Maximum Likelihood Inference of Neuronal Dynamics under Noisy and Intermittent Observations using Sequential Monte Carlo EM Algorithms

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

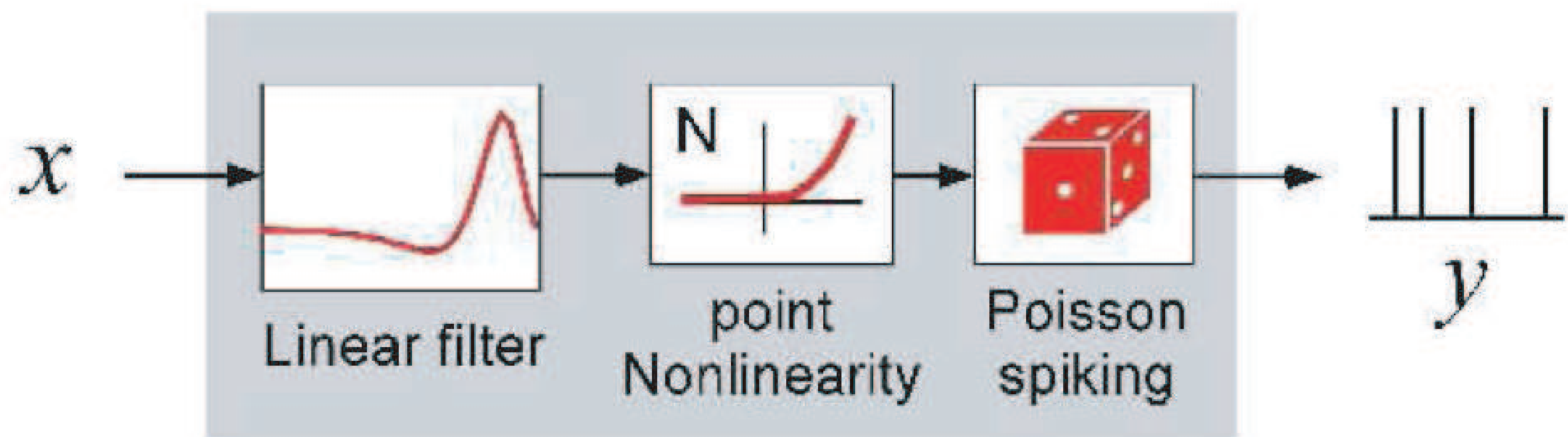
[1]Johns Hopkins University
[2]Columbia University

Motivation

Although technology continues to progress improving the information experimentalists acquire about populations of neurons, the data remains noisy. Because the distribution of noise is, in general, non-Gaussian; and biophysically realistic neural models are nonlinear, sophisticated nonlinear systems identification tools must be developed to infer both (i) the unobserved/hidden states and (ii) the parameters governing the dynamics. By writing models of neurons in terms of observation states (eg, photon or spike counts), and hidden states (eg, voltage or conductance), an Expectation–Maximization (EM) algorithm is a natural framework. By writing the dynamics as Markov Chains, the EM approach is greatly simplified. Because the states are continuous with nonlinear dynamics, exact expectations are not feasible, and we therefore approximate them using Sequential Monte Carlo (SMC) algorithms. This approach is very general and can be applied to a variety of models with increasing complexity.

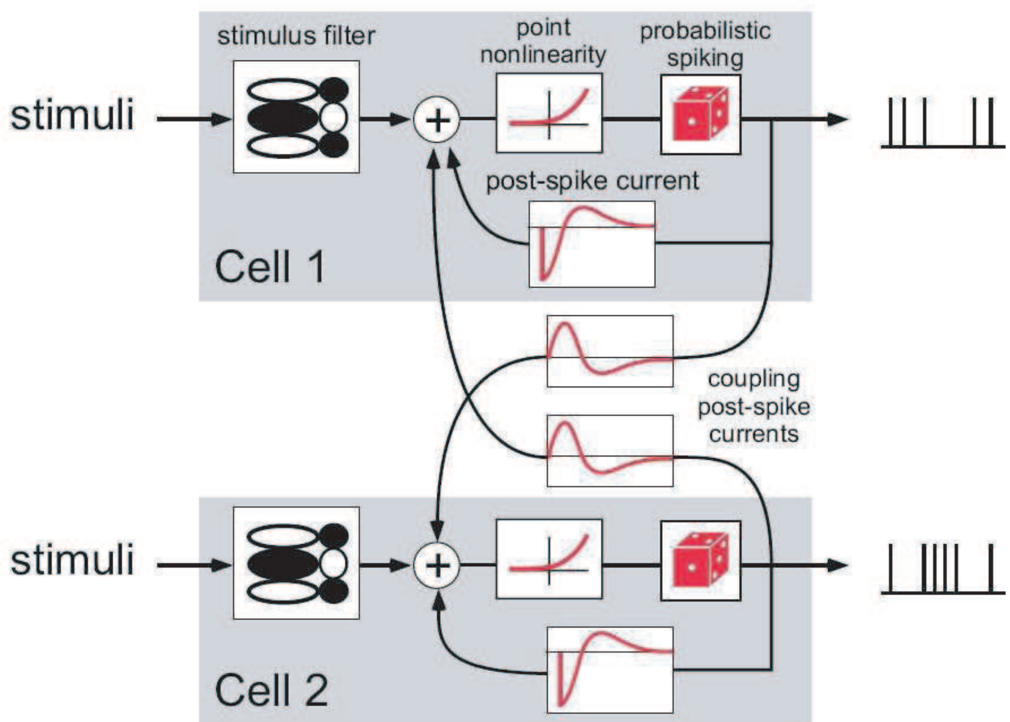
Generalized Linear Models (GLM)

- Cascade Linear–Nonlinear–Poisson (LNP) Models
- Unique global extrema
- Great success in spike train data



Generalizations

- Spike history terms for arbitrary dynamics (e.g., facilitation)
- Spike history terms from other neurons for population effects
- *Noisy and intermittent observations*
- Maintains properties of GLM (e.g., log concave, etc.)
- Populations of neurons are Hidden Markov Models



Hidden Markov Models (HMM)

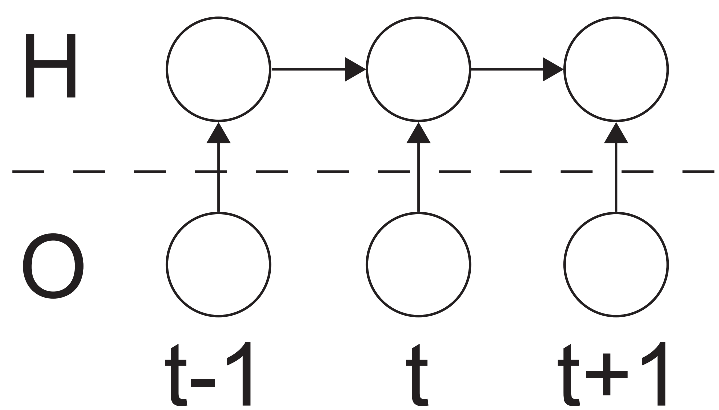
Markov Property

$$P[H_t, H_0, \ldots, H_{t-1}] = P[H_t | H_{t-1}] \tag{1}$$

Hidden Markov Model

$$P[H_{0:T}, O_{0:T}] = P[O_{0:T} | H_{0:T}] P[H_{0:T}] \tag{2}$$

$$= P[H_0] \prod_{t=1}^{T} P[O_t | H_t] P[H_t | H_{t-1}] \tag{3}$$



Expectation–Maximization (EM)

Expectation Step: Write the expected value of the *joint* likelihood, $Q(\theta, \theta')$

