odor pulses evoke different responses, likely reflecting differences in alignments to the breathing phase. In addition, we found that neurons from the population cluster according to either the estimated impulse responses or the stimulus responses, suggesting the presence of distinct neural populations in piriform cortex that could have distinct roles in the processing of information in olfactory search.

Experimental setup. An experimental session consists of ∼250 trials. In each trial, a custom device delivered 50 ms odor pulses (red dots in Fig. 2(a)) of the same peak concentration to the animal's nose at a Poisson-distributed pulse rate between 0.5-4 pulse/s for 5 s. Neural activity in the animal's anterior piriform cortex was recorded with a custom-built 32-channel tetrode drive at a 30 kHz sampling rate using the Open Ephys recording system [1]. Single-unit spiking activities (black dots in Fig. 2(a)) were isolated using Kilosort2 [2]. We isolated 5-40 single units in each recording session. At the end of each session, the entire bundle of tetrodes was lowered by 40 μm to obtain a new set of neurons for the subsequent session. We recorded $C = 388$ neurons during $S = 17$ behavioral sessions. We excluded neurons with very low and high firing rate and used $C = 221$ neurons for the subsequent analysis.

Methods. We downsampled the data to a resolution of 1 ms. Let $\mathbf{s}^{j,c} \in \{0,1\}^{4500}$ be the vector of spikes from neuron c in trial j from 0.5 to 5 s. We conducted our analysis using 50 ms bins. We denote the odor indicator by $\mathbf{o}^j \in \{0,1\}^{100}$. We model the spike counts $\mathbf{y}^{j,c} \in \mathbb{R}^{N=90}$ as a Poisson process (i.e., $\mathbf{y}^{j,c} \sim \text{Poisson}(\boldsymbol{\mu}^{j,c})$). Let $\mathbf{h}_c \in \mathbb{R}^{20}$ be a 1 s long dictionary element (impulse response) that characterizes the activity of neuron c, and $\mathbf{x}^{j,c}$ a sparse vector whose non-zero entries and amplitudes indicate the presence of an odor pulse and modulate the strength of neural response, respectively. We model the spiking rate as $\boldsymbol{\mu}^{j,c} = \exp\left(\mathbf{h}_c * \mathbf{x}^{j,c} + a^{j,c}\right)$ where $a^{j,c}$ determines the baseline activity of neuron c which we heuristically estimate from the trials. The sparsity of the occurrences of odor pulses motivates us to enforce sparsity on $\mathbf{x}^{j,c}$ using the ℓ_1 norm. Our goal is to learn an impulse response $\mathbf{h}_c$ (a convolutional dictionary with one filter) for each neuron by minimizing the negative log-likelihood $-\sum_{j=1}^{J} \log P(\mathbf{y}^{j,c} \mid \mathbf{h}_c, \mathbf{x}^{j,c})$. We optimize the following objective

$$\min_{\mathbf{h}_c, \{\mathbf{x}^{j,c}\}_{j=1}^{J}} \quad \sum_{j=1}^{J} -(\mathbf{h}_c * \mathbf{x}^{j,c} + a^{j,c})^{\mathrm{T}} \mathbf{y}^{j,c} + \mathbf{1}^N \mathbf{e}^{\mathbf{h}_c * \mathbf{x}^{j,c} + a^{j,c}} + \lambda \|\mathbf{x}^{j,c}\|_1 \qquad \text{s.t. } \|\mathbf{h}_c\|_2 = 1,\ \mathbf{x}^{j,c} \geq 0. \tag{1}$$

*This work is accepted to *Computational and Systems Neuroscience, 2021.*

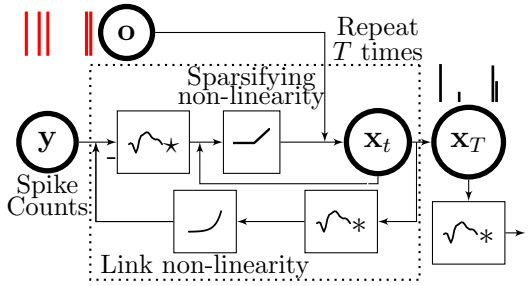

Figure 1: DCEA architecture.

