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.

# 1 Introduction

Synapses play an important role in neurophysiology, although detecting and characterizing these synapses is not easy. The extracellular spiking activities of neurons provide a way for synapses detection, and it is usually done by analyzing the cross-correlogram(Barthó et al., 2004; Fetz et al., 1991). The cross-correlogram is a histogram of postsynaptic spiking times forward and backward to presynaptic spikes, which provides an estimate of cross-correlation between pre- and post-synaptic neurons(Perkel et al., 1967). When a monosynaptic occurs, the postsynaptic spiking probability will often increase or decrease after a presynaptic spike, for an excitatory or inhibitory synapse. This corresponds to a transient and short-latency peak or trough in the cross-correlogram. However, investigating the cross-correlogram directly may not be enough for modern neural data analysis. By using multielectrode arrays, the extracellular spiking of hundreds of neurons can be recorded simultaneously. These large-scale recorded neurons can have tens of thousands of potential synapses between them, but both spiking and synapses can be sparse. This makes it difficult for synapses detection by simply plotting the cross-correlograms. Some recent research tried to solve this problem by modeling spike cross-correlations to handle the sparse information for detection(Kobayashi et al., 2019; Ren et al., 2020).

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. Changes in synaptic strength occur over multiple timescales and due to different biophysical mechanisms(Zucker & Regehr, 2002). For instance, on timescales of a few milliseconds, the synaptic strength may decrease by consumption of neurotransmitters, or may increase by influx of calcium into presynaptic axon terminal, which increases neurotransmitters release probability. In contrast, increases or decreases in receptor density or structural changes occur on timescales of minutes to hours. 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, higher presynaptic rate will produce a weaker synaptic strength, if the synapse has short-term depression(Swadlow & Gusev, 2001).

The short- and long-term dynamics in synapses can be roughly studied by investigating the cross-correlations or postsynaptic spiking probability, under different presynaptic inter-spike interval (ISI)(Carandini et al., 2007; Csicsvari et al., 1998; Fujisawa et al., 2008; Mantel & Lemon, 1987; Swadlow & Gusev, 2001; Usrey et al., 2000) and different recording time respectively. Several authors have proposed models for estimating either long-term(Linderman et al., 2014; Stevenson & Kording, 2011) or short-term(Chan et al., 2008; English et al., 2017; Ghanbari et al., 2017) changes in synaptic weights. Unlike English et al.(English et al., 2017), the short-term plasticity estimation is conditioned on the previous presynaptic inter-spike intervals (ISIs) in Ghanbari et al.(Ghanbari et al., 2020), and this provides only an incomplete picture for synaptic strength inference. Here we introduce a statistical model that estimates both long- and short-term changes simultaneously. The unexplained changes in postsynaptic neuron are treated as the baseline firing rate. The long-term changes in both synaptic weight and baseline is estimated by the point process adaptive smoothing(Eden et al., 2004), while the short-term plasticity is modeled by additive effects that depend on the presynaptic spike timing(Stevenson & Kording, 2011). By several simulations, we show that this model provides a flexible framework for tracking short- and long-term variation in spike transmission, and its necessary to estimate both long-term and short-term effects simultaneously.

# 2 Methods

Here we introduce an extension of a Poisson GLM that aims to describe the coupling between a pre- and postsynaptic neuron(D R Brillinger, 1988; David R. Brillinger, 1992). While many previous studies(Chan et al., 2008; English et al., 2017; Ghanbari et al., 2017; Linderman et al., 2014; Stevenson & Kording, 2011) have modeled static coupling between neurons, our goal here is to describe a time-varying synaptic strength with both short- and long-term changes.

We model the postsynaptic spiking in discrete time as a doubly stochastic Poisson process with time-varying parameters. Partitioning the total recording time T into evenly-spaced bins , such that with time steps , we denote the total number of presynaptic spikes in as and represents the number of spikes observed in . and denote the same calculations for the postsynaptic neuron’s spikes. For small enough both and take values of 0 or 1 and can be viewed as spike indicators for time bin .

Previous models of static coupling between neurons typically include the recent spiking history of both the presynaptic input (the coupling effect) and the postsynaptic neuron itself.

|  |  |  |
| --- | --- | --- |
|  |  | (2.1) |
|  |  | (2.2) |

where is the conditional intensity of postsynaptic neuron at , given the recent spiking of the postsynaptic neuron and presynaptic neuron , steps into the past. For the model parameters, defines the fixed baseline firing rate and the coupling and history effects are weighted by and **,** for the post- and pre-synaptic neurons respectively.

Here we extend this static model to include fast and slow dynamics in the coupling effect and model the postsynaptic neuron’s firing rate as

|  |  |  |
| --- | --- | --- |
|  |  |  |
| (2.3) |
|  |  | (2.4) |

In our model, the spiking history of postsynaptic neuron is captured by the unstructured time-varying baseline . In the static coupling model, the main goal is to accurately infer the shapes of the filters, here we assume that the shape is fixed , but is weighted by an additional factor that varies over time. Additionally, we aim to partition the variations in into a long-term component and a short-term component that scales the long-term synaptic weight multiplicatively. Here we model transient increases/decreases in as a function of the presynaptic spike timing to account for short-term synaptic facilitation/depression. In the absence of presynaptic activity, returns to a base value of . To simplify the notation, we write , , and as , , and below.

