Tracking fast and slow changes in synaptic weights from simultaneously observed pre- and postsynaptic spiking

Ganchao Wei1\* and Ian H. Stevenson2,3

1Department of Statistics, University of Connecticut

2Department of Psychological Sciences, University of Connecticut

3Department of Biomedical Engineering, University of Connecticut

\*Corresponding Author: ganchao.wei@uconn.edu

# Abstract

Synapses change on multiple timescales, ranging from milliseconds to minutes, due to a combination of both short- and long-term plasticity. Here we develop an extension of the common Generalized Linear Model to infer both short- and long-term changes in the coupling between a pre- and post-synaptic neuron based on observed spiking activity. We model short-term synaptic plasticity using additive effects that depend on the presynaptic spike timing, and we model long-term changes in both synaptic weight and baseline firing rate using Point Process Adaptive Smoothing. Using simulations, we first show that this model can accurately recover time-varying synaptic weights 1) for both depressing and facilitating synapses, 2) with a variety of long-term changes (including realistic changes, such as due to STDP), 3) with a range of pre and post-synaptic firing rates, and 4) for both excitatory and inhibitory synapses. We also show the importance of tracking fast changes in synaptic weights, slow changes in synaptic weights, and unexplained variations in baseline firing simultaneously. Omitting any one of these factors can lead to spurious inferences for the others. Altogether, this model provides a flexible framework for tracking short- and long-term variation in spike transmission.

# Introduction

Detecting synapses from spikes… (cite https://journals.physiology.org/doi/full/10.1152/jn.01170.2003, <https://link.springer.com/chapter/10.1007/978-1-4615-6622-9_1>).

[Explain cross-correlogram] (cite <https://www.cell.com/biophysj/pdf/S0006-3495(67)86597-4.pdf>).

More recent large-scale detection (https://journals.physiology.org/doi/abs/10.1152/jn.00066.2020, <https://www.nature.com/articles/s41467-019-12225-2>). Short-term dynamics can be studied in vivo (<https://www.nature.com/articles/nn.2134.pdf?origin=ppub>, https://www.nature.com/articles/nn0401\_402)

In cases where synapses can be reliably identified from spikes, one possibility is that these recordings can be used to examine changes in synaptic strength over time. These changes occur over multiple timescales and due to different biophysical mechanisms (CITE STP/LTP reviews, eg. https://www.annualreviews.org/doi/abs/10.1146/annurev.physiol.64.092501.114547?casa\_token=SxwZDj40v10AAAAA%3ADEfvxd8dVMRJmA5GdkhGF6Fd2i0HTjBJW0ARMT-vGM0KPhGZY4f9GustaCq40QT9ry\_f\_LU7ajU). To understand learning rules and to make sure our models generalize it is important to separate the short- and long-timescale effects. It is also important to note that fluctuations in the presynaptic rate can create fluctuations in synaptic strength on much longer timescales. For instance, short-term synaptic depression… Several authors have proposed models for estimating long-term (CITE Linderman… Stevenson and Kording, 2011) and short-term (CITE Ghanbari PLoS CB, English et al https://www.sciencedirect.com/science/article/pii/S0896627317309029, Song/Berger…) changes in synaptic weights. Here we introduce a model that estimates both long- and short-term changes simultaneously.

(Unlike English et al.) Ghanbari j neuro – conditioning on the previous presynaptic ISI provides only an incomplete picture. [maybe a discussion point]

Here we develop a statistical model…

# Methods

Inference of synaptic transmission is usually modeled based on Generalized Linear Model (GLM). Here we introduce an extension of a Poisson GLM that9,13 aims to describe both short- and long-term changes in the coupling between a pre- and postsynaptic neuron.

## Model

We model the postsynaptic spiking in discrete time as a doubly stochastic Poisson process with time-varying parameters. Partition the recording time T into , such that with time steps . Denote the total number of presynaptic spikes in as , therefore represents spikes in . The postsynaptic notations and are defined similarly. Make small enough such that both and can only take 0 or 1, and hence we can view and as spike indicators at time . Further, we define the pre-synaptic firing indexes , such that . For , . Therefore, the inter-spike interval between and spikes is .

The postsynaptic neuron’s firing is generated by the following Poisson linear model:

|  |  |  |
| --- | --- | --- |
|  |  | (x.1) |
|  |  | (x.2) |

where is the conditional intensity function at , which depends on baseline firing rate , synaptic plasticity and synaptic connection . We further divide the synaptic plasticity into long-term plasticity (LTP, ) and short-term plasticity (STP, ).

