Maximum Entropy Modeling Of Multi-Neuron Firing Patterns in V1

Ifije E. Ohiorhenuan, Jonathan D. Victor.

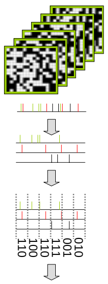
Department of Neurology and Neuroscience, Weill Medical College, Cornell University, New York, NY

Introduction

- Pairs of neurons in the visual cortex are correlated over a few tens of milliseconds.
- The implications of these correlations on the higher-order structure of cortical networks is not known.
- Maximum entropy techniques have demonstrated that the structure of multi-neuron firing patterns in the retina is dominated by interactions between pairs of neurons (Schneidman et al. 2006 and Shlens et al. 2006).
- Here, we implement a similar MaxEnt analysis of multi-neuron firing patterns from tetrode recordings in the primary visual cortex of the anesthetized macaque.

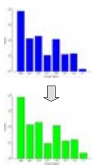
Methods

Data Collection

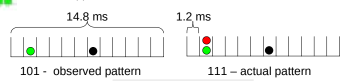


- 6 macaques under sufentanil/propofol anesthesia and vecuronium paralysis
- Tetrode recordings of V1 - 21 different sites with 3-5 simultaneously recorded single units. 16 triplets, 3 quadruplets, 2 quintuplets, 70 neurons in total
- Binary m-sequence modulated checkerboard visual stimulus. 67.6 Hz frame rate (14.8 ms frame length). 8-16 repeats of 60.6 s stimulus and its contrast inverse
- Spikes are sorted and spike trains binned into 14.8 ms bins.
- Multi-neuron firing patterns were counted and analyzed as below

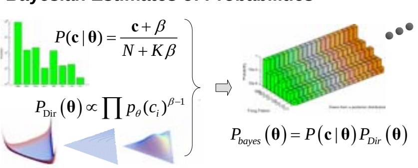
Lock-out Correction



- Spike sorting prevents detection of multiple spike that are within a 1.2 ms lock-out window
- This leads to a systematic underestimate of multi-neuron spiking events within a 14.8 ms bin.
- Correction assumes that within a bin all possible assortments of k spikes into n slots are equally likely but only $\binom{n}{k}$ are observable.



Bayesian Estimates of Probabilities



$$P(\mathbf{c}\,|\,\boldsymbol{\theta}) = \frac{\mathbf{c}+\beta}{N+K\beta}$$

$$P_{\mathrm{Dir}}(\boldsymbol{\theta}) \propto \prod p_{\theta}(c_i)^{\beta-1}$$

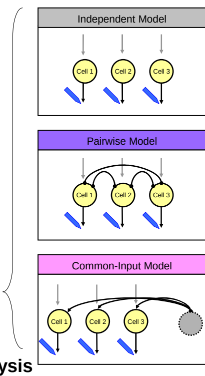
$$P_{bayes}(\boldsymbol{\theta}) = P(\mathbf{c}\,|\,\boldsymbol{\theta})\,P_{Dir}(\boldsymbol{\theta})$$

Since a range of probability distributions could have generated the observed firing counts, we use Markov Chain Monte Carlo simulations to sample from the *a posteriori* distribution under different priors.

Maximum Entropy Models

$$H(P) = -\sum_{i=1}^{K} p(\theta_i)\log_2 p(\theta_i) \quad \text{s.t} \ \langle f(x)\rangle = \sum_{i=1}^{n} p_i f(x_i)$$

- The MaxEnt procedure generates canonical joint firing distributions that are consistent with a certain set of constraints $f(x)$ but are otherwise as unstructured as possible – i.e. maximize $H(P)$
- We construct 3 models:
  1) Independent Model
  2) Pairwise-Constrained Model
  3) Common-Input Model with constraints on the rate at which each neuron fires alone and at which it fires in participation with any ensemble



Information Theoretic Analysis

- To evaluate the models we calculate the K-L distance between the observed firing distributions and the model predictions
- Another measure of performance is the log-likelihood ratios between models which can be computed from the K-L distance
- To assess the fraction of all correlations in the observed data captured by the models, we calculate the connected information

