Model-based optimal inference of spike times and calcium dynamics given noisy and intermittent calcium-fluorescence imaging

Joshua T. Vogelstein[1] **and Liam Paninski**[2]

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

Using calcium sensitive fluorescence to study neural dynamics is becoming increasingly popular both in vitro and in vivo, at the level of individual spines, dendrites, boutons, neurons, or populations of neurons. While the data collected from these experiments are time-varying fluorescence images, the signals of interest are the precise spike times and/or the intracellular calcium concentrations, $[\text{Ca}^{2+}]_t$, of the observable neurons. Unfortunately, determining the true value of these hidden signals is a difficult problem for a number of reasons. First, observations are noisy. This is a problem unlikely to be solved in the near future, as a major source of the noise is photon *shot noise*, which reflects the quantal nature of light. Second, observations may be undersampled. When using epifluorescence, this problem can be solved with faster cameras. Similarly, when using confocal microscopy, faster cameras and spinning disk technology enable sufficiently fast imaging. However, when using two-photon microscopy (2PM), images are reconstructed from serial scans across the imaging plane. Array based technologies and faster scanning can help alleviate this problem, but may also exacerbate the third problem: low signal-to-noise ratio. Fourth, the relationship between fluorescence observations and $[\text{Ca}^{2+}]_t$ is nonlinear, especially for fluorescent proteins. This has placed undesirable and unnecessary restrictions on the calcium indicators used for analysis, as the standard analytical tools assume a linear relationship between $[\text{Ca}^{2+}]_t$ and fluorescence.

We develop a technique based on a first-principles approach incorporating a well defined probabilistic "forward-model" of the signals of interest and the imaging process. In particular, we assume a biophysically based nonlinear model governing spiking activity, $[\text{Ca}^{2+}]_t$, and fluorescence observations. We then develop a Sequential-Monte-Carlo Expectation-Maximization (SMC-EM) algorithm, designed to optimally infer the $[\text{Ca}^{2+}]_t$ and spike times, given the observed fluorescence signals. This strategy allows us to infer $[\text{Ca}^{2+}]_t$ using only a single trial. By considering empirically derived noise distributions, the algorithm provides errorbars on the inference of the $[\text{Ca}^{2+}]_t$ and spike trains, thus acknowledging the inherent uncertainty of these estimates, given the nature of the data. Further, this framework provides a natural way of incorporating a stimulus (if present) to improve the accuracy of the inference. Finally, because our approach fits the model parameters directly to each observed cell's fluorescence data, we do not need to train the algorithm using simultaneously acquired intracellular electrophysiology and $[\text{Ca}^{2+}]_t$ data for each cell, but rather the parameters are automatically recovered.

Simulated results indicate that this approach can successfully infer spike times from a range of noisy, intermittent data, even as the fluorescence signal saturates, or with multiple spikes per image frame. Experimental verification of this approach using different indicators is in progress.

Acknowledgments
The authors would like to thank Q. Huys, M. Nikitchenko, K. Svoboda, B. Watson, R. Yuste, D. Greenberg, D. Yue, M. Tadross, E. Young, and S. Mihalas for helpful discussions. Support for JV was provided by NIDCD DC00109. LP is supported by an NSF CAREER award and by an Alfred P. Sloan Research Fellowship.