Inferring the Structure of Populations of Neurons using a Sequential Monte Carlo EM Algorithm

Joshua T. Vogelstein[1], Kechen Zhang[1], Bruno Jedynak[1], and Liam Paninski[2]

[1]Johns Hopkins University, [2]Columbia University

A fundamental goal of neuroscience is to be able to construct models of a population of neurons acting in concert to perform nonlinear operations. A primary difficulty hindering progress towards this goal is the paucity of computational tools designed with population data in mind. Cross-correlation based ideas become computationally intractable due to the combinatorial explosion. Fitting phenomenological models to multicell data has received some attention, but we take a more biophysically realistic approach. Our models have two distinguishing features to this effect. First, we consider experimentally derived noise distributions act on the neurons, which can be non-Gaussian (e.g., synaptic failures). Second, we assume the existence of unobserved data (i.e., hidden states), in addition to the observed data (or states). These qualities naturally lead to a Sequential Monte Carlo Expectation-Maximization approach (SMC-EM). The expectation step requires computing the likelihood of the joint probability of both the hidden and observed state trajectories. By assuming that the hidden states are Markov, these distributions are approximated using SMC. The maximization step requires computing the maximum likelihood estimates (MLEs) of the model parameters. By choosing models with likelihood functions that are jointly convex in the parameters, we can guarantee convergence. The MLEs for those parameters for which no closed-form solutions exist are determined using standard convex optimization tools.

We apply this approach to data assumed to have been acquired using one of two population recording techniques: i) arrays of extracellular electrodes, or ii) 2-photon microscopy (2PM). When using electrodes, the observed states are spike times, and the hidden states are voltages of each neuron and conductances of each synapse. Each neuron is modeled with stochastic-leaky-integrate-and-fire (IF) dynamics, with noisy conductance-based synapses (noise due to failures or number of neurotransmitters released). When using 2PM, the data are noisy images of calcium signals, which are themselves noisy low-pass filtered versions of the spike times. Neurons in this case are modeled as Linear-Nonlinear-Poisson (LNP) functions, such that the external stimulus is linearly filtered, its output is exponentiated giving firing rate, and spikes are generated with Poisson statistics according to that rate. Presynaptic spikes operate on postsynaptic firing rate via linear filters.

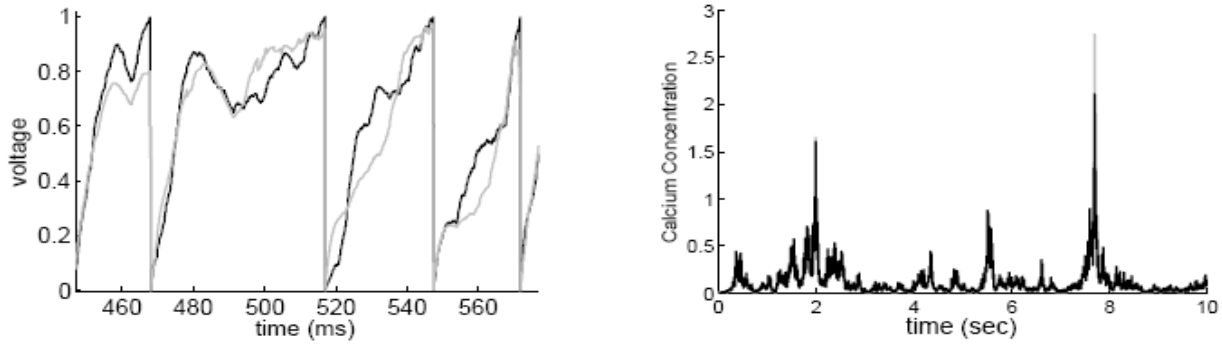

Figure 1: Mean particle hidden state trajectories (gray lines) coincide well with actual hidden state trajectories (black lines). Left: voltage for IF model. Right: calcium concentration for LNP model.

Acknowledgments: This work was supported by NIH grant DC000109.