## Estimation of Synaptic Connection

The shape of synaptic connection is described by the alpha-function , with time delay and drop-off . Therefore, the synaptic connection is the convolution of on the pre-synaptic spikes, i.e. . The alpha function is estimated by modeling cross-correlogram, with consideration of both slow effects due to background fluctuations, and fast effect by synaptic connection. Briefly, the model is described as:

(x.3)

where is count rate for cross-correlogram. The slow effect is described as , with baseline and a linear combination of basis functions . The fast effect is described as , where is the connection strength from pre-synaptic neuron to post-synaptic neuron. The fast effect is estimated by a two-stage optimization on penalized Poisson log-likelihood (see details in Ren et al., 2020).

## Estimation of baseline firing rate and synaptic weight

The baseline firing rate and synaptic plasticity (, and ) are changing in two different time- scales: long-term changes ( and ) and short-term changes ().

Different methods are used to estimate changes in these two timescales. The long-term effects are estimated by point process adaptive smoothing, while the short-term effect is estimated by additive effects depending on presynaptic inter-spike intervals (ISIs). The estimation of STP gives a generalized bilinear model (GBLM). These two sets of estimations are combined by alternating optimization, i.e. holding one set of parameters fixed while updating the other and alternating between the two optimizations.

### Long-Term Effects Estimation: Point Process Adaptive Smoothing

Denote the joint parameters for long-term effects as and the remaining term as . Further, define as . Therefore, the conditional intensity function can be re-written as: . The mean and covariance for are estimated by adaptive smoothing, which has two steps: 1) forward algorithm: adaptive filtering, and 2) backward algorithm: Rauch-Tung-Striebel (RTS) smoothing. The model for parameters evolution is as follows:

(x.4)

where is a system evolution matrix and represents Gaussian noise with covariance at . The evolution equation (x.4) gives model prediction (prior) for parameters, at each time step.

Step One: Forward Algorithm (Adaptive Filtering)

Let and be the model predicted mean and covariance, i.e. prior estimations, of parameters at . Let and be the adaptive filtering estimated mean and covariance after observations, i.e. posterior estimations, at . Since , and . Therefore, the adaptive filtering is:

|  |  |  |
| --- | --- | --- |
|  |  | (x.5) |
|  |  | (x.6) |
|  |  | (x.7) |
|  |  | (x.8) |
|  |  | (x.9) |

Equations (x.5) – (x.7) are prior predictions, while equations (x.8) and (x.9) are posterior corrections.

Step Two: Backward Algorithm (RTS Smoothing)

Let and be RTS smoothing estimates for mean and covariance at . The RTS smoothing is:

|  |  |  |
| --- | --- | --- |
|  |  | (x.10) |
|  |  | (x.11) |
|  |  | (x.12) |

To make the algorithm numerically stable, the estimation of mean can be equivalently written as:

(x.13)

In this analysis, is set as time-constant . To give unbiased estimations, is set as identity matrix .

Selection of

The performance of adaptive smoothing is highly affected by , and improper will even make the algorithm diverge. Usually, the is chosen based on previous knowledge. When there’s no sufficient knowledge for , it is usually adjusted by manual trial-and-error approaches. Although we can use the EM algorithm14 to estimate , the convergence is notoriously slow, even with an accelerator15.

Here we choose to estimate by maximizing prediction likelihood, i.e. likelihood under . The is assumed to be diagonal, i.e. noise for ()and ()are unconditionally independent. Besides estimating by direct two-dimensional optimization, the optimized can be approximated by sequential one-dimensional optimization. Although and are assumed to be unconditionally independent, they are not independent conditioning on data. Therefore, we need to be careful to the order of one-dimensional approximation. Since both synaptic connection () and STP () are small, and most synaptic connections are 0, the values of is small. Therefore, the values of has negligible influence on estimation. Based on this observation, we can do one-dimensional approximation as follows: 1) fix and get MLE ; 2) fix as and get MLE .

