Functional Models of Mouse Visual Cortex

Ekaterina Taralova
Columbia University
New York, NY 10023
et2495@columbia.edu

Tony Jebara
Columbia University
New York, NY 10023
jebara@cs.columbia.edu

Rafael Yuste
Columbia University
New York, NY 10023
rmy5@columbia.edu

Abstract

We develop machine learning methods to decipher the neural code that links perception with the firing of neurons in the cerebral cortex. We deploy Bayesian graphical models for holistic neuron data analysis to provide a framework for understanding the neural circuitry, which is unattainable with current single neuron methods. The proposed models are data-driven and capture the probabilistic conditional dependencies between the neural activity and the visual stimuli.

1 Introduction

Despite a century of intense research, neuroscience still lacks computational models of visual stimuli and the spatio-temporal firing patterns they elicit in the visual system of animals or humans. Traditional imaging techniques like functional magnetic resonance imaging (fMRI) indirectly measure responses of large numbers of neurons (voxels). However, to understand how neural activity gives rise to cognition, we need to model how collections of hundreds of neurons engage in complex interactions over time. In contrast to whole-brain or micro scale imaging of a small number of neurons, new imaging methods like two-photon calcium imaging can record the firing of hundreds to thousands of neurons in living animals. Imaging the operation of neural circuits at this mesoscale will enable better understanding of the cortical information processing algorithms relating the activity of neurons to visual stimuli.

We have conducted targeted neurophysiological experiments that interrogate the operation of mesoscale cortical computing circuits in awake behaving mice using two-photon calcium imaging. We record spontaneous activity and activity evoked by a database of visual stimuli ranging from drifting gratings, wavelets, composite textures and complex scenes. Using this neurophysiology data we train Bayesian graphical models to capture the spatio-temporal functional dependencies between neurons. The result is a compact representation that links visual stimuli to neural activity. We validate the models in the tasks of encoding (predicting) the neural activity and decoding (classifying) the visual stimuli.

2 Dataset

We have collected Calcium image data from layer 2/3 in mouse visual cortex in an awake, behaving animal. The mouse was head-restrained on a rotary treadmill, while visual stimuli were presented and neuronal activity recorded. The mouse was expressing the genetically encoded calcium indicator GCaMP6 ”fast.” We recorded spontaneous activity and evoked activity. We see strongly conserved activity in visual cortex simply when presenting drifting gratings and textures, even without a behavioral readout. We use the spatiotemporal NMF source separation code developed by Pnevmatikakis et. al. to automatically extract regions of interest (ROIs) and spike events from the resulting time series [3, 4].

3 Graphical Models

The MAP inference problem and the computation of the partition function are NP-hard in general. As a result, approximations are often necessary in practice. The Bethe free energy is a standard approximation to the so-called Gibbs free energy that is motivated by ideas from statistical physics. The approximation has been generalized to include different counting numbers that result in alternative entropy approximations [5]. In particular, these counting numbers can be chosen such that approximation is convex. The popularity of the Bethe approximation is due mostly to the observation that there is a BP-like message-passing scheme whose fixed points correspond to stationary points of the reweighted free energy. When the approximation is convex, each fixed point corresponds to a global optimum, and we recover the typical Bethe free energy approximation. If the standard Bethe approximation is convex, then we can leverage tools from convex optimization in an attempt to minimize it. If the Bethe approximation is not convex, belief propagation or a convergent algorithm can be used in an attempt to find a local optimum of the Bethe approximation [7]. Alternatively, if the graphical model has a maximum degree at most $\mathcal{O}(\log n)$, then there exists a fully polynomial time approximation scheme (FPTAS) to compute a local minimum of the Bethe approximation [1]. The drawback of these algorithms is that, even if they converge, they are only guaranteed to find a local optimum of the Bethe approximation. In recent work [6], we show how we can design methods that provably find/approximate the optimum of the Bethe approximation in the non-convex case. We obtained a FPTAS for any attractive binary pairwise model without any restriction on topology, with much better running time guarantees. The method also applies to models that are not attractive, thereby reducing the approximation task to a multi-label MAP inference problem.