Maximization Step: Maximize $Q(\theta, \theta')$

$$\widehat{\theta} = \max_{\theta} Q(\theta, \theta') \tag{4}$$

$$= \max_{\theta} E_{P(H_{0:T}|O_{0:T})}[\log P_{\theta}(O_{0:T}, H_{0:T})] \tag{5}$$

$$= \max_{\theta} \sum_{H_0} \cdots \sum_{H_T} \log P_{\theta}(O_{0:T}, H_{0:T}) P_{\theta'}(H_{0:T}|O_{0:T}) \tag{6}$$

EM for HMM

- Joint likelihood greatly simplifies
- Monotonically increasing
- Recursive algorithms available

$$
\begin{aligned}
\widehat{\boldsymbol{\theta}} = \max_{\boldsymbol{\theta}} \sum_{H_0} P_{\boldsymbol{\theta}'}\big[H_0 \mid O_{0:T}\big] \times \ln P_{\boldsymbol{\theta}}\big[H_0\big] \\
+ \sum_{t=1}^{T} \sum_{H_t} \sum_{H_{t-1}} P_{\boldsymbol{\theta}'}\big[H_t, H_{t-1} \mid O_{0:T}\big] \times \ln P_{\boldsymbol{\theta}}\big[H_t \mid H_{t-1}\big] \\
+ \sum_{t=0}^{T} \sum_{H_t} P_{\boldsymbol{\theta}'}\big[H_t \mid O_{0:T}\big] \times \ln P_{\boldsymbol{\theta}}\big[O_t \mid H_t\big]
\end{aligned} \tag{7}
$$

Sequential Monte Carlo (SMC)

- Efficiently (sometimes optimally) sample from HMM
- Many samples approximate the distribution
- SMC–EM Algorithm

$$\widehat{\boldsymbol{\theta}} = \max_{\boldsymbol{\theta}} \sum_{i_0} P_{\boldsymbol{\theta}'}\left[H^{(i_0)} \mid O^{(0:T)}\right] \times \ln P_{\boldsymbol{\theta}}\left[H^{(i_0)}\right]$$

$$+ \sum_{t=1}^{T} \sum_{i_t} \sum_{j_{t-1}} P_{\boldsymbol{\theta}'}\left[H^{(i_t)}, H^{(j_{t-1})} \mid O^{(0:T)}\right] \times \ln P_{\boldsymbol{\theta}}\left[H^{(i_t)} \mid H^{(j_{t-1})}\right]$$

$$+ \sum_{t=0}^{T} \sum_{i_t} P_{\boldsymbol{\theta}'}\left[H^{(i_t)} \mid O^{(0:T)}\right] \times \ln P_{\boldsymbol{\theta}}\left[O^{(t)} \mid H^{(i_t)}\right] \quad (8)$$

Forward–Backward Approach

Forward Sampling

$$H_t \sim P[H_t | H_{t-1}, O_{0:t}] \tag{9}$$

Backwards Recursion

$$P[H_t, H_{t-1} | O_{0:T}] = \\ P[H_{t-1} | O_{0:T}] \frac{P[H_t | H_{t-1}] P[H_{t-1} | O_{0:t-1}]}{\int dH_{t-1} P[H_t | H_{t-1}] P[H_{t-1} | O_{0:t-1}]} \tag{10}$$

