Does the neuronal noise in cortex help generalization?

Anonymous Author(s)
Affiliation
Address
email

Abstract

Neural activity is highly variable in response to repeated stimuli. We used an open dataset, the Allen Brain Observatory, to quantify the distribution of responses to repeated natural movie presentations. A large fraction of responses are better fit by log-normal distributions or Gaussian mixtures with two components. These distributions are similar to those from units in deep neural networks with dropout. Using a separate set of electrophysiological recordings, we constructed a population coupling model as a control for state-dependent activity fluctuations and found that the model residuals also show non-Gaussian distributions. We then analyzed responses across trials from multiple sections of different movie clips and observed that the noise in cortex aligns better with in-clip versus out-of-clip stimulus variations. We argue that noise is useful for generalization when it moves along representations of different exemplars in-class, similar to the structure of cortical noise.

1 Introduction

One of the hallmarks of neural codes is the high level of trial-to-trial variability [1, 2]. This variability has been studied using multiple stimuli [3], along with its relation to attention [4] and other behavioral variables [5]. Previous theories on the possible role of noise center on its potential usefulness in inference [6]. In the field of machine learning, noise can have a regularizing effect and enable better model generalization (e.g. dropout [7]). Here, we explore the hypothesis that networks of cortical neurons use noise with the goal of building general representations from a small number of exemplars.

2 Results

Noise distribution in the Allen Brain Observatory [8]. We used the 30 s long "natural movie one" stimulus. The movie was presented 10 times over 3 imaging sessions ($N = 30$ trials). We analyzed all excitatory cells in the dataset, excluding cells with a mean trial-to-trial correlation below zero, which resulted in $N = 11,428$ cells. We split the movie into non-overlapping 1 s epochs, and computed the mean dF/F response for each cell over each epoch. The "preferred" stimulus for each cell was defined as the epoch which elicited the max mean dF/F response over all trials.

We analyzed trial-to-trial variability in neural responses to each cell's "preferred" stimulus across visual areas. We find that the majority of cells are best fit by log-normal distributions or Gaussian mixtures with two components, with 40.4% of cells showing dropout-like distributions (see Methods). We performed a similar analysis on units from a convolutional neural network[1] trained on CIFAR-10 [9], using each unit's "preferred" image. Dropout ($p = 0.5$) was used in all layers during training and inference, which may act as a form of Bayesian approximation [10]. We find that two-component

[1]Network architecture: conv5-10, conv5-10, conv5-20, conv5-20, fc-320, fc-50, where 'conv' represents convolutional layers (kernel size-features) and 'fc' represents fully connected layers (features)

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

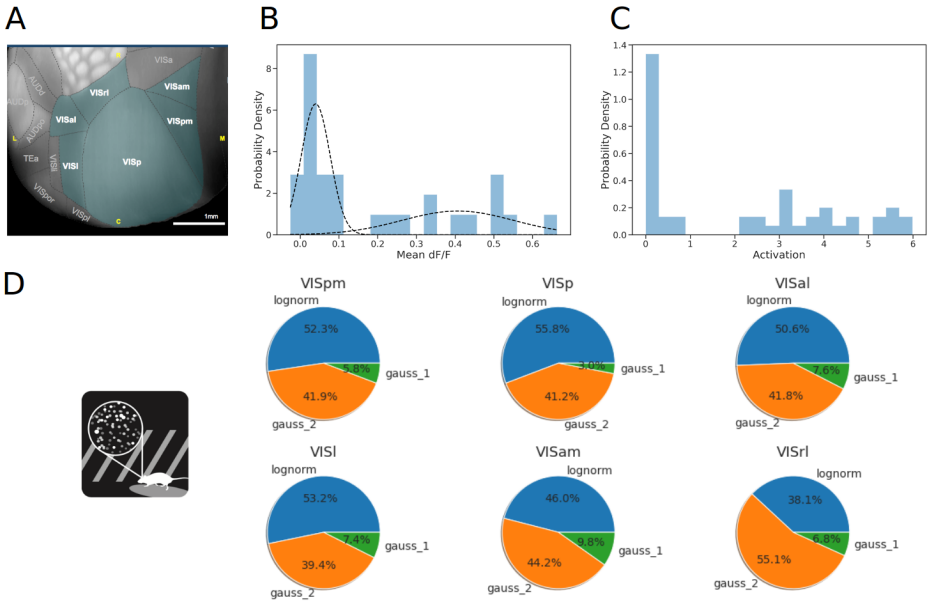

Figure 1: Neural variability in the Allen Brain Observatory. (A) Visual areas where recordings were performed. (B) The distribution of responses for an example cell. The two components of the Gaussian mixture are shown in dotted black lines. (C) The distribution of activations for an example unit in the neural network (second convolutional layer, fourth feature map). (D) Percent of cells in each of the recorded visual areas, broken down by the best-fitting distribution.

Gaussian mixtures can also capture the responses from the network with dropout. Figure 1 shows example response distributions, and a summary of the distribution fits across cells in our dataset.