This problem can be solved by alternating between sparse coding (i.e., solving for $\mathbf{x}^{j,c}$ given an estimate of the dictionary), and a dictionary update step (i.e., learn a better $\mathbf{h}_c$ given the estimated codes). Recent work [3] has shown a one-to-one correspondence between (1) and deep convolutional exponential-family autoencoder (DCEA), shown in Fig. 1. The encoder performs sparse coding (i.e., estimates the odor onsets and the amplitude of response), and the decoder reconstructs the rate $\boldsymbol{\mu}^{j,c}$ using $\mathbf{h}_c$, which is learned by backpropagation. In our results below, we use the knowledge of the support for the code (i.e., timing of the pulses is known). Hence, the encoder unfolds T iterations of proximal gradient descent on the objective which is $\mathbf{x}_t^{j,c} = \mathbf{o}^j \cdot \mathrm{ReLU}_{\frac{\lambda}{L}}\left(\mathbf{x}_{t-1}^{j,c} + \frac{1}{L}\mathbf{h}_c \star (\mathbf{y}^{j,c} - \mathbf{e}^{\mathbf{h}_c * \mathbf{x}^{j,c} + a^{j,c}})\right)$ where $\star$ is correlation operator. The amplitudes of non-zero elements in $\mathbf{x}_T^{j,c}$ determine the strength of neural response to the stimulus. We leave the case of unknown support for future work. We initialize $\mathbf{h}_c$ using the peristimulus time histogram (PSTH) of the aligned raster (Fig. 2(d)) which is shown in (Fig. 2(c)) in black.

Results. Fig. 2 shows the data and results for one neuron. Given $\mathbf{h}_c$ (Fig. 2(c) green), Fig. 2(b) shows spike counts $\mathbf{y}^{j,c}$ of one trial (black) and the estimated rate $\boldsymbol{\mu}^{j,c}$ (orange). Fig. 2(e) shows the odor onsets (red) and the estimated code (black) from a single trial. This figure shows that the neuron does not respond to all pulses with the same strength. For instance, the pulses around 4 s do not evoke a response. We may attribute this to the effect of spiking history or position of these pulses with respect to the animal's breathing cycle. The KS plot from Fig. 2(f) demonstrates that the model fits the data well. Below, we categorize the neurons into groups based on their impulse responses or sparse codes.

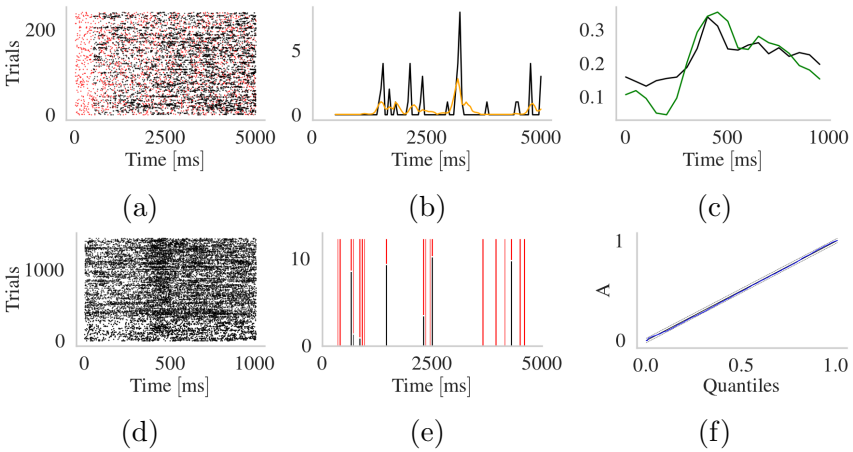

Figure 2: One neuron. (a) Odors (red) and spikes (black). (b) Spike counts (black) and estimated rate (orange) for a trial. (c) PSTH of aligned raster (black) and the dictionary (green). (d) Aligned raster given odor onsets. (e) Odor events (red) and the code (black) for a trial. (f) goodness-of-fit (dotted lines shows 95% confidence interval).

