COMMUN. MATH. SCI.
Vol. 17, No. 5, pp. 1177–1192

ANALYSIS OF SPIKE-DRIVEN PROCESSES THROUGH ATTRIBUTABLE COMPONENTS*

HORACIO G. ROTSTEIN† AND ESTEBAN G. TABAK‡

In memory of Professor David Shenou Cai

Abstract. Postsynaptic neuron activity at both the sub and suprathreshold level is analyzed through the combination of: (1) the numerical simulation of a simple leaky integrate-and-fire model forced by both constant frequency and Poisson-distributed presynaptic spike-trains,(2) the transformation of the model's response into sequences describing non-summation effects in subthreshold and the probability of spiking within a time-window in suprathreshold dynamics, (3) for constant frequency input, the analysis of these sequences through an autoregressive linear model, and (4) for non-uniform input, their analysis through attributable components. It is found that the attributable component methodology can reproduce the dynamics on testing data, effectively replacing the original dynamical model, and that the optimal order of both the autoregressive and the attributable component model, is an indicator of the relative strength of the underlying depression and facilitation mechanisms.

Keywords. time series; history-dependent processes; synaptic short-term plasticity; dimensional reduction; attributable components.

AMS subject classifications. 62-07; 62G08; 92C20; 37N25.

1. Introduction

Many naturally occurring dynamical processes are driven not by a continuous external forcing but by discrete, punctuated events. The prototypical example arises in neuronal dynamics, where each neuron receives as input not a continuous current but spike-trains from other neurons through synaptic connections. Analyzing such processes typically involves not only the development of conceptual models that explain their behavior, but also of models that are exclusively data-driven, seeking to capture the dynamics underlying a set of observations so as to be able to reproduce the dynamics and make predictions accordingly. Ideally, such data-driven models can be combined with field expertise to uncover mechanisms underlying the observed dynamics.

The purpose of this article is to explore the construction of data-driven models of spike-driven processes via attributable component analysis, a recently developed methodology for the explanation of variability in data [1]. This will be developed in the particular context of the postsynaptic activity of individual neurons. It will be shown that the attributable component methodology provides an effective non-parametric tool for reproducing both the sub and suprathreshold dynamic response of a cell to excitatory input spikes from another cells, and that the optimal number of spikes in the immediate past used by the model provides an indicator of the relative levels of facilitation and depression mechanisms operating at the synaptic level.

The dynamic behavior of neuronal circuits results from the cooperative activity of the participating neurons and the synaptic network connectivity [2–4]. These complex processes involve the intrinsic properties of the individual neurons (ionic currents, nonlinearities, time scales), the type of neurotransmitter involved (AMPA excitation,

*Received: May 30, 2018; Accepted (in revised form): May 4, 2019.

†Federated Department of Biological Sciences, New Jersey Institute of Technology and Rutgers University, Newark, NJ, USA and Institute for Brain and Neuroscience Research, New Jersey Institute of Technology, Newark, NJ, USA (horacio@njit.edu). https://biology.njit.edu/faculty/horacio

‡Courant Institute of Mathematical Sciences, New York University, New York, NY, USA (tabak@cims.nyu.edu). https://math.nyu.edu/faculty/tabak

$GABA_A$ inhibition) and the so-called synaptic plasticity, which describes the changes in the efficacy of a synaptic connection over time [5–7].

Synaptic connections are highly complex processes. Schematically, they consist of a presynaptic neuron (the "sender"), a postsynaptic neuron (the "receiver") and the proper synapse (the "directed arrow" from the sender to the receiver). Each of these components possesses intrinsic dynamics with varying degrees of complexity. A crucial step in understanding the dynamics of synaptic connections is to determine how the spiking patterns of the presynaptic neurons (presynaptic spikes) shape the response of the postsynaptic neurons both at the subthreshold (membrane potential) and suprathreshold (spike) levels. Computationally, this could be pursued from different perspectives. The classical modeling approach focuses on determining the properties of the postsynaptic patterns given a presynaptic spike-train and the properties of the synaptic connection. The spike decoding approach focuses on reconstructing the presynaptic spike train (e.g., stimulus) from the postsynaptic pattern (e.g., sensory neuron) [8–10]. A third approach focuses on determining the properties of the synaptic dynamics given the postsynaptic pattern and the presynaptic spike-train [11–15]. This has been termed synaptic decoding. Naturally, all cases require knowledge of the intrinsic properties of the postsynaptic cell.

It is particularly important to understand how the synaptic short-term plasticity (STP) contributes to this process. STP refers to the changes experienced by the synaptic efficacy over time that reflect the history of the presynaptic activity [5,6,16,17]. In terms of biophysical models, the synaptic connectivity is captured by the so-called synaptic function S. In the absence of any synaptic dynamics the magnitude S_{max} of S is constant over time (see Figures 3.1-A1 and -A2, red). In the presence of solely synaptic depression S_{max} decreases over time (Figures 3.1-B1 and -B2, red), while purely synaptic facilitation has the opposite effect. The combination of both may produce more complex patterns (Figures 3.1-C1 and -C2, red).