## Estimating the Synaptic Filter

Here we estimate the shape of the static presynaptic filter by directly modeling the cross-correlogram between the pre- and postsynaptic spikes, similar to Ren et al(Ren et al., 2020). Briefly, we assume that the shape of the synaptic connection is described by an alpha function , with a latency and time constant . These parameters are then estimated by modeling the cross-correlogram as a combination of a slow background correlation and a fast, transient effect of the synaptic connection:

|  |  |  |
| --- | --- | --- |
|  |  | (2.5) |
|  |  | (2.6) |
|  | and | (2.7) |

where denotes the expected rate of coincidences from the cross-correlogram with baseline and a linear combination of smooth basis functions accounting for slow changes in cross-correlation, at time bin . Here we use cubic B-spline bases with equally spaced knots for the smooth basis and model the time-range in bins. The fast effect is described as where is the connection strength from pre-synaptic neuron to post-synaptic neuron, and the alpha function is convolved with the auto-correlation of the presynaptic neuron to account for the effects of presynaptic dynamics. The parameters are estimated by maximizing the Poisson log-likelihood. We use random restarts, since the objective function is non-convex in the latency and time-constant parameters.

After fitting the latency and time constant of the synaptic filter to the cross-correlogram we assume that these parameters are fixed when modeling the long- and short-term changes in synaptic strength. This simplifying assumption allows us to model rescaling of the basic presynaptic input, given by and avoid a computationally intensive non-convex optimization of the full likelihood with respect to and .

## Estimating long-term changes in baseline firing rate and synaptic weight

To estimate the time-varying baseline firing rate and the effects of synaptic plasticity (, and ) we use two distinct strategies. For the long-term effects we estimate and by point process adaptive smoothing(Eden et al., 2004; Rauch et al., 1965), and for the short-term effects we estimate using an additive model that depends on the presynaptic inter-spike intervals (ISIs)(Stevenson & Kording, 2011). To estimate all of the effects together we use an alternating optimization – we hold the long-term effects constant while updating the short-term effects then hold the short-term effects constant while updating the long-term effects and repeat this alternating pattern until convergence.

### Estimating Long-Term Effects

To model the long-term changes in baseline firing and synaptic strength we assume that the parameters evolve over time with noisy, linear dynamics. Denoting the parameters as a vector , we assume that the model parameters evolve over time following

|  |  |  |
| --- | --- | --- |
|  |  | (2.8) |

where is a system evolution matrix and represents Gaussian noise with covariance at . The conditional intensity can be re-written as where . We then use adaptive smoothing to track estimates of the parameters given the observed spiking of the pre- and postsynaptic neurons. During a forward step, we first approximate the distribution using adaptive filtering(Eden et al., 2004). Then, during a backward step, we approximate the distribution using Rauch-Tung-Striebel (RTS) smoothing(Rauch et al., 1965). In both cases, we approximate the distribution over using a multivariate Gaussian.

For adaptive filtering, we assume an initial mean and covariance and . We first propagate the estimated mean and covariance forward in time according to the process model

|  |  |  |
| --- | --- | --- |
|  |  | (2.9) |
|  |  | (2.10) |

Here and denote the predicted mean and covariance given observations up to time . We then update the mean and covariance based on the observed spiking at time .

|  |  |  |
| --- | --- | --- |
|  |  | (2.11) |
|  |  | (2.12) |
|  |  | (2.13) |

Here and are the resulting mean and covariance after incorporating the observation at . These updates were previously derived for a general Poisson model here (Eden et al., 2004).

Given the estimates from adaptive filtering we then step backwards to find smooth estimates of the parameters. Here we use updated based on the RTS method

|  |  |  |
| --- | --- | --- |
|  |  | (2.14) |
|  |  | (2.15) |
|  |  | (2.16) |

where and denote smoothed estimates for the mean and covariance at , and to make the algorithm numerically stable, we use an equivalent update

|  |  |  |
| --- | --- | --- |
|  |  | (2.17) |

In the results that follow we assume that the process covariance is constant , and that the parameter evolution is a random walk .

The performance of adaptive smoothing is highly affected by , and choosing improperly can prevent the algorithm from converging. Estimating from the data itself using the EM algorithm(Ananthasayanam et al., 2016) is notoriously slow, even with an accelerator(Du & Varadhan, 2020). Here we choose to estimate by maximizing prediction likelihood, i.e. likelihood under . To simplify, we assume that is diagonal with independent noise for () and ().

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 .

