

Implementing efficient "shotgun" inference of neural connectivity from highly sub-sampled activity data

Daniel Soudry, Suraj Keshri, Patrick Stinson, Min-hwan Oh, Garud Iyengar, Liam Paninski

October 26, 2015

Inferring connectivity in neuronal networks remains a key challenge in statistical neuroscience. The “common input” problem presents a major roadblock: it is difficult to reliably distinguish causal connections between pairs of observed neurons versus correlations induced by common input from unobserved neurons. Available techniques allow us to simultaneously record, with sufficient temporal resolution, only a small fraction of the network. Consequently, naive connectivity estimators that neglect these common input effects are highly biased. In a recent work [1] we propose a “shotgun” experimental design, in which we observe multiple sub-networks briefly, in a serial manner. Thus, while the full network cannot be observed simultaneously at any given time, we may be able to observe much larger subsets of the network over the course of the entire experiment, thus ameliorating the common input problem.

To perform network inference given this type of data, in which only a small fraction of the network is observed in each time bin, we developed a scalable Bayesian method. The method is derived using a cavity-style approximation to the gradient of the expected loglikelihood - in a generalized linear model for a spiking recurrent neural network. Simulation demonstrated that the shotgun experimental design can eliminate the biases induced by common input effects. Networks with thousands of neurons, in which only a small fraction of the neurons is observed in each time bin, can be quickly and accurately estimated, achieving orders of magnitude speed up over previous approaches.

[1] Soudry, D., Keshri, S., Stinson, P., Oh, M., Iyengar, G., & Paninski, L. (2015). Efficient “Shotgun” Inference of Neural Connectivity from Highly Sub-sampled Activity Data. *PLOS Computational Biology*, 11(10), e1004464.

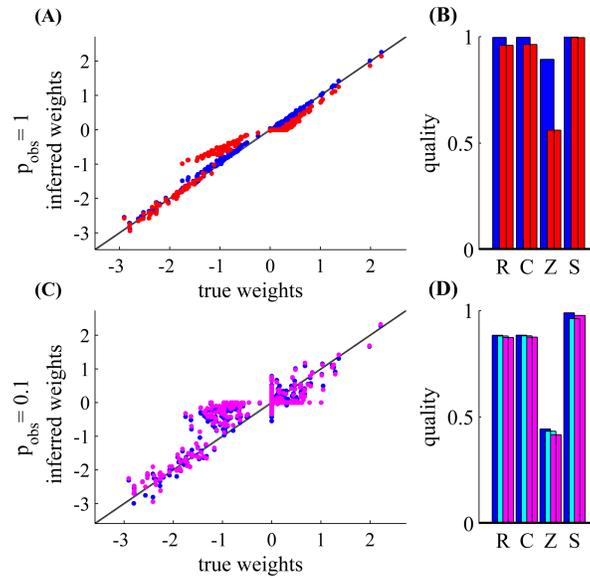


Figure 1: The proposed Expected LogLikelihood (ELL) based estimator is statistically efficient. **(A)**: The ELL-based method (blue) compares favorably to the standard MAP estimate (red) when spikes are fully observed (both with a similar L1 prior), even though it is several orders of magnitude faster. **(B)** Quality of estimation measures (higher values correspond to better estimates) - S = sign detection, Z = zero detection, C = correlation, $R = \sqrt{R^2}$. *Bottom (C,D)*: We compare the ELL based method (blue) to a Expectation Maximization (EM) approach (in which we use Gibbs sampling to infer missing spikes), when only 10% of the spikes are observed. We show the results after one (cyan) and two (magenta) EM steps. The EM steps (which are very slow to compute) do not improve over the ELL-based method. Parameters: $N = 50$ neurons, $T = 1.4$ hours.