Noise distribution and state dependence in electrophysiological recordings. In vivo recordings were performed in the visual cortex of awake, head-fixed mice using Neuropixels probes [11]. The repeated natural movie stimulus was 81 s long, consisting of 11 shorter clips ranging from 4 to 9 s each ($N = 98$ trials). All spike data were acquired with a 30-kHz sampling rate and recorded with the Open Ephys GUI. A 300-Hz analog high-pass filter was present in the Neuropixels probe, and a digital 300-Hz high-pass filter (3rd-order Butterworth) was applied offline prior to spike sorting. Spike times and waveforms were automatically extracted from the raw data using Kilosort2. After filtering out units with "noise" waveforms using a random forest classifier trained on manually annotated data, all remaining units were packaged into the Neurodata Without Borders format for further analysis. This resulted in a total of $N = 936$ units across three mice.

One potential source of variability is state-dependent changes in neural activity [5], which we control for by using a population coupling model (see Methods). We again analyzed trial-to-trial variability by fitting different distributions to the neural response residuals from this model. We find that the majority of cells are still better fit by either log-normal distributions or Gaussian mixtures, even when including Poisson and negative binomial distributions (Figure 2).

Trial-to-trial variability mimics in-class exemplar changes. We randomly choose 10 non-overlapping 200 ms long sections ("exemplars") from each of the 11 movie clips. We exclude the first 1 s following a clip transition to avoid onset transient effects. We define the activity for neuron i of exemplar j in clip k during trial n as the spike count during the presentation of the exemplar $a_{i,j,k,n}$. The signal for neuron i of exemplar j in clip k is the average over trials $s_{i,j,k} = \langle a_{i,j,k,n} \rangle_n$. The noise for neuron i of an exemplar j in clip k during trial n is the activity minus the signal $n_{i,j,k,n} = a_{i,j,k,n} - s_{i,j,k}$. We also define the following neural subspaces:

The **exemplar coding subspace** for an exemplar and clip is defined as the set of neurons for which the signal is larger than the mean, $E_{j,k} = \{i | s_{i,j,k} > \langle s_{i,j,k} \rangle_{j,k}\}$. The **clip coding subspace** is defined as the set of neurons for which the signal is larger than the mean for more than half

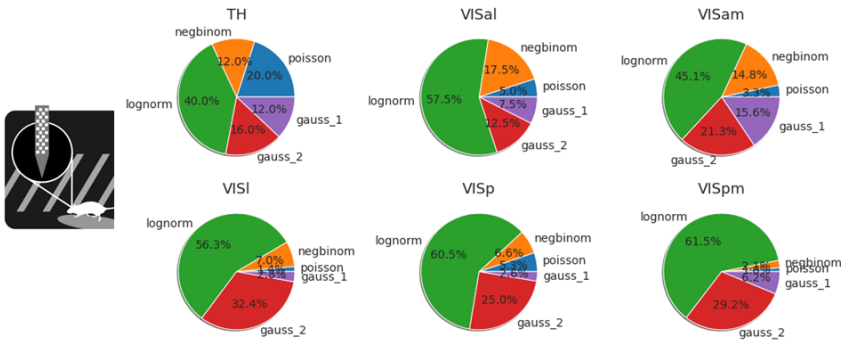

Figure 2: Neural variability in the electrophysiological recordings after controlling for state dependence. A large fraction of cells are still best fit by log-normal distributions or Gaussian mixtures.

the exemplars in the clip, $C_k = \{i | mean_j(s_{i,j,k} > \langle s_{i,j,k} \rangle_{j,k}) >= 0.5\}$. The **clip variance subspace** is defined as the set of neurons for which the variance of the signal for exemplars in-clip is larger than out-of-clip, $V_k = \{i | std_j(s_{i,j,k}) > std_{j,k}(s_{i,j,k})\}$. The **noise subspace** is defined as the set of neurons for which the absolute value of the noise is larger than its standard deviation, $N_{j,k,n} = \{i | abs(n_{i,j,k,n}) > std_{j,k,n}(n_{i,j,k,n})\}$.

Our hypothesis is that noise for an exemplar in clip k_1 should lie in the clip k_1 subspace, but not the clip k_2 subspace. We found that the distance between the noise subspace for exemplars in clip k_1 is smaller for the clip k_1 versus the the clip k_2 variance subspace, suggesting that noise aligns better with in-clip variations. We also found that noise aligns with the same clip coding subspace and the exemplar coding subspace (Figure 3). We computed these measures for each mouse and visual area with at least 20 reliable neurons (9 areas passed this threshold). The differences are in the right direction and statistically significant in each individual area analyzed ($p < 0.05$).

Relation to dropout and new avenues of machine learning research. Dropout has been shown to be an effective regularization technique that prevents model overfitting and reduces feature co-adaptation [7]. The non-Gaussian distributions we observed in the data inspired us to use a subspace analysis. As dropout-like noise generates projections in neuronal space, eliminating some neurons altogether, it is a natural place to focus on for the analysis of how trial-to-trial variability and noise aligns with different neural subspaces. Future work will study how different forms of subspace-aligned noise may help deep neural networks generalize better from fewer examples.