Several models have been used to investigate the mechanisms underlying synaptic transmission and short term dynamics [18], ranging from detailed models [19] to phenomenological models of different types [12, 20–24]. These models are physiologically interpretable based on the assumptions on which they rely. On the other extreme, nonparametric models (lacking any assumption about the underlying process) have been used to capture the input-output relationships associated to STP with minimal, rather generic assumptions [14] about the memory process.

In this paper we combine attributable components analysis –described in the section on methods below– with simple dynamical systems tools to understand the dynamic structure embedded in synaptic decoding in the presence of STP (synaptic depression, facilitation, or both). We focus on the relatively simple architecture consisting of a passive postsynaptic neuron receiving a presynaptic spike-train input. Each presynaptic spike generates a prototypical membrane potential response (Figures 3.1-A1, blue) whose height is assumed to vary due to the effects of synaptic summation (Figure 3.1-A2, blue), depression and facilitation (Figures 3.1-B and -C, blue). The difference A_j (indexed over the input spike times t_j) between the heights of the postsynaptic responses to a given input spike in the presence and absence of STP (once the summation effects have been removed) generates a finite sequence [12, 13]. Each point in the domain of this sequence (except for $j = 1$) has an associated value for the input inter-spike interval (ISI). In order to analyze the spiking response to the presynaptic input spikes we use a spiking probability metric adapted from [11] that assigns to each input spike a spiking probability P_j that a spike occurs within some bin after the jth input spike.

Using simple passive postsynaptic cell and STP models we show that the evolution of $\{A_j\}$, viewed as a discrete dynamical system evolving over the index set j, can be captured by a linear autoregressive map with ISI-dependent parameters when the input frequency is constant. These parameters can be estimated from the synthetic data generated by the model. For either pure depression or facilitation the sequence $\{A_j\}$ is either decreasing or increasing, and the map that generates this sequence is one-dimensional, corresponding to a first-order model. When both depression and facilitation are present, the sequence of $\{A_j\}$ can overshoot, and the corresponding map can be two- or three-dimensional – of second or third order – depending on the relative levels between the two processes. This higher-dimensionality reflects the interaction between the two history-dependent processes. The differences in dimensionality obtained by using two different parameter sets reflect differences in the relative levels of depression and facilitation. Interestingly, this shows that in order to produce the present value of A_j the system needs information up to three steps back depending on the STP scenario. For constant input frequencies, the evolution of P_j shows qualitatively similar behavior to that of A_j.

We then investigate the properties of the subthreshold and spiking responses for the more realistic Poisson-distributed input spike-trains. Here the sequence to analyze involves not only the $\{A_j\}$ or $\{P_j\}$, but also the spiking times $\{t_j\}$ of the presynaptic neuron. We develop a procedure whereby, given the input sequence $\{t_j\}$, we first generate the response of the postsynaptic neuron through the numerical solution of a current-balance equation model, and decode it in terms of $\{A_j\}$ or $\{P_j\}$, depending on whether we are analyzing sub or suprathreshold responses. We use attributable component analysis to build a non-linear autoregressive model that computes A_j (or P_j) in terms of n_s prior values of A_j and of the corresponding inter-spike intervals $\Delta_j = t_j - t_{j-1}$.

This article is structured as follows: after this introduction, Section 2 (Subsection 2.1) describes the leaky integrate-and-fire model used for the postsynaptic cell, the stochastic model for the presynaptic spike-trains, and the short-term dynamics of synaptic depression and facilitation. Subsection 2.2 summarizes the attributable component methodology developed in [1]. Section 3 describes the results obtained for both history-dependent membrane potential response patterns and spiking response patterns. In both cases, we first analyze the response to uniform spike-trains of constant frequency, showing how the combined effects of facilitation and depression give rise to distinct transient patterns and to autoregressive dynamics of different orders for the reduced variables A and P. Then we consider stochastic spike-trains drawn from Poisson distributions, and show how the attributable component methodology can robustly reproduce the response of the postsynaptic neuron, and how the optimal number of steps in the past that the model considers relates to the relative amounts of facilitation and depression. Finally, we summarize the results in Section 4, where we also discuss directions for further development.