### Estimating Short-Term Effects

In addition to the slow changes in baseline firing and the synaptic weight, we also aim to model fast changes in synaptic weights due to short-term synaptic plasticity. These changes occur on timescales much faster than the typical postsynaptic ISI (~10 ms), and cannot be accurately tracked by adaptive smoothing. Rather than using smoothing, we thus model short-term synaptic plasticity using a parametric model previously introduced in Ghanbari et al.(Stevenson & Kording, 2011)

|  |  |  |
| --- | --- | --- |
|  |  | (2.18) |
|  |  | (2.19) |

Here is the presynaptic firing time and the denotes the inter-spike interval (ISI) between the and presynaptic spikes, and is a nonlinear function of presynaptic ISI, which describes how the synaptic strength increases or decreases following the a pair of presynaptic spikes with a specific ISI. decreases synaptic strength and mimicking short-term synaptic depression, while creates increases in synaptic strength akin to facilitation. The cumulative effects of STP are modeled by a convolution of with the presynaptic spikes, and we assume that the effects decay exponentially with rate . We parametrize the shape of using a linear combination of raised-cosine basis functions .Note that, since we can rewrite the short-term synaptic effect , optimizing (when the long-term effects are fixed) is simply a matter of fitting a GLM.

Additionally, assuming that the long-term effects are fixed allows us to approximate standard errors for the cumulative effects of STP and the modification function (defined as ) itself

|  |  |  |
| --- | --- | --- |
|  |  | (2.20) |
|  |  | (2.21) |

To summarize, altogether, we have the following parameters: , , , and , with hyper-parameters defining the raised-cosine basis and timescale for short-term modifications, and the long-term weight process covariance . To fit these parameters we first estimate the synaptic latency and time-constant directly from the cross-correlogram. Then we optimize the remaining parameters by alternating between fitting the short-term parameters assuming , fixed and fitting the long-term parameters , assuming fixed.

## Simulations

Here we validate the model using simulated pre- and postsynaptic spiking. If not otherwise specified, presynaptic spike times are generated by a homogeneous Poisson process with a firing rate of 5Hz. The postsynaptic neuron is simulated as a conditionally Poisson process defined in equation 2.3 and 2.4. The observed time length is 20min if not specified, and we use bin size 1ms, throughout.

For simulating spike-timing-dependent plasticity (STDP) the use a long-term modification function that depends on the relative timing of pre- and postsynaptic spikes. Here we use a double-exponential modification function, based on the STDP observed in cortical and hippocampal slices. In this case, each pair of pre- and post-synaptic spikes modifies the synapse by

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |
| (2.22) |
|  |
|  |

And we set , , and . We further add an additional long-term decaying factor that pushes the synaptic weights back to 1, as in Stevenson and Kording(Stevenson & Kording, 2011). Namely,

|  |  |  |
| --- | --- | --- |
|  |  | (2.23) |

where is the indicator function, and we set .

In the following results, 5 raised-cosine bases with non-linear stretching peaks are used to model STP modification function within [0, 600ms] in 1ms bins. When fitting the model without tuning of , it is set as in all results except for those in section “Selection of Hyper-parameter in Adaptive Smoothing”.

Code for the model and all simulations is available at http://github.com/weigcdsb/GBLM\_SMOOTH/.

# 3 Results

Here we present simulation results illustrating how this model performs, how the hyperparameters can be efficiently optimized, and why tracking *both* short- and long-term effects can prevent spurious inference of synaptic dynamics. The recording lengths are 20min and presynaptic firing rates are 5Hz, if not specified.

## Simultaneous short-term and long-term changes in synaptic weights

Although most experimental paradigms aim to manipulate and/or observe only a single synaptic timescale, short-term synaptic dynamics (depression and facilitation) and long-term synaptic dynamics (LTD and LTP) coexist and have distinct physiological mechanisms. To illustrate how these dynamics can coexist we simulate an excitatory synaptic connection between a pre- and postsynaptic neuron where the synaptic weight increases instantaneously midway through a 20min recording while also having short-term synaptic depression on fast timescales (Fig 1). Presynaptic spike timing is generated using a Poisson GLM with a history effect that mimics refractoriness, while postsynaptic spike times are generated by our full model. The overall cross-correlogram between the pre- and postsynaptic shows a short latency, fast onset peak where the probability of postsynaptic spiking increases following each presynaptic spike (Fig 1B, right). Qualitatively, this peak is like putative excitatory connections observed in large-scale multielectrode recordings.

Several previous studies have characterized the short-term and long-term dynamics of spike transmission by partitioning or splitting the overall cross-correlogram to use only a subset of presynaptic spikes. For instance, comparing the cross-correlogram generated using only the presynaptic spikes before the change-point to the cross-correlogram generated using the spikes after the change-point reveals the increase in synaptic strength (Fig 1B). By using presynaptic spikes that are preceded by a specific inter-spike interval we can observe the effects of short-term synaptic depression (Fig 1D). In this case, presynaptic spikes that occur recently after another spike tend to have lower efficacy compared to spikes that occur following a long period of silence, where the synaptic resources have had an opportunity to recover (Fig 1C). Although partitioning the cross-correlogram can reveal clear evidence of long- and short-term changes in spike transmission, it is often unclear what the best partitioning should be and how to combine evidence from multiple types of partitions (e.g. time-based and ISI-based) into a single description. It is also important to note that in the simulation here the firing rate for presynaptic neuron is set to 5Hz and the baseline firing rate of postsynaptic neuron is set to 20Hz. In cases where firing rates are lower and/or when synaptic efficacy is weaker the partitioned cross-correlograms may be too noisy to obtain meaningful estimates of the synaptic weight.