### Estimation of Short-Term Effect: Additive Effects Depending on presynaptic ISIs

Since the short-term effects are changing too fast, it’s impossible to track them by adaptive smoothing. Here, we model the short-term synaptic plasticity by the following linear model:

|  |  |  |
| --- | --- | --- |
|  |  | (x.14) |
|  |  | (x.15) |

, where is the presynaptic firing time. is a nonlinear function of presynaptic ISI, which gives amplitude of short-term effect under different ISIs. The STP is modeled by convolution of on presynaptic spikes, with exponentially decaying of rate . The nonlinearity of is captured by linear combinations of raised-cosine basis functions , which are functions of presynaptic ISI.

To simplify notations, define , and hence . The short-term modification function is defined as . The variances for STP and modification function can be calculated as:

|  |  |  |
| --- | --- | --- |
|  |  | (x.16) |
|  |  | (x.17) |

# Results

In this study, 1) we generated an example under the assumed GBLM (equation x.1) and implemented commonly used methods (i.e. cross-correlogram and efficacy plots) to analysis the simulated data. 2) Then we implemented our estimation method into several simulations, to see the fitting results under different situation. 3) We further discussed influence of presynaptic firing rates, model misspecification and hyper-parameter selection for long-term effects tracking (i.e. in point process adaptive smoothing, see details in Method). The observed time length is 20min if not specified.

## Traditional Methods for Synaptic Weight Inference

In this section, we generated data under the assumed GBLM in equation x.1. The presynaptic firing rate is 5Hz, while the baseline firing rate for postsynaptic neuron is constantly 20Hz. The LTP is step-changing at 10min, while keeping synaptic weight constantly as 3 before 10min and 5 after 10min. The STP is depression. Traditionally, the baseline firing rate and LTP are detected by splitting cross-correlogram by recording time, and the STP is detected by splitting cross-correlogram according to presynaptic inter-spike intervals (ISIs). Plotting post-synaptic firing efficacy against presynaptic ISIs can also show STP. Figure 1 shows the analysis results by implementing these traditional methods. These traditional methods are easy to implement and can give powerful insight about synaptic weights. However, the traditional methods are just plotting for these effects. That means, we cannot give accurate and detailed inferences on the synaptic weights. Also, the interval estimation is not possible for traditional methods. Here, we provide a way to model and track synaptic weights in detail, and this makes interval estimation possible.

|  |
| --- |
|  |
| **Fig 1. Traditional methods for synaptic weights analysis. (A)** In this example the simulated LTP is step-changing at middle recording time (T/2). **(B)** The change of LTP can be visualized by splitting the cross-correlogram, before and after T/2. The cross-correlation between pre- and postsynaptic spiking is higher after T/2, which suggests a stronger coupling. **(C)** The simulated STP can be viewed by plotting the postsynaptic neuron efficacy against presynaptic inter-spike interval (ISI). Shorter presynaptic ISIs correspond to lower postsynaptic efficacies, which shows that the STP is depression here. The efficacy after T/2 consistently dominates the efficacy before T/2, and this shows abruptly increase of LTP after T/2. **(D)** By splitting cross-correlogram according to T and presynaptic ISI, we can show LTP and STP simultaneously. The cross-correlation is higher for later recording period (>T/2) and longer presynaptic ISIs (> median of simulated ISI). The baselines are the same across all split plots, which suggest a constant postsynaptic firing rate. |

## Simulations

In this section, we use several simulations to see the inference results under different situations, i.e. synapses with 1) different LTP, 2) different STP, 3) different postsynaptic baseline firing rates and 4) different synaptic types. In the following simulations, the presynaptic firing rate are 5Hz.

Inference for long-term and short-term plasticity

First, we use two simulated examples to show inference of linear- and sinusoidal-changing LTP. The baseline firing rate for postsynaptic neuron is constantly 20Hz. The synaptic weight for linear-changing LTP changes linearly from 1.5 to 3.5, and the weight for sinusoidal-changing LTP oscillates around 2.5 with period of 10min. The STP is depression. The LTPs are further shown by splitting cross-correlogram for quartiles of recording time. The left two panels in Figure 2 (A and B) show fitting results for these two simulations.

|  |
| --- |
|  |
| **Figure 2. Inference of different long- and short-term plasticity.** These plots show fitting results for different LTP (orange) and STP (blue). The STP is shown by modification function. The dashed lines show standard error from point estimations. **(A)** The fitting results under different types of LTP, i.e. linear- and sinusoidal-changing. The postsynaptic baseline firing rates are constant and STPs are depression. **(B)** These two LTPs and fitted values can also be visualized by splitting cross-correlogram for quartiles of recording time. **(C)** The fitting results under different types of STP, i.e. depression, facilitation and no STP. The postsynaptic baseline firing rates are constant and LTPs are step-changing at middle recording time. **(D)** Split cross-correlograms for quartiles of presynaptic ISIs show these STPs and fitted values. |