2. Methods

2.1. Models.

2.1.1. Postsynaptic cell: leaky integrate-and-fire model. The current-balance equation for the postsynaptic cell is given by

$$C\frac{dV}{dt} = -g_L(V - E_L) + I_{app} - I_{syn} + I_{noise}, \tag{2.1}$$

where t is time (ms), V represents the voltage (mV), C the specific capacitance (μF/cm^2), g_L the leak conductance (mS/cm^2), I_{app} the tonic (DC) current (μA/cm^2), $I_{noise} = \sqrt{2D}\eta(t)$ represents white noise (delta correlated with zero mean), and I_{syn} is an excitatory synaptic current of the form

$$I_{syn} = G_{ex}\, S\,(V - E_{ex}). \tag{2.2}$$

Here G_{ex} is the maximal synaptic conductance (mS/cm^2), $E_{ex} = 0$ is the reversal potential for AMPA excitation, and the synaptic variable (S) obeys a kinetic equation of the form

$$\frac{dS}{dt} = -\frac{S}{\tau_{dec}}, \tag{2.3}$$

where $\tau_{dec} = 3$ (ms) is the decay time of excitation. Each presynaptic spike instantaneously raises S to some value S_{max} which varies depending on the properties of the short-term dynamics (depression and/or facilitation) and is given by $S_{max} = ru$ defined below. We refer the reader to [2, 3] for additional details on biophysical (conductance-based) models.

2.1.2. Presynaptic spike-trains. We model the spiking activity of the presynaptic cell as a spike train by providing the presynaptic spike times $t_1, t_2, \ldots, t_N$. We consider two types of input spike-trains: uniformly- and Poisson-distributed. The uniform spike-trains are characterized by their interspike interval (ISI) or its reciprocal, the spiking frequency. Poisson distributed inputs are characterized by their mean spiking frequency or the associated exponential distribution of ISIs.

2.1.3. Short-term dynamics: synaptic depression and facilitation. We use a phenomenological model introduced in [22]. Short-term depression and facilitation are described by two independent variables: r and u that obey kinetic equations of the form

$$\frac{dr}{dt} = \frac{1-r}{\tau_{dep}} \qquad \text{and} \qquad \frac{du}{dt} = \frac{U-u}{\tau_{fac}}, \tag{2.4}$$

respectively, where τ_{dep} and τ_{fac} are characteristic time constants and $U = 0.1$. Each presynaptic spike instantaneously decreases r by an amount equal to the current value of the product ru and increases u by an amount equal to $U(1-u)$. These variables enter into the current-balance equation through S_{max}, which is given by the product ru. In the absence of depression, $r = const$ $(= 1)$, and in the absence of facilitation, $u = U = 0.1$.

2.2. Attributable components. Attributable components is a recently developed methodology [1] for estimating the conditional mean $\bar{x}(z)$ of a variable x dependent on covariates $z = (z_1, \ldots, z_L)$. Although the x and each z_l can be variables of quite general type, we will limit the description here to the particular case in which they are all real and scalar, as this will be the situation in the application to history dependent processes under consideration.

Given n data pairs of the form $\{x^i, z^i\}$, one seeks the conditional expectation $\bar{x}(z)$, which may be characterized as the minimizer of the empirical variance

$$\bar{x}(z) = \operatorname*{argmin}_f \sum_{i=1}^{n} \left\| x^i - f\left(z^i\right) \right\|^2, \tag{2.5}$$

as in least-square regression. The challenge is how to describe the multivariate function $f(z)$ in a way that is at the same time general (minimally parametric), robust (not prone to overfitting), and efficiently computable. The proposal in [1] is to perform the equivalent to a low-rank tensor factorization

$$f(z) = \sum_{k=1}^{d} \prod_{l=1}^{L} f(l)^k(z_l) \tag{2.6}$$

for variables z_l that are not necessarily discrete, as are the rows and columns in conventional matrix factorization. When the z_l are real variables, (2.6) is analogous to a truncated sum of the separated-variable solutions used for instance in linear partial differential equations.

In order to characterize each function $f(l)^k(z_l)$ in (2.6), we introduce a grid $\left\{z_{gl}{}^j\right\}$, not necessarily uniform, and write each observed value z_l^i in the form

$$z_l^i = \sum_j \alpha(l)_i^j z_{gl}{}^j, \quad \alpha(l)_i^j \geq 0, \quad \sum_j \alpha(l)_i^j = 1, \tag{2.7}$$

where for each i only two $\alpha(l)_i^j$ are nonzero: those corresponding to the two grid points $z_{gl}{}^j$ immediately surrounding z_l^i. Then we model $f(l)^k(z_l)$ through piecewise linear interpolation on the given grid:

$$f(l)^k\left(z_l^i\right) = \sum_j \alpha(l)_i^j V(l)_j^k, \quad V(l)_j^k = f(l)^k\left(z_{gl}{}^j\right). \tag{2.8}$$

Thus the parameters that characterize $f(z)$ are the L matrices $V(l)$. Yet one should not let these be completely free in the minimization of (2.5), as this could lead to overfitting, especially when the grids z_g chosen are very fine with respect to the number n of sample points available. Instead, the algorithm proposed in [1] penalizes the non-smoothness of $f(l)^k(z_l)$ through an extra term added to the combination of (2.5), (2.6) and (2.8):

$$\min_V \sum_i \left(x^i - \sum_k \prod_{l\in L} \sum_j \alpha(l)_i^j V(l)_j^k \right)^2 + \lambda \sum_{l=1}^{L} \sum_k \left(\prod_{b\in L, b\neq l} \|V(b)^k\|^2 \right) V(l)^{k\prime} C^l V(l)^k, \tag{2.9}$$

where the matrices C^l model the square norm of finite difference approximations to combinations of derivatives of $f(l)^k(z)$, typically the first and second derivative.