|  |
| --- |
|  |
| **Fig 1. Short-term and long-term plasticity in an excitatory synapse. (A)** 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. |

Using a model-based approach the short- and long-term synaptic weights can be estimated simultaneously to create a unified description of postsynaptic spiking. Here we track the long-term changes in synaptic weight using adaptive smoothing, and we fit the short-term changes in synaptic weights using a modification function that describes the additive effects for different presynaptic ISIs (see Methods).

Using adaptive smoothing, the model simply updates its estimate of the synaptic weight based on the observed spike transmission. If postsynaptic spikes follow presynaptic spikes more than expected the parameter for the long-term synaptic weight increases, and if postsynaptic spikes follow presynaptic spikes less than expected the parameter decreases. The hyperparameter provides a constraint on how fast the long-term synaptic weight parameter can change. This unstructured approach allows the model to track a wide variety of patterns (Fig 2A). Since the model estimates a synaptic effect for every presynaptic spike, it also provides efficacy estimates for arbitrary partitions of the cross-correlogram (Fig 2B).

For the short-term synaptic weight, we use a more structured approach that depends on the specific pattern of presynaptic ISIs. Following each presynaptic spike, we assume that the short-term synaptic weight increases or decreases by an amount that depends on the interspike interval (ISI) preceding that spike. In the absence of presynaptic spikes, we assume that the short-term weight decays (exponentially) back to a baseline weight. These changes can occur rapidly but are constrained by the presynaptic spike timing. Within the model, presynaptic spikes following an interval of 100ms, for instance, will always modify the short-term synaptic weight by the same amount. However, these structured short-term changes allow a range of behaviors, including short-term facilitation and depression-like effects (Fig 2C). These effects occur simultaneously with any long-term changes in synaptic weight, and, again, allow for efficacy estimates with arbitrary partitions of the cross-correlogram (Fig 2D).

|  |
| --- |
|  |
| **Figure 2. Inference of different long- and short-term plasticity.** These plots show fitting results for different LTP and STP (in 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. |

During natural ongoing spiking activity there are unlikely to be large shifts in long-term synaptic strength like Fig 1 and 2. In the absence of external stimulation or task demands, long-term synaptic changes are more likely to result from ongoing patterns of activity. Here we simulate a synapse with spike timing-dependent plasticity (STDP). The modification function is a traditional double-exponential function, similar to data from cortical and hippocampal slices, where the synapse is strengthened when postsynaptic spikes follow presynaptic spikes and weakened when presynaptic spikes follow postsynaptic spikes. With the presynaptic spike timing again coming from a Poisson GLM, STDP induces slow fluctuations in the synaptic strength of the full model (Fig 3A). These changes are accurately tracked by the adaptive smoother, even though the smoother does not model pre-post spike timing explicitly. As before, these slow changes also exist simultaneously with short-term synaptic depression (Fig 3B). Although the short-term synaptic effects are visible when partitioning the cross-correlograms (Fig 3C), it is unclear how the long-term fluctuations in STDP could be similarly partitioned.

|  |
| --- |
|  |
| **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 firing rate for postsynaptic neuron is set to 15Hz and the STP is depression. Dashed lines show standard error. **(A)** and **(B)** show fitted results for LTP and STP (modification function). **(C)** The STP can be visualized by splitting cross-correlogram according to quartiles of presynaptic ISIs. |

The previous simulations are all for excitatory synapses. However, the model works similarly for inhibitory synapses. In practice, the “sign” of the synapse is determined by the signs of the long-term and short-term synaptic effects. For simplicity, we assume that the short-term synaptic effect is positive and decays to a baseline value of such that the sign of the long-term effect effectively determines whether the synapse is excitatory or inhibitory . As before, both short- and long-term synaptic effects can be estimated when simulating from the full model (Fig 4).

|  |
| --- |
|  |
| **Figure 4. Inference of inhibitory synapse.** These plots show one example for estimations on inhibitory synapse. The baseline firing rate is for postsynaptic neuron is set to 400Hz, the LTD changes instantaneously midway and STP is depression. Dashed lines show standard error. **(A)** and **(B)** show fitted results for LTD and STP (modification function). By splitting cross-correlogram for quartiles of recording time **(C)** and presynaptic ISIs **(D)**, we can show corresponding LTD and STP. |

## Variation in pre- and postsynaptic firing rates

In addition to modeling fast and slow changes in synaptic strength we also model slow, unexplained fluctuations in the baseline postsynaptic firing rate. Here we include the baseline parameter in the adaptive smoother, and, as with the long-term synaptic weight, it is constrained to vary slowly using the hyperparameter .