However, the linear- and sinusoidal-changing LTPs are far from real situation. To investigate the fitting results for more realistic LTPs, we further simulate an LTP as the spike-timing dependent plasticity (STDP) and implement our methods. The baseline firing rate for postsynaptic neuron is constantly 15Hz. The modification function for LTP is a traditional double-exponential function that can accurately model STDP in cortical and hippocampal slices. The STP is depression. Similarly, the LTP is shown by the split cross-correlogram by recording length. Figure 3 shows the fitting results.

|  |
| --- |
|  |
| **Figure 3. Inference under STDP generated LTP.** These plots show one example with LTP generated by spike-timing dependent plasticity (STDP) model. The modification function for LTP in STDP model is a double-exponential function. The baseline is constant and STP is depression, and STP is shown by modification function. The simulated values and data are plot in black, while the fitted values are plot in yellow for baseline, orange for LTP, and blue for STP. The dashed lines show standard error from point estimations. **(A), (B) and (C)** show fitted results for baseline, LTP and STP. **(D)** The STP can be visualized by splitting cross-correlogram for quartiles of presynaptic ISIs. |

Then we see the fitting performance under different STPs, i.e. depression, facilitation and no STP cases. Again, the baseline firing rate for postsynaptic neuron is constantly 20Hz. The LTP is step-changing at 10min, and the STPs are set as depression, facilitation and constantly one (no STP) accordingly. The STPs are further shown by splitting cross-correlogram for quartiles of presynaptic ISIs. The right two panels in Figure 2 (C and D) show fitting results for these three simulations.

Inference for postsynaptic baseline firing rates

The long-term dynamics of postsynaptic firing rates can be caused by unexplained factors. To build the unbiased model, we put the unexplained long-term dynamics into baseline line firing rate and make inference on it. Here, we show the fitting results for linear- and sinusoidal-changing postsynaptic baseline firing rates. To see the influence of LTP on baseline estimations, we further simulated step- and linear-changing LTP in each case.

The baseline firing rates for postsynaptic neuron 1) linearly change from 7Hz to 55Hz or 2) oscillate sinusoidally around 20Hz with period of 10min in log-space. The step- and linear-changing LTPs are considered for each case. The STP is depression. The LTPs and baselines are further shown by splitting cross-correlogram for quartiles of recording time. Figure 4 shows fitting results for these four simulations. By these simulations, the baselines are accurately recovered, and dynamics in LTP will not influence baseline estimations. Therefore, we can separate LTP and unexplained long-term dynamics successfully.

|  |
| --- |
|  |
| **Figure 4. Inference of different postsynaptic baseline firing rates.** These plots show fitting results of postsynaptic baseline firing rates (yellow). The STP is shown by modification function. The fitted values for LTP and STP are shown in orange and blue. The dashed lines show standard error from point estimations. **(A)** The fitting results under different types of baseline, i.e. linear- and sinusoidal-changing. The LTPs are step-changing at middle recording time and STPs are depression. **(C)** The same simulation settings as (A), except that the LTP changes linearly. **(B) and (D)** Split cross-correlograms for quartiles of recording time show dynamics of baseline and LTP. |

Inference for inhibitory synapse

The previous simulations are all for excitatory synapses. Here, we show one example for inhibitory synapse. The setting is similar to previous examples, i.e. constant baseline, step-changing LTP and depression STP, but the LTP weight changes from -9 to -3 in this case. Figure 5 shows the fitting results. We can see the fitting results are generally good but a bit worse than excitatory synapses, since we will observe less postsynaptic firing in inhibitory synapses.

|  |
| --- |
|  |
| **Figure 5. Inference of inhibitory synapse.** These plots show one example for estimations on inhibitory synapse, and the STP is shown by modification function. In this example, the baseline firing rate is constant, the LTP jumps at mid-recording-time and STP is depression. The simulated values and data are plot in black, while the fitted values are plot in yellow for baseline, orange for LTP, and blue for STP. The dashed lines show standard error from point estimations. **(A), (B) and (C)** show fitted results for baseline, LTP and STP. By splitting cross-correlogram for quartiles of recording time **(D)** and presynaptic ISIs **(E)**, we can show corresponding LTP and STP. |