Linear–Nonlinear–Poisson Calcium Model

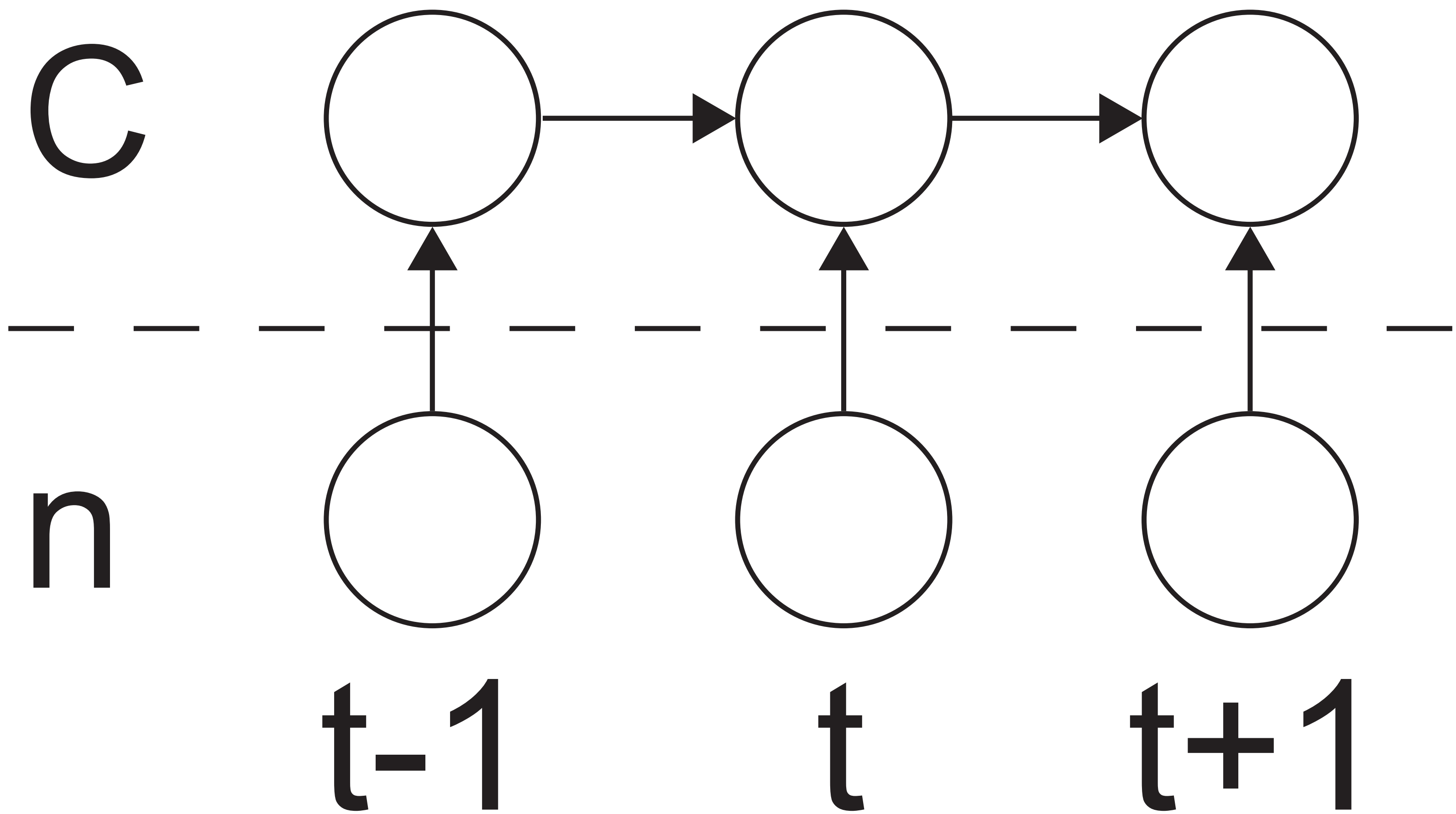
LNP–Ca: Model Specification

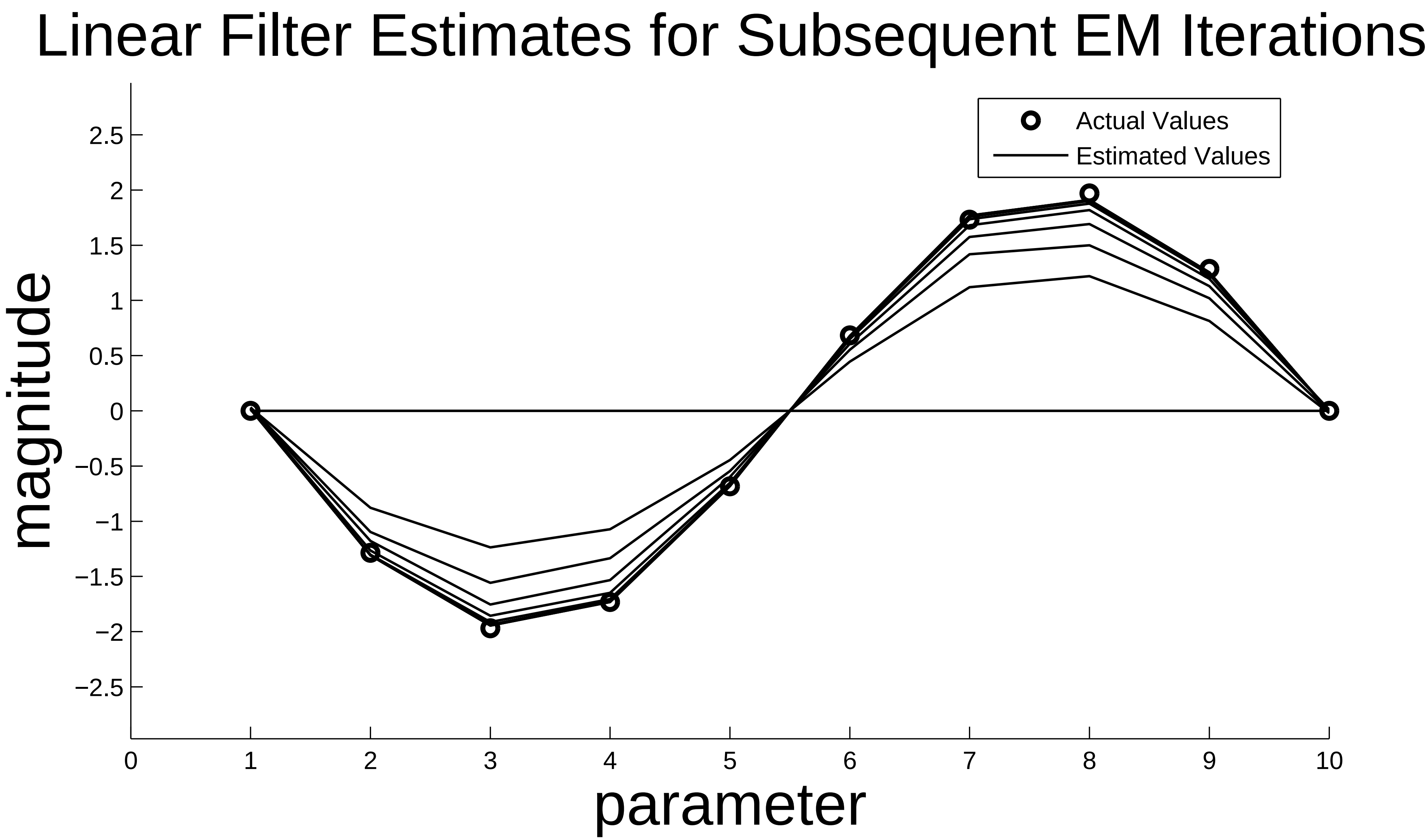
$$\lambda(t) = f\big(b + \boldsymbol{k}'\boldsymbol{x}(t)\big) \tag{11}$$

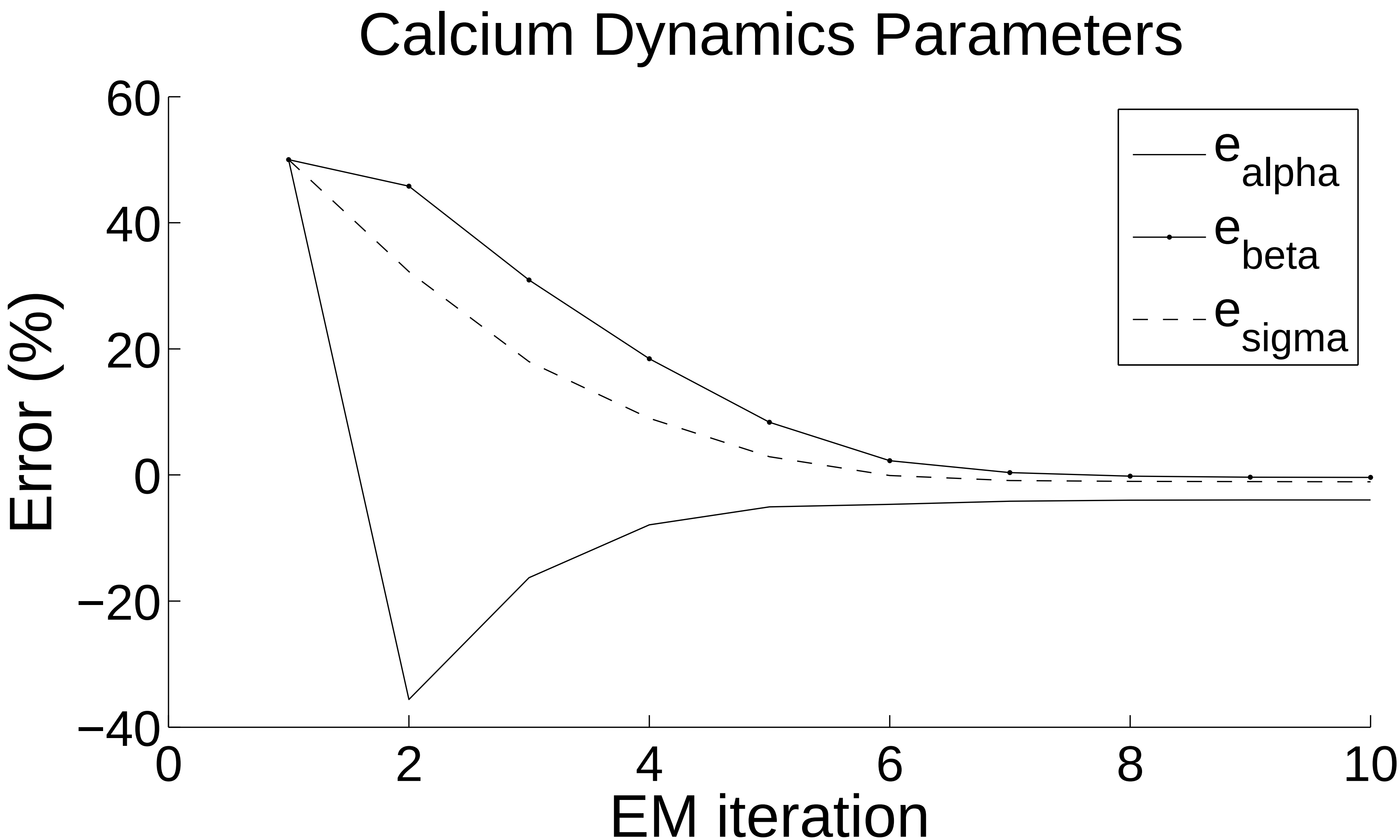
$$n(t) \sim \mathcal{P}\big[n(t); \lambda(t)dt\big] \tag{12}$$