To illustrate the effects of a time-varying baseline we simulate, again, a presynaptic neuron with Poisson GLM spiking that provides excitatory input to a postsynaptic neuron simulated by the full model. Here the synapse undergoes short-term synaptic depression as well as long-term changes in strength, as before, and, in addition, the firing rate of the postsynaptic neuron fluctuates (Fig 5). In general, the baseline is estimated more accurately and with higher certainty than the long-term synaptic weight. Information about the presence or absence of postsynaptic spikes is always available, but information about the synaptic weight is only available in the short time window following each presynaptic spike. This difference also creates a kind of separability where estimates of the baseline are not particularly influenced by the changes in synaptic weight and estimates of synaptic weight are not influenced by the changing baseline (Fig 5A and C).

As in the previous examples, since there is an estimate of the postsynaptic rate at every time, using a model-based approach allows for reconstruction of arbitrary partitions of the cross-correlogram. When the baseline fluctuates it can also have an influence on the cross-correlogram (Fig 5B and D). Note that here, the baseline affects both the base level of the cross-correlogram as well as the magnitude of the synaptic effect. With the exponential output nonlinearity, the time-varying baseline acts like a time-varying gain and may be useful to mimic slow fluctuations in excitability due to neuromodulation or brain state.

|  |
| --- |
|  |
| **Figure 5. Inference of different postsynaptic baseline firing rates.** These plots show fitting results of postsynaptic baseline firing rates. **(A)** The fitting results under different types of baseline, i.e. linear- and sinusoidal-changing. The LTPs are jumps at midway and STPs (shown in modification functions) are depression. **(C)** The same simulation settings as in **(A)**, except that the LTP changes linearly. **(B)** and **(D)** Split cross-correlograms for quartiles of recording time show dynamics of baseline and LTP. |

Since information about the synaptic strength is only available when there is a presynaptic spike, the presynaptic firing rate influences how accurately the synaptic weight can be estimated. Higher presynaptic firing rates allow the long-term synaptic strength to be estimated more accurately and with less uncertainty (Fig 6). To illustrate this feature of the model we simulate presynaptic spiking with a Poisson GLM whose rate changes abruptly, either increasing (Fig 6A) or decreasing (Fig 6B). The postsynaptic neuron spikes according to the full model, in this case, with short-term synaptic facilitation. The baseline postsynaptic rate and long-term synaptic weight are both constant in this simulation. However, the accuracy and variance of the adaptive smoothing estimates for the long-term weight change substantially when the presynaptic rate changes. Estimates for the short-term modification function will also depend on the presynaptic spike rate to some extent, but the impact on accuracy is not as strong, since the short-term model pools information from presynaptic across the entire recording depending on their ISIs.

|  |
| --- |
|  |
| **Figure 6. Influence of presynaptic firing rates on LTP estimation.** In these two examples, the baseline firing rates are constantly 55Hz, LTP are constant and the STP are 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)** 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. Dashed lines show standard error. |

## Omitted variable bias

In the examples above we have shown how estimates of the baseline and synaptic weight are largely separable, since the synaptic effects are constrained to follow presynaptic spikes, and estimates of the short-term and long-term synaptic effects are largely separable, since short-term effects are constrained to be ISI-dependent. However, omitting components of the model can result in spurious estimates for the long-term synaptic weight.

To illustrate how omitting effects can bias estimation, we again simulate a presynaptic neuron with Poisson GLM spiking that provides synaptic input to a postsynaptic neuron whose firing is determined by the full model. Here we simulate a synapse with short-term synaptic depression and a constant long-term weight. Although in the previous examples the presynaptic baseline has been held constant, in this case we simulate a fluctuating baseline for both the pre- and postsynaptic neurons. Slow fluctuations in the presynaptic rate induce slow changes in the synaptic weight that are purely due to short-term synaptic effects (Fig 7). For the depressing synapse simulated here, a high presynaptic rate causes the synapse to be in a chronically depleted state, while a lower presynaptic rate allows the synapse to recover from depression. In this case, the short-term synaptic weight is negatively correlated with the presynaptic rate.

These slow fluctuations in the short-term synaptic weight are not necessarily problematic and can be accurately estimated with the full model (Fig 7A). However, if we instead fit a model that only includes a long-term synaptic weight and omits the short-term effect, the slow fluctuations are misattributed to the long-term weight (Fig 7B). The adaptive smoother effectively tracks these changes when they are not otherwise explained. Since natural neural activity contains these types of slow fluctuations in presynaptic rate (as a function of brain state, for instance) there are likely to be slow fluctuations in synaptic strength that can be explained away as byproducts of short-term synaptic mechanisms. Modeling short- and long-term effects simultaneously can allow these “explainable” fluctuations to be separated from other phenomena, such as LTP/LTD or STDP.