3 Discussion

In the first part of the paper, we observed complex, non-Gaussian distributions in the responses of neurons even for their preferred stimulus. In the second part of the paper, we found that trial-to-trial noise for an exemplar in a clip aligns better with exemplar-by-exemplar variation in the same clip than for other clips. We believe that research into the structure and role of biological noise will be useful for developing new methods to train neural networks with better generalization capabilities.

4 Methods

Noise distribution fitting. For each cell, we quantified the distribution of neural responses across trials by fitting different distributions using the *scikit-learn* package. We fit one- and two-component Gaussian mixtures, log-normal, Poisson, and negative binomial distributions. We only fit the Poisson and negative binomial distributions to the electrophysiological recording data since these require discrete count data. We used the Akaike information criterion (AIC) to select between model fits, although other information theoretic measures yielded qualitatively similar results. We used a bootstrap parametric cross-fitting test with $N = 10,000$ samples to determine the significance of the two-component Gaussian mixture model fits. For cells best fit by the two-component Gaussian mixture, we performed an additional test to determine whether their response distributions were

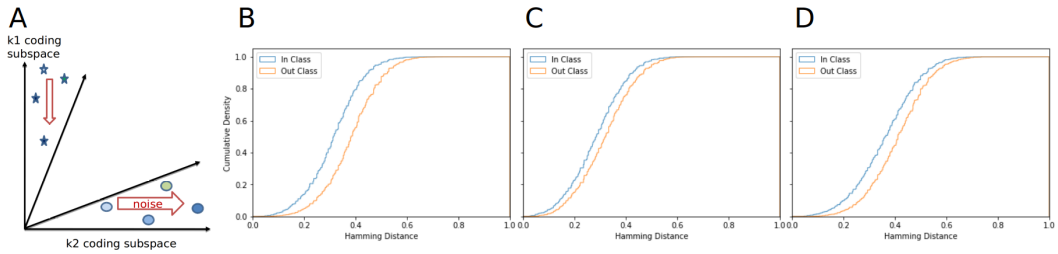

Figure 3: Noise aligns with in-clip variations. (A) Noise in cortex for an exemplar is aligned with the coding subspace for the clip it belongs to (clip coding subspace) and to the variations in the coding subspace between exemplars in the same clip (clip variance subspace). (B-D) Hamming distance between the noise subspace and the clip variance subspace (B), clip coding subspace, and exemplar coding subspace (D). Distances are lower in-class versus out-of-class, suggesting a shared subspace.

dropout-like. For each of these cells, we calculated a z-score on the component with the lower mean, and those cells with z-scores less than two (meaning their means are not significantly different than zero) were counted as cells with dropout-like response distributions.

Population coupling model. If neural variability is exclusively the result of state fluctuations, it should be captured by the coupling of each cell's activity to a lower-dimensional representation of population activity. We isolated each neuron and clustered the activity of the other neurons into 100 clusters. We used agglomerative clustering from the *scikit-learn* package with average linkage and the pairwise Pearson correlation coefficient of single-trial activities. The average activity of neurons within each cluster was used as predictors for the single-trial activity of the held-out neuron. For each neuron, we then fit a generalized linear model with the Gaussian family and an identity link function using the *statsmodels* package. We split the single-trial activities into two equal halves, using the first half for training and the second half for testing. The difference between the model predicted responses and the experimentally observed responses was used to calculate the residual activity for each neuron. The distribution of residuals was then fit using the methods outlined above.

References

[1] D J Tolhurst, J A Movshon, and A F Dean. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision research*, 23(8):775–85, 1983.

[2] M N Shadlen and W T Newsome. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of neuroscience*, 18(10):3870–96, 5 1998.

[3] Robbe L T Goris, J Anthony Movshon, and Eero P Simoncelli. Partitioning neuronal variability. *Nature Neuroscience*, 17(6):858–865, 6 2014.

[4] Tatiana A Engel, Nicholas A Steinmetz, et al. Selective modulation of cortical state during spatial attention. *Science*, 354(6316):1140–1144, 2016.

[5] Carsen Stringer, Marius Pachitariu, et al. Spontaneous behaviors drive multidimensional, brainwide activity. *Science (New York, N.Y.)*, 364(6437):255, 4 2019.

[6] Rodrigo Echeveste and Máté Lengyel. The Redemption of Noise: Inference with Neural Populations. *Trends in neurosciences*, 41(11):767–770, 2018.

[7] Nitish Srivastava, Geoffrey Hinton, et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 2014.

[8] Saskia E J de Vries, Jerome Lecoq, et al. A large-scale, standardized physiological survey reveals higher order coding throughout the mouse visual cortex. *bioRxiv*, page 359513, 6 2018.

[9] Alex Krizhevsky. Learning Multiple Layers of Features from Tiny Images. Technical report, 2009.

[10] Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. 6 2015.

[11] James J. Jun, Nicholas A. Steinmetz, et al. Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551(7679):232–236, 11 2017.