$$dC(t) = -\alpha C(t)dt + \sigma dB_t + \beta n(t) \tag{13}$$

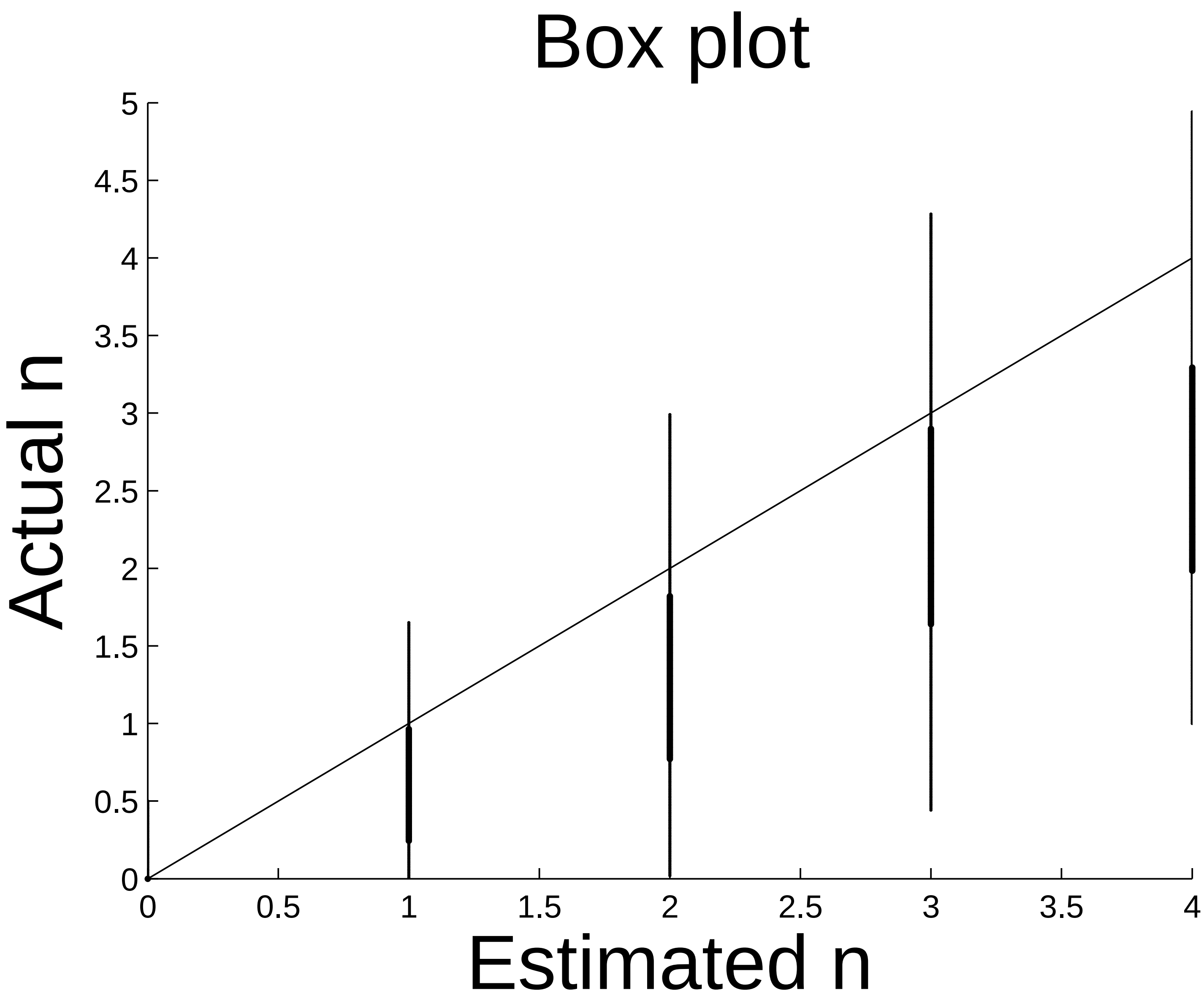
LNP–Ca: Graphical Model




Linear Filter Estimates for Subsequent EM Iterations
Actual Values
Estimated Values
magnitude
parameter
2.5
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
−2.5
0
1
2
3
4
5
6
7
8
9
10



Calcium Dynamics Parameters
Error (%)
EM iteration
60
40
20
0
−20
−40
0
2
4
6
8
10
e_{alpha}
e_{beta}
e_{sigma}



Box plot
Actual n
Estimated n
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

Linear–Nonlinear–Poisson Intermittent Calcium Observations Model

LNP–ICa: Model Specification

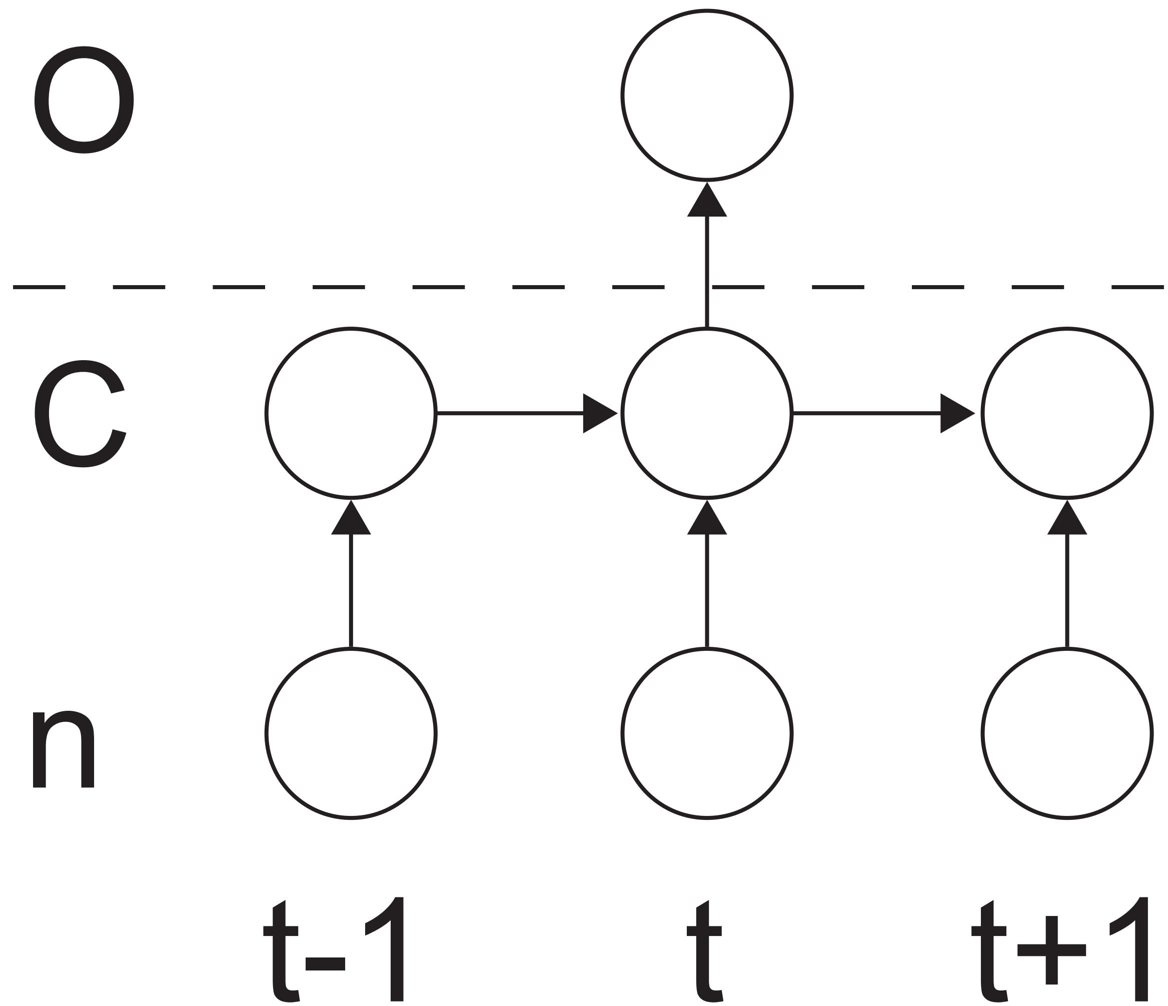
$$\lambda(t) = f\big(b + \boldsymbol{k}'\boldsymbol{x}(t)\big) \tag{14}$$

