From calcium sensitive fluorescence movies to spike trains

Joshua T Vogelstein[1], **Baktash Babadi**[2], **Bredon O Watson**[2], **Rafael Yuste**[2,3], **Liam Paninski**[2,4]

[1]Department of Neuroscience, Johns Hopkins School of Medicine, [2]Department of Neuroscience, Columbia University, [3] Howard Hughes Medical Institute, [4]Department of Statistics, Columbia University

joshuav@jhu.edu

Abstract

Motivation *Calcium imaging is quickly becoming a prominent paradigm to collect data in neuroscience. To maximally utilize the power of this technique, complementary analytical tools can be built.*

Goal *We aim to develop analytical tools to facilitate inferring spike trains from fluorescent observations, fitting tuning curves, and inferring population connectivity, given only short sequences of possibly very noisy, low temporal resolution, and saturating fluorescence images.*

Solution *By framing the problem as a state-space problem, we can utilize tools developed by the statistics community for related problems. In particular, we develop a (i) fast filter utilizing a tridiagonal trick and interior point methods to approximate the MAP spike train, (ii) particle filter to infer the probability of spiking in each frame, (iii) a population version of our particle filter to infer connectivity.*

Conclusions

Our fast filter can accurately approximate the most likely spike train given the fluorescence data in $O(T)$ time. When fluorescence saturates, stimuli are present, or temporal resolution is unsatisfactorily slow, our particle filter can infer the probability of a spike in each time bin. If a small population of neurons are imaged simultaneously, our population particle filter can learn the effective connectivity of the observable neurons.

References

1) Vogelstein JT, Watson BO, Packer AM, Yuste R, Jedynak B, and Paninski L. *Spike inference from calcium imaging using sequential Monte Carlo methods.* In Press at Biophysical Journal.

2) Vogelstein JT, Babadi B, Packer AM, Yuste R, and Paninski L. *Fast methods for spike inference from calcium imaging.* In preparation.

Acknowldgments

Support for JTV was provided by NIDCD DC00109. LP is supported by an NSF CAREER award, by an Alfred P. Sloan Research Fellowship, and a McKnight Scholar Award.

1. simple

1.1 a simple model

$$F_t = \alpha C_t + \beta + \sigma\varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0,1)$$
$$C_t = \gamma C_{t-1} + \eta + \rho n_t, \qquad n_t \sim \text{Poisson}(n_t; \lambda\Delta)$$

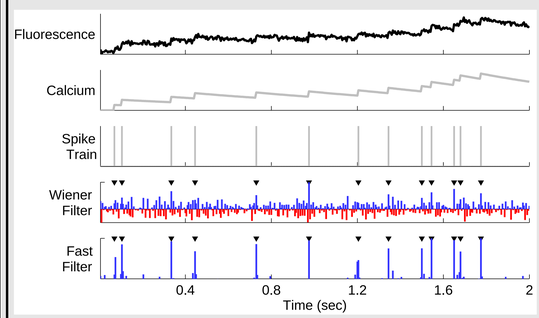
1.2 a simple filter

We would like to solve the following optimization problem:

$$\widehat{\boldsymbol{n}}_{MAP} = \underset{n_t \geq 0 \forall t}{\operatorname{argmax}}\, P(\boldsymbol{n}|\boldsymbol{F}) = \underset{n_t \geq 0 \forall t}{\operatorname{argmax}}\, P(\boldsymbol{F}|\boldsymbol{n})P(\boldsymbol{n})$$

where $\boldsymbol{n}$ is the spike train and $\boldsymbol{F}$ is the fluorescence observations. To impose the non-negative constraint, we use an interior point method, making this a concave constrained optimization approach. To solve it effeciently, we make use of the tridiagonal structure of the Hessian, and use Gaussian elimination to take Newton-Raphson steps in $O(T)$ time. A naïve approach would require $O(T^3)$.

1.3 a simple result



ABOVE: A comparison of our fast filter with the Wiener filter. The Wiener filter is the optimal linear filter (under the assumption that spikes are distribution according to a Gaussian process). By imposing a non negativity constraint, we can drastically reduce the effective noise level upon filtering the fluorescence trace.

RIGHT: A comparison of our particle filter with the Wiener filter. When the fluorescence signal approaches saturation, the Wiener filter struggles to differentiate between image frames with and without a spike (third panel). Our particle filter, however, still correctly infers the precise spike times (fourth panel). Artificially decreasing the temporal resolution decreases the inference precision (fifth panel). However, incorporating the time-varying stimulus into our model, we can recover much of the lost precision (bottom panel).