Omitting the fluctuations in the baseline can also result in spurious estimates of the synaptic strength (Fig 7C). Here, when the model is fit while assuming that the baseline is constant, the adaptive smoother aims to account for higher/lower than expected firing rate using the long-term synaptic weight. Although the synaptic effect is limited to the brief interval following each presynaptic spike, misattributing the rate changes to fluctuations in the synaptic weight improves the overall likelihood. In this case, there is a strong positive correlation between the (omitted) baseline fluctuations and the misestimated long-term synaptic weight. Tracking unexplained variation in the postsynaptic rate is thus likely to be important for accurately tracking synaptic weights in experimental data.

In these scenarios, partitions of the cross-correlogram may serve as a useful check on how well the model describes spike transmission in specific time-periods or as a function of ISI. The model that omits STP, for instance, will fail to explain the observed depression as a function of ISI, while the model that assumes a constant baseline will fail to model the changing base level of the cross-correlograms when they are partitioned over time.

|  |
| --- |
|  |
| **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 20 Hz, and the presynaptic firing rate fluctuated around 8Hz. The STP is depression. Standard error is shown by dashed lines. **(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 be misattributed to LTP. **(C)** Similarly, the LTP estimation will capture STP fluctuation when we miss to estimate STP. |

## Optimizing the adaptive smoother

For the baseline and long-term synaptic weight, the process noise covariance defines the timescales for tracking. In the examples above we use a fixed during inference. In general, however, should be matched to the underlying timescale of the process. If is too small the estimated baseline and synaptic strength will be oversmoothed and potentially meaningful changes may be underestimated. On the other hand, if is too large the estimated baseline and synaptic strength undersmoothed and may reflect noise. Here we assume that is diagonal and show how the variance of the baseline and synaptic weight and can be optimized by maximizing the prediction likelihood (see Methods).

To illustrate the process of optimizing we simulate from the full model and sample the baseline and long-term synaptic weight using a Gaussian random walk with a ground truth . Although the full likelihood always improves with larger values of , the prediction likelihood has a single maxima as a function of and (Fig 8A). We find that changes in have a much larger effect on the prediction likelihood than changes in , presumably due to the fact that the synaptic weight only influences the likelihood in the short interval follow each presynaptic spike. As mentioned in the method and shown in the example of omitted variable bias (Figure 7), values of has negligible influence on prediction likelihood under fixed but no vice versa. Therefore, 2D optimization problem can be solved by optimizing the two variance hyperparameters sequentially with a specific order. Full 2D optimization of the prediction likelihood (bounded gradient descent in this example) accurately recovers the true process noise, as does a simple 1D optimization scheme where we first optimize with then optimize using the optimized in the first step. The corresponding tracks of the baseline and long-term synaptic weights under different (Figure 8B) show oversmoothness and undersmoothness.

We find that in a series of short, 10min simulations with pre- and postsynaptic rates be 5Hz and 20Hz, the true values of can be accurately recovered (Fig 8C) by maximizing the prediction likelihood. As before, the accuracy in tracking the baseline is typically higher than the accuracy in tracking the synaptic weight. The accuracy in recovering the baseline hyperparameter is also typically higher than the accuracy in recovering the synaptic weight hyperparameter.

|  |
| --- |
|  |
| **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) . The blue dot is and the green dot is . The one-dimensional approximated 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 under , 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 10 min in panel C. |
|  |

# 4 Discussion

Here we introduce a statistical model that aims to simultaneously track short- and long-term changes in synaptic weights from spike observations. Using simulations, we show that our model can successfully recover time-varying synaptic weights for both excitatory and inhibitory synapses with a variety of short-term and long-term plasticity patterns. The LTP estimation accuracy depends on presynaptic firing rates: the more presynaptic neuron fires, the more accurate the LTP estimation. We found that the three pieces (unexplained variations in baseline, LTP and STP) should be estimated simultaneously. If we omit any one of these factors, the inferences for others will be biased. Finally, we provide two efficient way, i.e. 2D MLE and 1D approximation, to choose the appropriate hyper-parameters in adaptive smoothing.

In GLMs, omitted variables can result in biased parameter estimates for the effect that are included in the model(Stevenson, 2018). Some previous research for synaptic strength inference had tried to tackle this problem by including postsynaptic history effects(Ghanbari et al., 2017, 2020; Stevenson & Kording, 2011; Truccolo et al., 2005). However, the postsynaptic history may not capture all effects that have impact on neural activity, and the history effects can also be biased by omitted variables. Therefore, we allow the postsynaptic baseline firing rates to be time-varying to compensate for slowly varying omitted variables. By conditioning on the adaptive baseline, the inference on synaptic strengths will less influenced by unexplained variation in the postsynaptic firing rate. However, the true postsynaptic rate is not necessarily identifiable(Amarasingham et al., 2015). Synaptic weights estimated from in vivo recordings are still likely to be affected by unobserved variables, particularly if these variables induce firing rate changes on fast timescales or on multiple timescales.