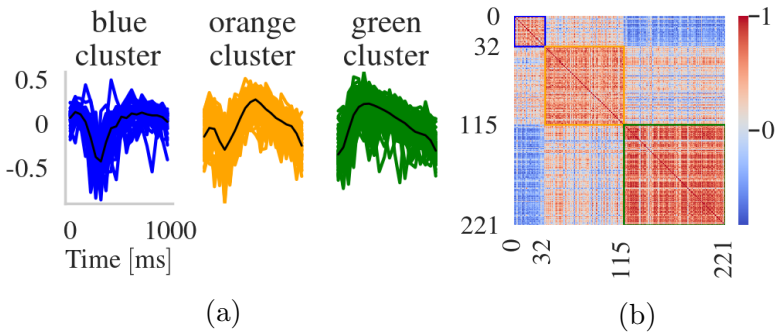

Figure 3: Dictionary (a) clusters and (b) similarity.

After learning impulse responses for all neurons, we standardized them (i.e., zero mean and norm one) and used spectral clustering to obtain three clusters. Fig. 3(a) visualizes the impulse responses for each cluster in a different color, where their representatives (mean) are shown in black. We observed mainly three types of responses: (i) one that starts to silence the neuron reaching minimal activity ∼300 ms after the odor release and then returns to baseline (blue cluster) (ii) another that results in an undershoot within ∼200 ms of an odor, followed by an increase in activity which decays after ∼500 ms (orange cluster), and (iii) finally one that shows a sharp increase in spiking activity and a decay after ∼300 ms (green cluster). Fig. 3(b) shows the cosine similarity of impulse responses; neurons within the identified clusters (inside the colored boxes) have high similarities, and neurons across clusters have a low similarity. Overall, the largest portion of the neurons come from the green cluster excited by the odor pulses. The smallest size cluster corresponds to neurons inhibited by the stimulus.

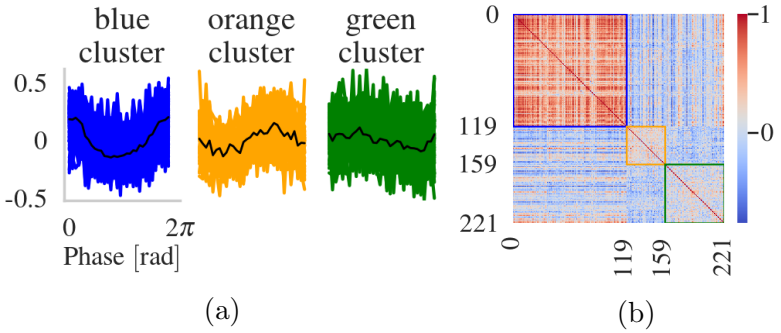

Figure 4: HBP (a) clusters and (b) similarity.

For each neuron, we extract sparse codes, $\mathbf{x}_T^{j,c}$, restricted to the support of odor pulses $\{\mathbf{o}^j\}_{j=1}^{221}$ from all trials. As the odor timing is different from one experiment session to another, we cannot compare the neurons across sessions given this vector. To enable comparison, we relate the code to the animal's breathing phase and construct a normalized weighted histogram of the breathing phase (HBP) from 0 to 2π using the code amplitudes. Given the standardized HBP, we then perform spectral clustering on HBP for the population of neurons from all experiment sessions. Fig. 4(a) presents the three clusters identified using HBP and Fig. 4(b) show its corresponding similarity matrix. These data show that different neurons are modulated by breathing phase in different ways, with one activation pattern (blue) dominating.

References

[1] Joshua H Siegle, Aarón Cuevas López, Yogi A Patel, Kirill Abramov, Shay Ohayon, and Jakob Voigts, "Open ephys: an open-source, plugin-based platform for multichannel electrophysiology," *Journal of Neural Engineering*, vol. 14, no. 4, pp. 045003, 2017.

[2] Marius Pachitariu, Nicholas A Steinmetz, Shabnam N Kadir, Matteo Carandini, and Kenneth D Harris, "Fast and accurate spike sorting of high-channel count probes with kilosort," in *Proc. Advances in Neural Information Processing Systems*, 2016, vol. 29.

[3] Bahareh Tolooshams, Andrew Song, Simona Temereanca, and Demba Ba, "Convolutional dictionary learning based auto-encoders for natural exponential-family distributions," in *Proc. the 37th International Conference on Machine Learning*, 2020, vol. 119, pp. 9493–9503.