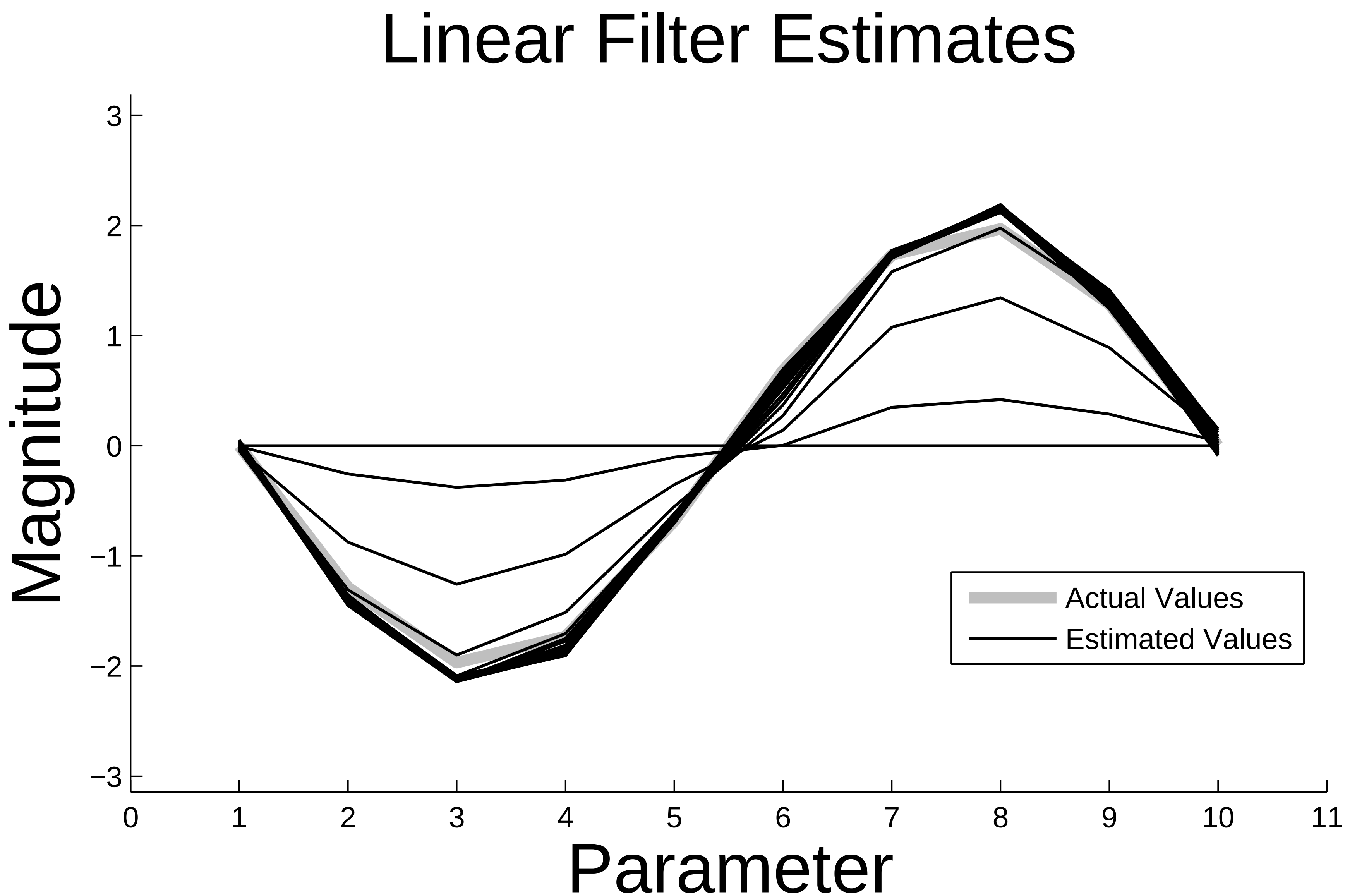
$$n(t) \sim \mathcal{P}\big[n(t); \lambda(t)dt\big] \tag{15}$$

$$dC(t) = -\alpha C(t)dt + \sigma dB_t + \beta n(t) \tag{16}$$

$$O(t) = \eta C(t) \qquad \text{whenever } t/r \in \mathcal{N} \tag{17}$$

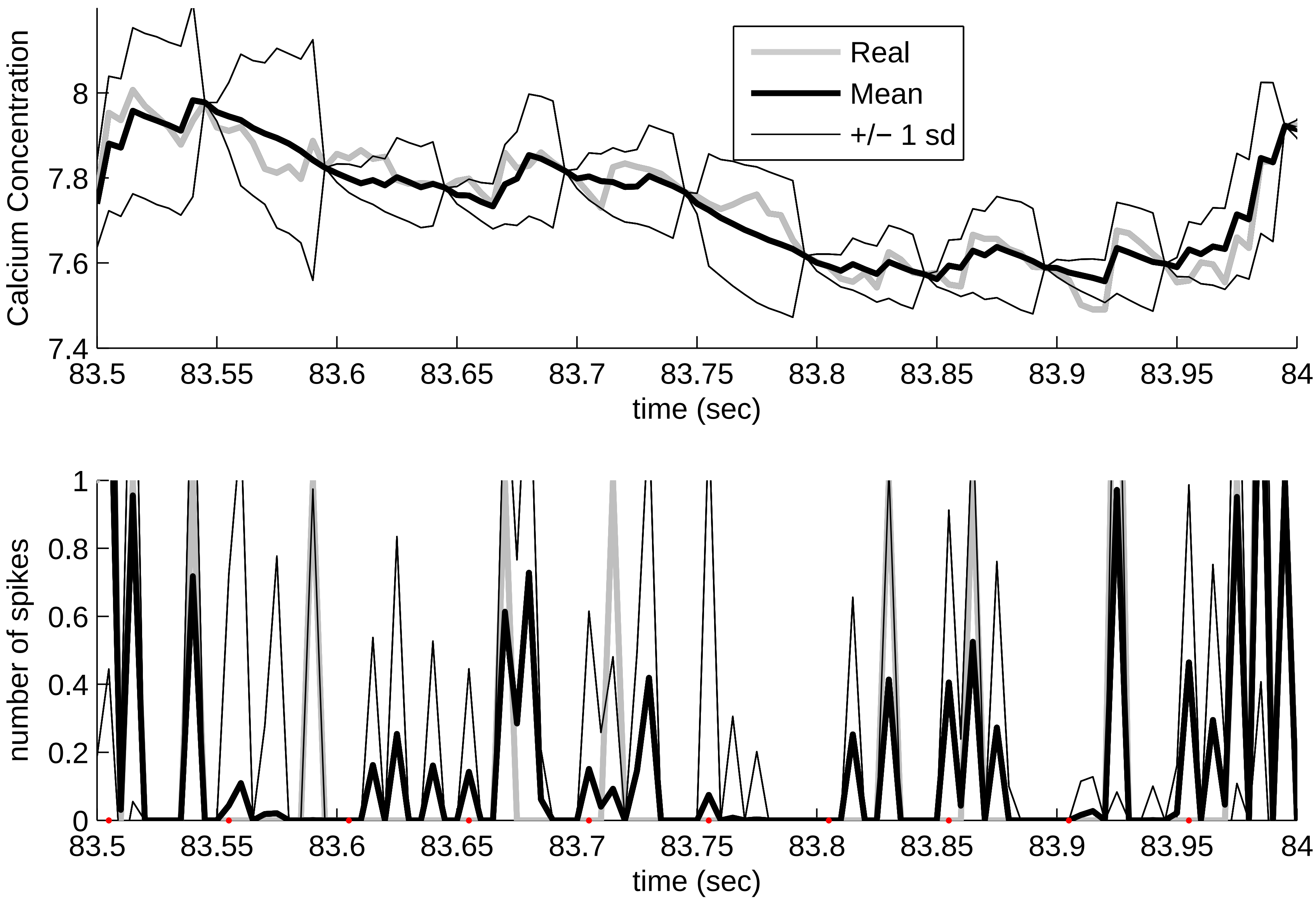
LNP–ICa: Graphical Model





Linear Filter Estimates
Magnitude
Parameter
3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9
10
11
Actual Values
Estimated Values