## Influence of Presynaptic Firing Rates

Since the LTP effect shows only when there is a presynaptic spike, the presynaptic firing rates will influence LTP estimation accuracy. The larger the firing rates, the more information for LTP estimation and therefore the estimation of LTP will be more accurate. Although STP effect also depends on presynaptic spikes, the STP depends directly on presynaptic ISIs but not the spiking activities itself. In other words, STP is estimated by combined information at different ISIs and presynaptic firing rates will not influence the estimation a lot. Figure 6 shows one example. In this example, the baseline and LTP are constant, and STP is facilitation to show more significant results. The recording time is divided into three parts evenly, and different presynaptic firing rates are assigned to each part respectively (i.e. 3Hz-15Hz-3Hz and 5Hz-0Hz-5Hz), to show influence of presynaptic firing rates. The results show that the more presynaptic spikes, the more accurate the LTP estimation is. The depression STP has similar pattern, but the results are not as significant as facilitation STP.

|  |
| --- |
|  |
| **Figure 6. Influence of presynaptic firing rates on LTP estimation.** In these two examples, the baseline firing rates and LTP are constant. The STP is facilitation to show more significant results. To show the influence of presynaptic firing rates, the recording time is divided into three parts evenly and different presynaptic firing rates are assigned to each part respectively. **(A)** In this simulation, the presynaptic firing rates are 3Hz-15Hz-3Hz. When the firing rate increase, the estimation of LTP is more accurate. **(B)** Now, the presynaptic firing rates are 5Hz-0Hz-5Hz. Lack of the presynaptic spikes leads to a variated LTP estimation. The variation in presynaptic firing rates will not influence estimation of baseline and STP a lot. |

## Model Misspecification

Although the mis-specified model may fit data well overall, it will give spurious inferences for specific synaptic weights. Figure 7 shows one example. In this example, there are fluctuations in baseline and STP, but the LTP is constant. When all three effects (baseline, LTP and STP) are estimated simultaneously, the model is fitted well. However, when we miss to estimate baseline, the fluctuations in baseline will flow into and even be enlarged in LTP estimation. Similarly, if we miss to estimate STP, the LTP estimation will capture the fluctuations in STP.

|  |
| --- |
|  |
| **Figure 7. Miss to estimate the baseline or STP causes spurious LTP estimations.** These plots show one example when missing estimations on baseline or STP. The postsynaptic baseline firing rate fluctuated around exp(3) = 20 Hz, and the presynaptic firing rate fluctuated around 8Hz. The STP is depression. The simulated values and data are plot in black, while the fitted values are plot in yellow for baseline, orange for LTP, and blue for STP. The dashed lines show standard error from point estimations. **(A)** Fitting results when all three effects (baseline, LTP and STP) are estimated simultaneously. **(B)** When we miss to estimate the baseline, the fluctuation in baseline will flow and enlarge in LTP estimation. **(C)** Similarly, the LTP estimation will capture STP fluctuation when we miss to estimate STP. |

## Selection of Hyper-parameter in Adaptive Smoothing

In long-term effects (baseline and LTP) estimations, the covariance for Gaussian noise defines the time scales for tracking. If is too small, the estimations are oversmoothed; if is too large, the estimations are noisy. Here, we estimated the hyper-parameter by maximizing the prediction likelihood (See details in method). The first two panels in figure 8 (A and B) show one example. Both baseline and LTP have Gaussian noise with variance be . Besides doing full 2D optimization (bounded gradient descent in this example), 1D approximated optimization results are also shown in the plot. The slices of prediction log-likelihood show that for baseline has large influence on LTP likelihood, but baseline likelihoods are nearly the same under different for LTP. This shows that 1D approximation is appropriate. As shown in this example, the 2D optimization and 1D approximation give similar results, and they are all close to true .