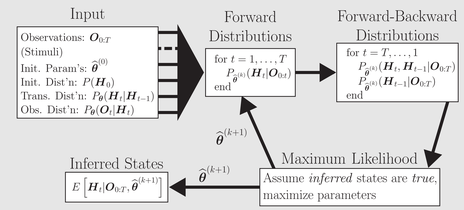
2. nonlinear

2.1 a nonlinear model

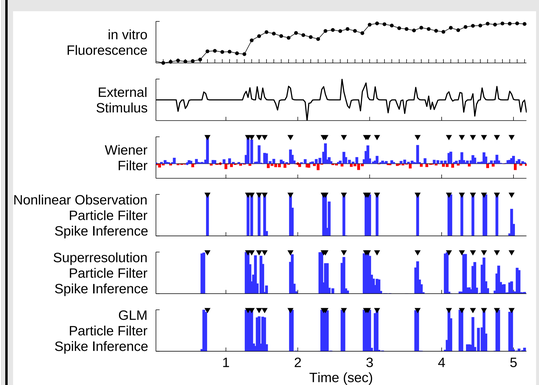
$$F_t = \alpha S([\text{Ca}^{2+}]_t) + \beta + (S([\text{Ca}^{2+}]_t) + \sigma_{Ft})\varepsilon_{F,t} \quad t \in \mathcal{T}_o$$
$$[\text{Ca}^{2+}]_t = \gamma[\text{Ca}^{2+}]_{t-1} + \eta + \rho n_t + \sigma_{ct}\sqrt{\Delta}\varepsilon_{c,t}$$
$$n_t \sim \text{Bin}\Big(n_t; f(b + \boldsymbol{k}'\boldsymbol{x}_t + \boldsymbol{\omega}'\boldsymbol{h}_t)\Big)$$
$$h_{l,t} = \gamma_l h_{l,t-1} + n_{t-1} + \sigma_h\sqrt{\Delta}\varepsilon_{l,t}$$

2.2 a nonlinear filter

Instead of finding $\boldsymbol{n}_{MAP}$, we can find the probability that there was a spike in any time bin, conditioned on the entire fluorescence time series: $P_{\boldsymbol{\theta}}(n_t|\boldsymbol{F})$. We then use a *forward-backward* approach to compute the desired distributions, approximating the forward recursion by using a sequential Monte Carlo algorithm (aka, a particle filter).



2.3 a nonlinear result



3. population

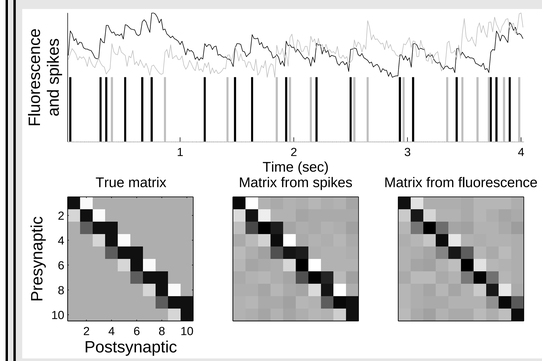
3.1 a population model

$$F_{i,t} = \alpha_i S([\text{Ca}^{2+}]_{i,t}) + \beta_i + (S([\text{Ca}^{2+}]_{i,t}) + \sigma_{F_i,t})\varepsilon_{F_i,t}$$
$$[\text{Ca}^{2+}]_{i,t} = \gamma_i[\text{Ca}^{2+}]_{i,t-1} + \eta_i + \rho_i n_{i,t} + \sigma_{c_i,t}\sqrt{\Delta}\varepsilon_{c_i,t}$$
$$n_{i,t} \sim \text{Bin}\Big(n_{i,t}; f(b_i + \boldsymbol{k}_i'\boldsymbol{x}_t + \sum_{j=1}^{N}\boldsymbol{\omega}_{ij}'\boldsymbol{h}_{j,t})\Big)$$
$$h_{i_l,t} = \gamma_{i_l} h_{i_l,t-1} + n_{i,t-1} + \sigma_{h_{i_l}}\sqrt{\Delta}\varepsilon_{h_{i_l},t}$$

3.2 a population filter

1: **for** $i = 1, \ldots, N$ **do**
2: Let $\widetilde{\boldsymbol{x}}_{i,t} = [\boldsymbol{x}_t, \boldsymbol{h}_t]$ and $\widetilde{\boldsymbol{k}}_i = [\boldsymbol{k}_i, \boldsymbol{\omega}_i]$
3: Use nonlinear filter to infer $P_{\boldsymbol{\theta}}(n_{i,t}|\boldsymbol{n}_{\setminus i}, \boldsymbol{F}_i)$ (i.e., conditioned on all spikes from all other neurons and fluorescence signal from neuron i)
4: Estimate $\widetilde{\boldsymbol{k}}_i$, let $\boldsymbol{\omega} \leftarrow \widetilde{\boldsymbol{k}}_i$(p:end) (where $\boldsymbol{x}_t \in \mathbb{R}^p$)
5: **end for**

3.3 a population result



ABOVE: Inferring network connectivity given noisy simulated calcium fluorescence data. TOP: A short segment of simulated spike trains and fluorescence activity from 2 of the 10 neurons in the population. BOTTOM: True connectivity matrix, connectivity matrix using the true spikes, connectivity matrix using only the fluorescence activity. A total of about 2000 spikes/neuron were used for this inference. White denotes excitatory connections; black denotes inhibitory; the same grayscale is used for each panel. All spikes are spontaneous (ie, no external drive was applied to the network). Note that it is harder to estimate the connectivity map given only noisy fluorescence as compared with the true spike trains. Nonetheless, the correct connectivity is obtained with only a few minutes of imaging data.