LNP–ICa: Density Estimation



Linear–Nonlinear–Poisson Noisy Intermittent Calcium Observations Model

LNP–NICa: Model Specification

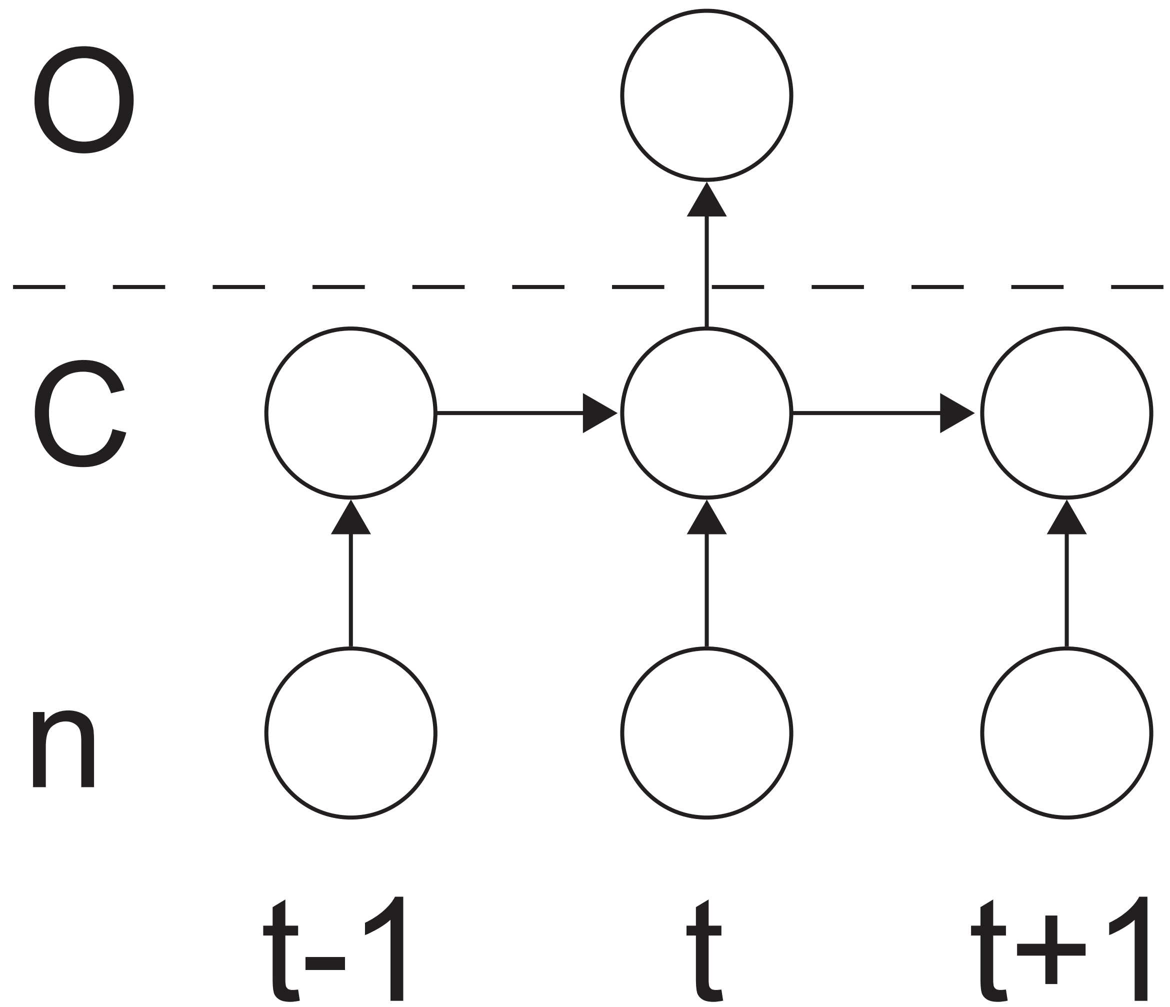
$$\lambda(t) = f\big(b + \boldsymbol{k}'\boldsymbol{x}(t)\big) \tag{18}$$

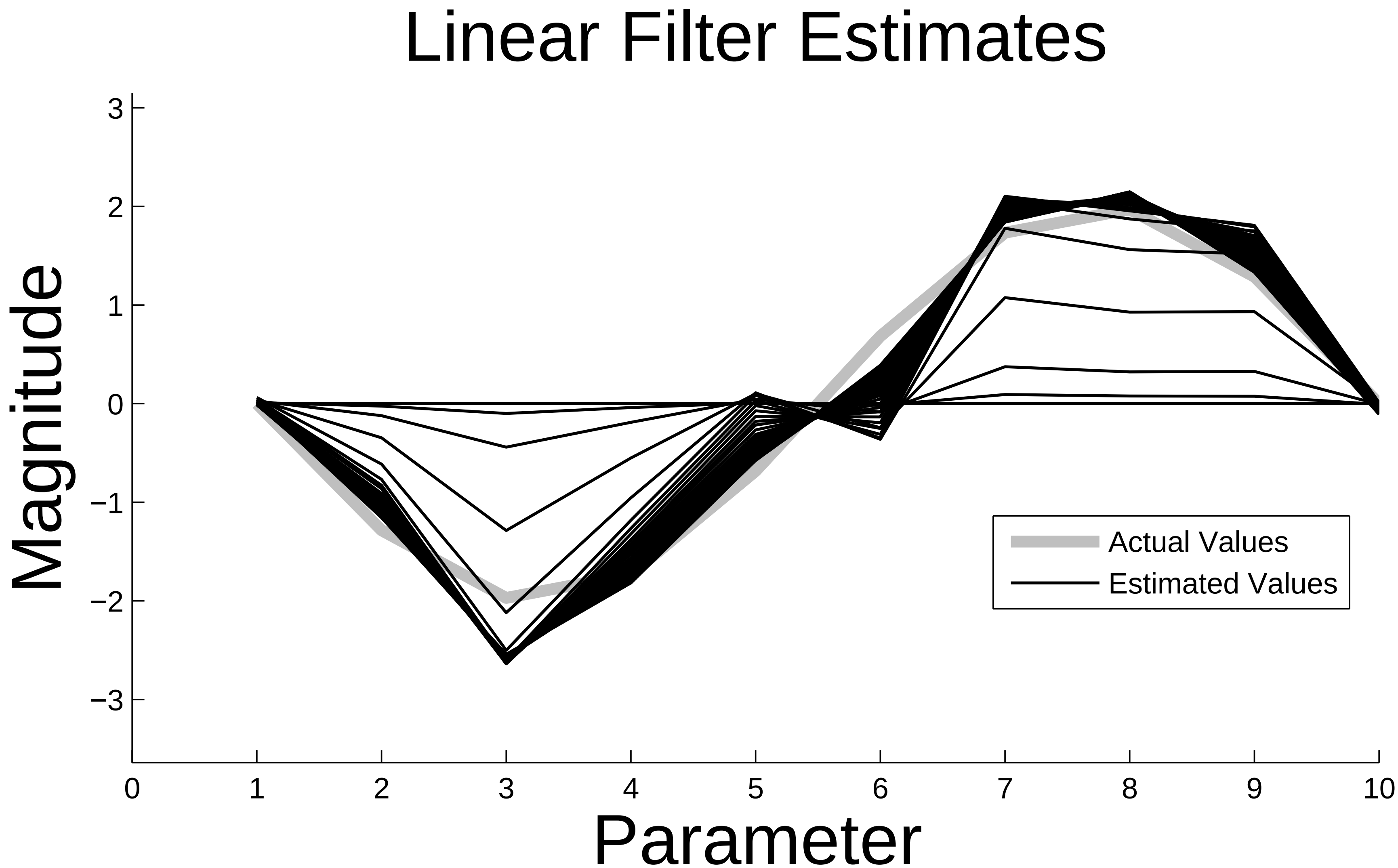
$$n(t) \sim \mathcal{P}\big[n(t); \lambda(t)dt\big] \tag{19}$$