Adaptive filtering for place fields(Brown et al., 2001) is widely used, although there are some limitations. Although the method is locally stable, the global stability should also be investigated. Moreover, some spiking activities are more or less variable than Poisson process, and this suggests the necessity for non-Poisson extension. For example, we can extend the method based on the Conway-Maxwell-Poisson (COM-Poisson)(Stevenson, 2016) likelihood to handle both over- and under-dispersion in spike counts.

This model aims to track time-varying synaptic weights from simultaneous extracellular recordings from a pre- and postsynaptic neuron. For the sake of simplicity, we assume that monosynaptic connections can be accurately identified. However, detecting synaptic connections from large-scale multielectrode recordings is not necessarily straightforward, particularly for weak connections or short recording times. Correlation…

Modeling a network of connected neuron could also benefit from modeling a state-space(Paninski et al., 2010; Smith & Brown, 2003). Specifically, assume there are Combining information from a population of spiking neurons could potentially allow for excitability fluctuations to be estimated at timescales faster than what can be estimated from a single neuron alone.

Several alternative statistical and biophysical models were used to describe LTP (e.g. STDP and Ca-based LTP/LTD(Graupner & Brunel, 2012)) and STP (e.g. TM(Costa et al., 2013)). 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. Also, recent studies(Deperrois & Graupner, 2020) show that short-term and long-term synaptic weight can together tune sensitive range of synaptic plasticity, and this suggest the necessity for considering effects in different timescale at the same time.

This model provides a flexible framework for tracking short- and long-term variation in spike transmission, and more detailed modeling can be done based on the framework. For example, the LTP is currently modeled in an unstructured way to provide an unbiased reference. We can replace it by some more biophysical meaningful models, such as STDP and Ca-based LTP/LTD(Graupner & Brunel, 2012), to make a more detailed inference. Extends static models of coupling between neurons.

# Acknowledgements

This work was supported by NSF CAREER 1931249 to IHS. Thanks to Naixin Ren for helpful comments.

# References

Amarasingham, A., Geman, S., & Harrison, M. T. (2015). Ambiguity and nonidentifiability in the statistical analysis of neural codes. *Proceedings of the National Academy of Sciences*, *112*(20), 6455–6460. https://doi.org/10.1073/PNAS.1506400112

Ananthasayanam, M. R., Mohan, M. S., Naik, N., & Gemson, R. M. O. (2016). A heuristic reference recursive recipe for adaptively tuning the Kalman filter statistics part-1: formulation and simulation studies. *Sādhanā*, *41*(12), 1473–1490. https://doi.org/10.1007/s12046-016-0562-z

Barthó, P., Hirase, H., Monconduit, L., Zugaro, M., Harris, K. D., & Buzsáki, G. (2004). Characterization of neocortical principal cells and interneurons by network interactions and extracellular features. *Journal of Neurophysiology*, *92*(1), 600–608. https://doi.org/10.1152/jn.01170.2003

Brillinger, D R. (1988). Maximum likelihood analysis of spike trains of interacting nerve cells. *Biological Cybernetics*, *59*(3), 189–200.

Brillinger, David R. (1992). Nerve Cell Spike Train Data Analysis: A Progression of Technique. *Journal of the American Statistical Association*. https://doi.org/10.2307/2290256

Brown, E. N., Nguyen, D. P., Frank, L. M., Wilson, M. A., & Solo, V. (2001). An analysis of neural receptive field plasticity by point process adaptive filtering. *Proceedings of the National Academy of Sciences*, *98*(21), 12261–12266. https://doi.org/10.1073/PNAS.201409398

Carandini, M., Horton, J. C., & Sincich, L. C. (2007). Thalamic filtering of retinal spike trains by postsynaptic summation. *Journal of Vision*, *7*(14), 20–20. https://doi.org/10.1167/7.14.20

Chan, R. H. M., Song, D., & Berger, T. W. (2008). Tracking temporal evolution of nonlinear dynamics in hippocampus using time-varying volterra kernels. *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *2008*, 4996–4999. https://doi.org/10.1109/IEMBS.2008.4650336

Costa, R., Sjostrom, P. J., & van Rossum, M. (2013). Probabilistic inference of short-term synaptic plasticity in neocortical microcircuits. *Frontiers in Computational Neuroscience*, *7*, 75. https://doi.org/10.3389/fncom.2013.00075

Csicsvari, J., Hirase, H., Czurko, A., & Buzsáki, G. (1998). Reliability and state dependence of pyramidal cell-interneuron synapses in the hippocampus: An ensemble approach in the behaving rat. *Neuron*, *21*(1), 179–189. https://doi.org/10.1016/S0896-6273(00)80525-5

Deperrois, N., & Graupner, M. (2020). Short-term depression and long-term plasticity together tune sensitive range of synaptic plasticity. *PLOS Computational Biology*, *16*(9), 1–25. https://doi.org/10.1371/journal.pcbi.1008265

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

Eden, U. T., Frank, L. M., Barbieri, R., Solo, V., & Brown, E. N. (2004). Dynamic Analysis of Neural Encoding by Point Process Adaptive Filtering. *Neural Computation*, *16*(5), 971–998. https://doi.org/10.1162/089976604773135069