Given the binary spike events for each visual stimuli, we perform structure learning and inference to learn Markov Random Field models. To render the inference more tractable, we use the Bethe partition function [2] for approximating the partition function. We use l1-regularization to perform structure estimation and iterative Frank-Wolfe methods to perform parameter estimation, cross-validating model parameters via model likelihood.

4 Preliminary Results

For each each experimental setting, namely, spontaneous activity, drifting gratings in four orientations ($0°$, $45°$, $180°$, $315°$) , we evaluate the models by predicting the values of each neuron in turn, given all the other neurons, on a withheld test dataset ($80 - 100$ test samples per experiment). We note that for spontaneous activity, the simpler tree model outperforms the loopy model. We hypothesize this is due to the random activity in this scenario, which might require more training data to estimate a reliable loopy model. However, for evoked activity, loopy models better capture the underlying process. The tree model has a low standard deviation for spontaneous activity (5), while loopy has a higher (14), but for evoked tree has much higher deviation ($14 - 24$), while loopy has low deviation (3). These results support previous observations that groups of neurons fire together during evoked stimuli, making the functional dependencies more prominent in the data.

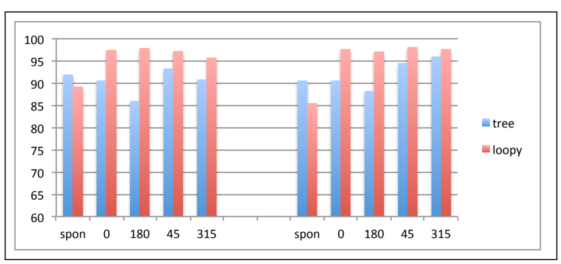

Figure 1: Mean accuracy of predicting values for each of the 135 on the test dataset. Loopy models perform better when the neural activity is driven by a stimulus

5 Future Work

We will model the dynamics of the system by unrolling the Bayesian networks in time to form Dynamic Bayesian networks, which have stationary dynamics (although some non-stationary dynamics will also be tested). By enforcing stationarity, it is possible to learn a compact model, which can potentially encode the temporal aspects of the neuronal data as well as temporal aspects of the visual stimuli. the Bethe approximation recovers marginals over latent variables as well as posterior distributions over parameters. It also allows straightforward introduction of latent variables and unsupervised learning without sampling, as well as other useful extensions.

References

[1] J. Shin. Complexity of bethe approximation. In *In Artificial Intelligence and Statistics*, 2012.

[2] K. Tang, N. Ruozzi, D. Belanger, and T. Jebara. Bethe learning of conditional random fields via map decoding. Technical Report 1503.01228, arXiv, 2015.

[3] J. T. Vogelstein, A. M. Packer, T. A. Machado, T. Sippy, B. Babadi, R. Yuste, and L. Paninski. Fast nonnegative deconvolution for spike train inference from population calcium imaging. *Journal of neurophysiology*, 104(6):3691–3704, Dec. 2010.

[4] J. T. Vogelstein, B. O. Watson, A. M. Packer, R. Yuste, B. Jedynak, and L. Paninski. Spike Inference from Calcium Imaging Using Sequential Monte Carlo Methods. *Biophysical Journal*, 97(2):636–655, July 2009.

[5] Y. Weiss, C. Yanover, and T. Meltzer. MAP estimation, linear programming and belief propagation with convex free energies. In *UAI 2007*, 2007.

[6] A. Weller and T. Jebara. Approximating the Bethe partition function. In *Uncertainty in Artificial Intelligence (UAI)*, 2014.

[7] A. Yuille and A. Rangarajan. The concave-convex procedure (CCCP). *Neural Computation*, 15(4):916–936, 2003.