$$dC(t) = -\alpha C(t)dt + \sigma dB_t + \beta n(t) \tag{20}$$

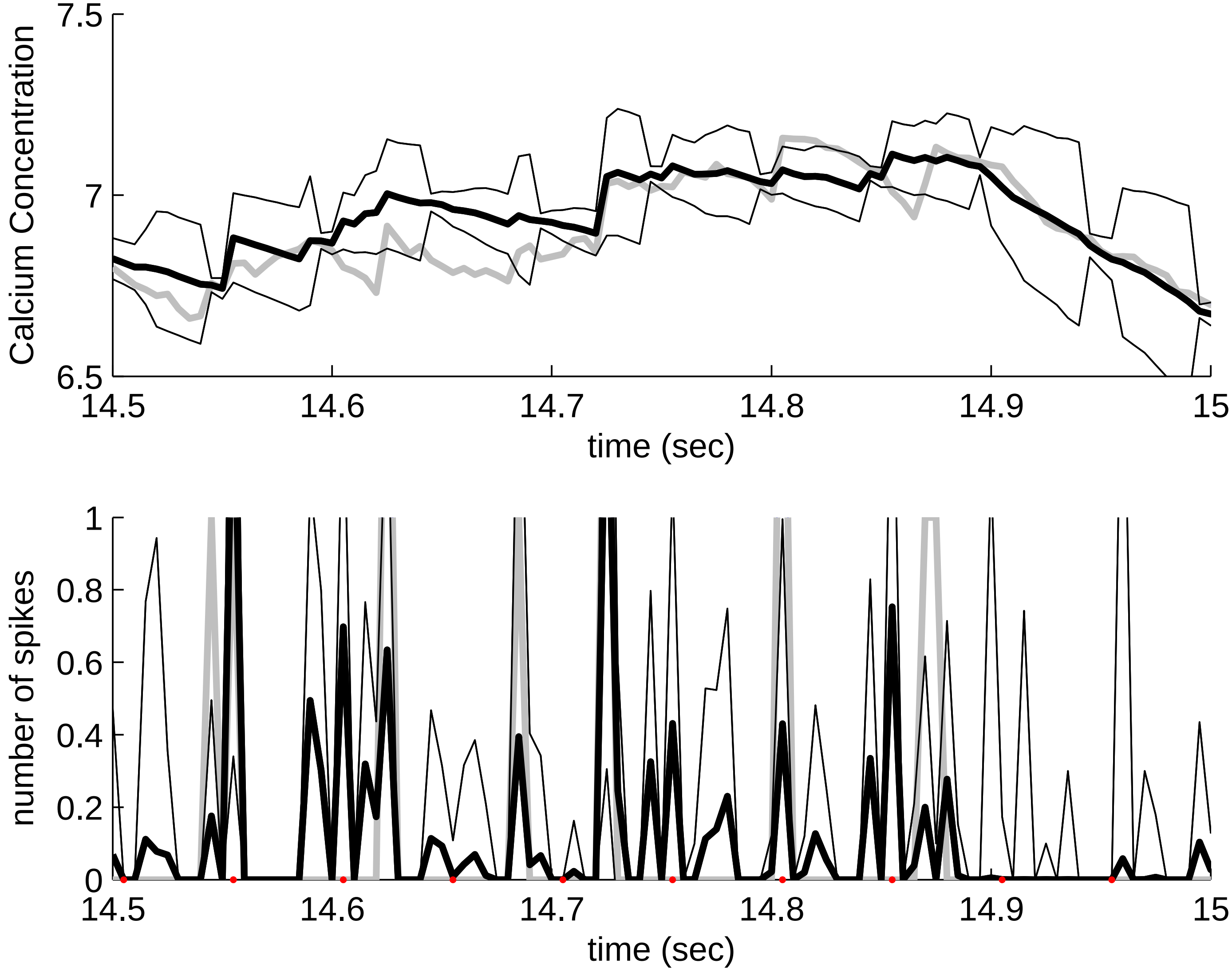
$$O(t) = \eta C(t) + \varepsilon B_t \qquad \text{whenever } t/r \in \mathcal{N} \tag{21}$$

LNP–NICa: Graphical Model




Linear Filter Estimates
Magnitude
Parameter
3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9
10
Actual Values
Estimated Values

LNP–NICa: Density Estimation



Linear–Nonlinear–Poisson with Spike–History and Noisy–Intermittent–Calcium Model

LNP–SH–NICa: Model Specification

$$\lambda(t) = f\left(b + \boldsymbol{k}'\boldsymbol{x}(t) + \sum_l \omega_l h_l(t)\right) \tag{22}$$

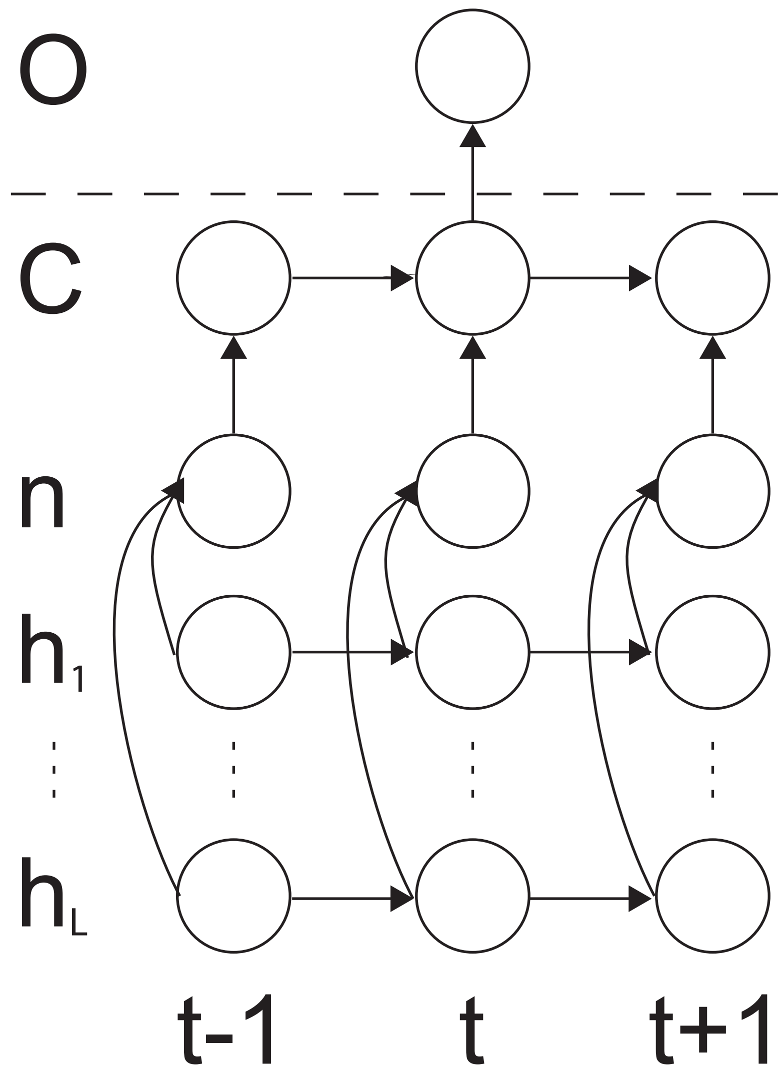
$$dh_l(t) = -\gamma_l h_l(t) + \sigma_h dB_{h,t} + n(t) \tag{23}$$