English, D. F., McKenzie, S., Evans, T., Kim, K., Yoon, E., & Buzsáki, G. (2017). Pyramidal Cell-Interneuron Circuit Architecture and Dynamics in Hippocampal Networks. *Neuron*, *96*(2), 505-520.e7. https://doi.org/10.1016/j.neuron.2017.09.033

Fetz, E., Toyama, K., & Smith, W. (1991). *Synaptic Interactions between Cortical Neurons* (pp. 1–47). Springer, Boston, MA. https://doi.org/10.1007/978-1-4615-6622-9\_1

Fujisawa, S., Amarasingham, A., Harrison, M. T., & Buzsáki, G. (2008). Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nature Neuroscience*, *11*(7), 823–833. https://doi.org/10.1038/nn.2134

Ghanbari, A., Malyshev, A., Volgushev, M., & Stevenson, I. H. (2017). Estimating short-term synaptic plasticity from pre- and postsynaptic spiking. *PLOS Computational Biology*, *13*(9), e1005738. https://doi.org/10.1371/journal.pcbi.1005738

Ghanbari, A., Ren, N., Keine, C., Stoelzel, C., Englitz, B., Swadlow, H. A., & Stevenson, I. H. (2020). Modeling the Short-Term Dynamics of in Vivo Excitatory Spike Transmission. *Journal of Neuroscience*, *40*(21), 4185–4202. https://doi.org/10.1523/JNEUROSCI.1482-19.2020

Graupner, M., & Brunel, N. (2012). Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proceedings of the National Academy of Sciences*, *109*(10), 3991–3996. https://doi.org/10.1073/pnas.1109359109

Kobayashi, R., Kurita, S., Kurth, A., Kitano, K., Mizuseki, K., Diesmann, M., Richmond, B. J., & Shinomoto, S. (2019). Reconstructing neuronal circuitry from parallel spike trains. *Nature Communications*, *10*(1), 4468. https://doi.org/10.1038/s41467-019-12225-2

Linderman, S., Stock, C. H., & Adams, R. P. (2014). A framework for studying synaptic plasticity with neural spike train data. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 27, pp. 2330–2338). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2014/file/4122cb13c7a474c1976c9706ae36521d-Paper.pdf

Mantel, G. W. H., & Lemon, R. N. (1987). Cross-correlation reveals facilitation of single motor units in thenar muscles by single corticospinal neurones in the conscious monkey. *Neuroscience Letters*, *77*(1), 113–118. https://doi.org/10.1016/0304-3940(87)90617-3

Paninski, L., Ahmadian, Y., Ferreira, D. G., Koyama, S., Rahnama Rad, K., Vidne, M., Vogelstein, J., & Wu, W. (2010). A new look at state-space models for neural data. In *Journal of Computational Neuroscience* (Vol. 29, Issues 1–2, pp. 107–126). Springer. https://doi.org/10.1007/s10827-009-0179-x

Perkel, D. H., Gerstein, G. L., & Moore, G. P. (1967). Neuronal Spike Trains and Stochastic Point Processes: II. Simultaneous Spike Trains. *Biophysical Journal*, *7*(4), 419–440. https://doi.org/10.1016/S0006-3495(67)86597-4

Rauch, H. E., Tung, F., & Striebel, C. T. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, *3*(8), 1445–1450. https://doi.org/10.2514/3.3166

Ren, N., Ito, S., Hafizi, H., Beggs, J. M., & Stevenson, I. H. (2020). Model-based detection of putative synaptic connections from spike recordings with latency and type constraints. *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00066.2020

Smith, A. C., & Brown, E. N. (2003). Estimating a state-space model from point process observations. *Neural Computation*, *15*(5), 965–991. https://doi.org/10.1162/089976603765202622

Stevenson, I. H. (2016). Flexible models for spike count data with both over- and under- dispersion. *Journal of Computational Neuroscience*, *41*(1), 29–43. https://doi.org/10.1007/s10827-016-0603-y

Stevenson, I. H. (2018). Omitted Variable Bias in GLMs of Neural Spiking Activity. *Neural Computation*, *30*(12), 3227–3258. https://doi.org/10.1162/neco\_a\_01138

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

Swadlow, H. A., & Gusev, A. G. (2001). The impact of “bursting” thalamic impulses at a neocortical synapse. *Nature Neuroscience*, *4*(4), 402–408. https://doi.org/10.1038/86054

Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., & Brown, E. N. (2005). A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects. *Journal of Neurophysiology*, *93*(2), 1074–1089.

Usrey, W. M., Alonso, J. M., & Reid, R. C. (2000). Synaptic interactions between thalamic inputs to simple cells in cat visual cortex. *Journal of Neuroscience*, *20*(14), 5461–5467. https://doi.org/10.1523/jneurosci.20-14-05461.2000

Zucker, R. S., & Regehr, W. G. (2002). Short-Term Synaptic Plasticity. *Annual Review of Physiology*, *64*(1), 355–405. https://doi.org/10.1146/annurev.physiol.64.092501.114547