The panel at right-bottom (C) show estimation performance under different combinations of true . Overall, the estimation of for baseline is better than that for LTP, since LTP fluctuations are shrink by small in the model. The observed time length is 10min for quick plotting. The longer the recording time, the more accurate the estimation is.

|  |
| --- |
|  |
| **Figure 8. Selection and influence of in point process adaptive smoothing. (A)** The heatmap shows prediction log-likelihood under different . The red upward-pointing triangle and dashed lines represent true . The orange dot represents the maximum prediction likelihood estimate (MLE) of . The blue dot is and the green dot is . The one-dimensional MLE is shown in blue downward-pointing triangle. The slices of prediction log-likelihood for LTP and baseline is shown besides the heatmap. **(B)** Corresponding fitted baselines and LTPs, for , and . When is too small, the estimations are over-smoothed; when is too large, the estimations are too noisy. **(C)** The performance of MLE (2D) under different combinations of true , shown by plotting MLE values against true values. Each combination has 5 replicates. The simulated recording time lengths are all 10 min. |

# Discussion

Here we examined a statistical model that aims to model short- and long-term synaptic changes simultaneously. We find that… [review results]

Adaptive baseline may also compensate for slowly varying omitted variables (cite Stevenson).

Multiple inputs (cite Okatan, Truccolo, Pillow, Harris, Volgushev). [How would the model change? What parameters]

Alternative models… STDP, Ca-based LTP/LTD (cite Graupner/Brunel), TM (cite Costa, )… However, the examples here highlight the need for inferring short- and long-timescale effects simultaneously. Ignoring the long-term effects of short-term plasticity or ignoring fluctuations baseline firing can lead to spurious estimates of the synaptic strength.

# References

1. Volgushev, M., Ilin, V. & Stevenson, I. H. Identifying and Tracking Simulated Synaptic Inputs from Neuronal Firing: Insights from In Vitro Experiments. *PLoS Comput. Biol.* **11**, (2015).

2. Ghanbari, A., Malyshev, A., Volgushev, M. & Stevenson, I. H. Estimating short-term synaptic plasticity from pre- and postsynaptic spiking. *PLOS Comput. Biol.* **13**, e1005738 (2017).

3. Chan, R. H. M., Song, D. & Berger, T. W. Tracking temporal evolution of nonlinear dynamics in hippocampus using time-varying volterra kernels. in *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* vol. 2008 4996–4999 (2008).

4. Stevenson, I. H. & Kording, K. Inferring spike-timing-dependent plasticity from spike train data. in *Advances in Neural Information Processing Systems* (eds. Shawe-Taylor, J., Zemel, R. S., Bartlett, P., Pereira, F. C. N. & Weinberger, K. Q.) vol. 24 (2011).

5. Abbott, L. F., Varela, J. A., Sen, K. & Nelson, S. B. Synaptic depression and cortical gain control. *Science (80-. ).* **275**, 221 (1997).

6. Graupner, M. & Brunel, N. Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proc. Natl. Acad. Sci.* **109**, 3991–3996 (2012).

7. Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P. & Brown, E. N. A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects. *J. Neurophysiol.* **93**, 1074–1089 (2005).

8. Harris, K. D., Csicsvari, J., Hirase, H., Dragoi, G. & Buzsáki, G. Organization of cell assemblies in the hippocampus. *Nature* **424**, 552–556 (2003).

9. Brillinger, D. R. Maximum likelihood analysis of spike trains of interacting nerve cells. *Biol. Cybern.* **59**, 189–200 (1988).

10. Okatan, M., Wilson, M. A. & Brown, E. N. Analyzing Functional Connectivity Using a Network Likelihood Model of Ensemble Neural Spiking Activity. *Neural Comput.* **17**, 1927–1961 (2005).

11. Weber, A. I. & Pillow, J. W. Capturing the Dynamical Repertoire of Single Neurons with Generalized Linear Models. *Neural Comput.* **29**, 3260–3289 (2017).

12. Paninski, L. Maximum likelihood estimation of cascade point-process neural encoding models. *Netw. Comput. Neural Syst.* **15**, 243–262 (2004).

13. Brillinger, D. R. Nerve Cell Spike Train Data Analysis: A Progression of Technique. *J. Am. Stat. Assoc.* (1992) doi:10.2307/2290256.

14. Ananthasayanam, M. R., Mohan, M. S., Naik, N. & Gemson, R. M. O. A heuristic reference recursive recipe for adaptively tuning the Kalman filter statistics part-1: formulation and simulation studies. *Sādhanā* **41**, 1473–1490 (2016).

15. Du, Y. & Varadhan, R. SQUAREM: An R Package for Off-the-Shelf Acceleration of EM, MM and Other EM-Like Monotone Algorithms. *J. Stat. Software; Vol 1, Issue 7*  (2020) doi:10.18637/jss.v092.i07.