$$n(t) \sim \mathcal{P}\big[n(t); \lambda(t)dt\big] \tag{24}$$

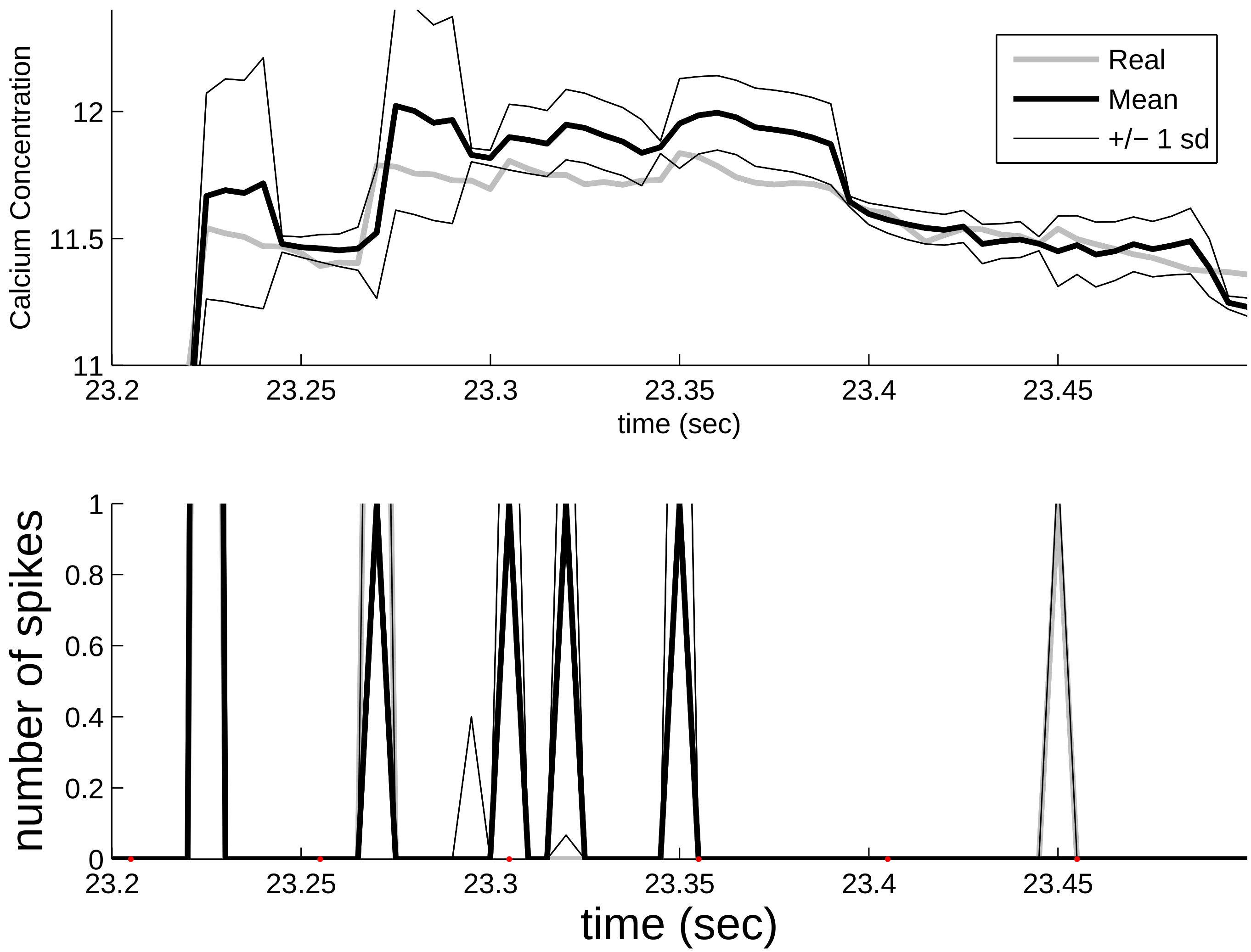
$$dC(t) = -\alpha C(t)dt + \sigma_C dB_{C,t} + \beta n(t) \tag{25}$$

$$O(t) = \eta C(t) + \varepsilon B_t \qquad \text{whenever } t/r \in \mathcal{N} \tag{26}$$

LNP–SH–NICa: Graphical Model



LNP–SH–NICa: Density Estimation



Discussion

- SMC–EM algorithms efficiently infer hidden states and parameters
- Conceptually straightforward generalization to populations of neurons
- Spike dropping scenario can be dealt with similarly
- Synaptic plasticity easily incorporated

References and Acknowledgements

[1] Q. Huys, M. Ahrens, and L. Paninski, "Efficient estimation of detailed single-neuron models," Journal of Neurophysiology, 96:872890, 2006.

[2] Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. Network: Computation in Neural Systems, 15:243262.

[3] Truccolo, W., Eden, U., Fellows, M., Donoghue, J., and Brown, E. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble and extrinsic covariate effects. Journal of Neurophysiology, 93:10741089.

[4] K. Ohki, S. Chung, Y. H. Chng, P. Kara, and R. C. Reid, Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex, Nature, vol. 433, pp. 597603, Feb 2005.

[5] R. D. Roorda, T. M. Hohl, R. Toledo-Crow, and G. Miesenbock, Video-rate nonlinear microscopy of neuronal membrane dynamics with genetically encoded probes, J Neurophysiol, vol. 92, pp. 609621, Jul 2004.

[6] L. R. Rabiner, A tutorial on hidden markov models and selected applications in speech recognition, Proceedings of the IEEE, vol. 72, pp. 257286, February 1989.

[7] A. Doucet, N. de Freitas, and G. Neil, Sequential Monte Carlo Methods in Practice. Springer, 2001.

[8] E. Yaksi and R. W. Friedrich, Reconstruction of firing rate changes across neuronal populations by temporally deconvolved ca2+ imaging, Nat Methods, vol. 3, pp. 377383, May 2006.

[9] Ramdya, Pavan and Reiter, Bettina and Engert, Florian, "Reverse correlation of rapid calcium signals in the zebrafish optic tectum in vivo," J Neurosci Methods, vol. 157, pp. 230–237, Oct 2006.

[10] W. We , M. J. Black, D. Mumford, Y. Gao, E. Bienenstock, J. P. Donoghue, "Modeling and Decoding Motor Cortical Activity Using a Switching Kalman Filter," IEEE Tran Biomed, vol. 51, pp. 933–942, June 2004

We would like to acknowledge Eric Young for helpful discussions. This work was partially supported by NIH grant DC000109.