Kullback-Leibler Distance

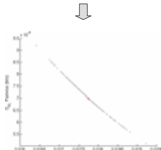
$$D_{\mathrm{KL}}(P\,\|\,Q) = \sum p_i \log_2\left(\frac{p_i}{q_i}\right)$$

Likelihood Ratio

$$\mathrm{LR} = 2^{-n\Delta\tau^{-1}\left(D_{\mathrm{KL}}(P\|Q_{\mathrm{Model\,2}}) - D_{\mathrm{KL}}(P\|Q_{\mathrm{Model\,1}})\right)}$$

Connected Information

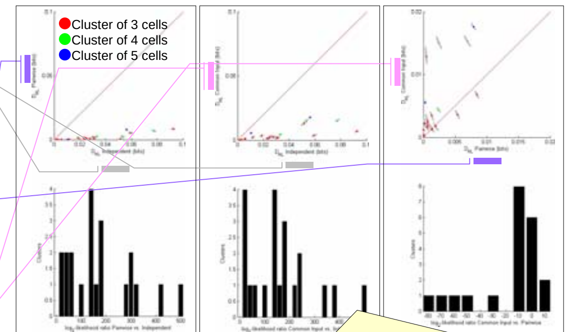
$$I_k = H_1 - H_k$$



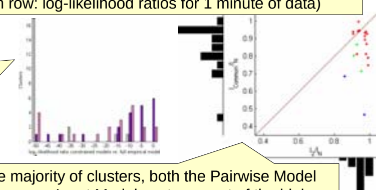
References
Schneidman et al., Nature 2006
Schneidman et al., Phys. Rev. Lett. 2003
Shlens et al., J.Neuroscience 2006
Kennel et al., Neural Comp. 2005

Results

Evaluation of MaxEnt Models



For most clusters, the observed firing distribution is significantly more likely to have been drawn from the Pairwise Model than the Independent & Common-Input Models . (Top row: K-L scatterplots, grey clouds are 3 s.d. bayesian confidence intervals. Bottom row: log-likelihood ratios for 1 minute of data)

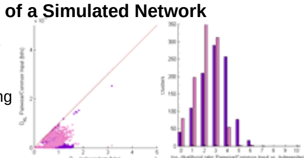


Log likelihood ratios for MaxEnt models vs. the Complete Empirical Model where all possible interactions are accounted for. After 1 minute of data, the Pairwise Model is effectively indistinguishable from the full model for 10 of 21 sites.

For the majority of clusters, both the Pairwise Model and Common-Input Model capture most of the higher order correlations in the observed data. However, the Pairwise Model (mean=0.95) captures a larger fraction of the correlations than the Common-Input Model (mean=0.85)

Control: Analysis of a Simulated Network

Performing the above analysis on a simulated independent network shows that while the Pairwise and Common-Input Models better fit firing distributions from an independent network, this difference is not significant.

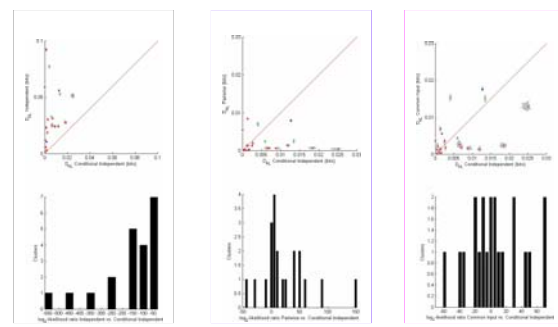


Conclusions

- Correlations between pairs of neurons lead to significant departures from independence for local populations of cortical neurons.
- While driving by a common input also captures much of the observed correlations, for most sites, interactions between pairs of neurons is a more complete model.
- Nevertheless, 50% of the neuronal clusters exhibited significant higher-than-second-order interactions, indicating that cortical circuits can manifest complex patterns of population activity

Influence of Stimulus-Dependent Correlations

- We compared the 3 models to an empirical Conditional Independent Model generated by trial-shuffling spike trains.
- This model effectively has 8190 parameters – one for each frame of the stimulus. Distributions from 13 of 21 clusters from the population have a likelihood of 2^{10}, or higher, of having been produced by the Pairwise Model than by the Conditional Independent Model.



Support: GM7739 and NEI1TEO7138 (IEO) 2RO1EY9314 (JDV)