In our application, the variable x^i to explain is given by either A^i or P^i, depending on whether we study subthreshold (membrane potential) response or spiking response patterns. The covariates z_l^i can be divided into two groups: previous values A^{i-k} or P^{i-k} with k ranging from (1) to n_s, and the corresponding time intervals Δ^{i+1-k}, with $\Delta^i = t^i - t^{i-1}$. Here n_s is a modeling parameter measuring the order of the discrete dynamical model, i.e. how far back in time it looks in order to make predictions. During training, all x^i and z^i are drawn from the data. During testing, only the time intervals Δ^i are given, while the x^i are inferred from the model, using, for the missing components of the z^i, the corresponding values of x^{i-k} previously inferred.

3. Results

Below we analyze the history-dependent neuronal response patterns to (AMPA) excitatory presynaptic spike-trains. We consider two cases: (i) postsynaptic membrane potential response patterns in the absence of spikes, and (ii) postsynaptic spiking response patterns. The first case includes subthreshold responses and more complex

types of responses where spikes are either ignored or filtered, for example when one considers the voltage envelope of bursting patterns. Because we are focusing on the history-dependent effects of the presynaptic activity, we use the relatively simple leaky integrate-and-fire model for the postsynaptic cell to avoid interferences with postsynaptic ionic currents (other than the leak current).

3.1. History-dependent membrane potential response patterns.

3.1.1. Uniformly distributed presynaptic spike-trains. Figure 3.1 shows representative membrane potential (V) response patterns to (AMPA) excitatory presynaptic spike-trains with constant frequency generated using the model described in Section 2.1. Each presynaptic spike generates a signal represented by the synaptic function S (red), which is the input to the current-balance Equation (2.1) in the postsynaptic cell through the synaptic current (2.2). This signal interacts with the neuronal intrinsic properties and produces a stereotypical voltage output V (blue). In the absence of any other factors V has the prototypical shape (alpha function or difference of exponentials) shown in (Figure 3.1-A1, blue). It is qualitatively similar to S, but V evolves on a slower time scale than S. As the input frequency increases, the voltage oscillation amplitude decreases, while the oscillation envelope may show different types of behavior depending on the synaptic properties.

For low enough frequencies, V decreases to zero before the next spike arrives. For higher frequencies, the so-called summation phenomenon amplifies the signal, in the sense that the oscillation peaks are larger the higher the input frequency, although the oscillation amplitude decreases with increasing input frequency (compare panels A1 and A2). Synaptic depression causes the amplitude of the synaptic function S to decrease with increasing number of spike times (Figure 3.1-B, red). For low enough input frequencies, the voltage response exhibits the same phenomenon (Figure 3.1-B1), while for higher input frequencies the interplay of summation and depression generates a peak in the voltage response (Figure 3.1-B2). The interplay of depression and facilitation generates a peak in the response of both S and V (Figure 3.1-C) under appropriate conditions (e.g., input frequency, τ_{dep}, τ_{fac}). The peak in S is the result of the interplay of depression and facilitation at the presynaptic level (the envelope of S remains at baseline). The peak in V depends on the additional interaction with the filtering properties of the postsynaptic cell. For low enough frequencies, the voltage response roughly mimics the S pattern (Figure 3.1-C1), while for higher frequencies summation generates an additional amplification of the voltage response (Figure 3.1-C2). We note that the dependence of the S and V patterns on the input frequency and the synaptic depression and facilitation time constants is highly complex, a full analysis is beyond the scope of this article.

Typically, the presynaptic spike times are not uniform (e.g., the ISIs are Poisson-distributed) and one has access to the voltage response V to a presynaptic spike train, but not to the synaptic functions S that are hidden. This together with the summation effects make the synaptic decoding, particularly the determination of the presence of STP, not straightforward. One approach that has been used in the literature is to approximate the voltage response by kernels that themselves approximate the stereotypical voltage responses to a single (isolated) input spike as in Figure 3.1-A1 (blue) [12]. Specifically, if the presynaptic spike times are $t_1, t_2, \ldots t_N$, then

$$V_{est}(t) = \sum_{j=1}^{N} K(t - t_j)[1 + A_j] \tag{3.1}$$

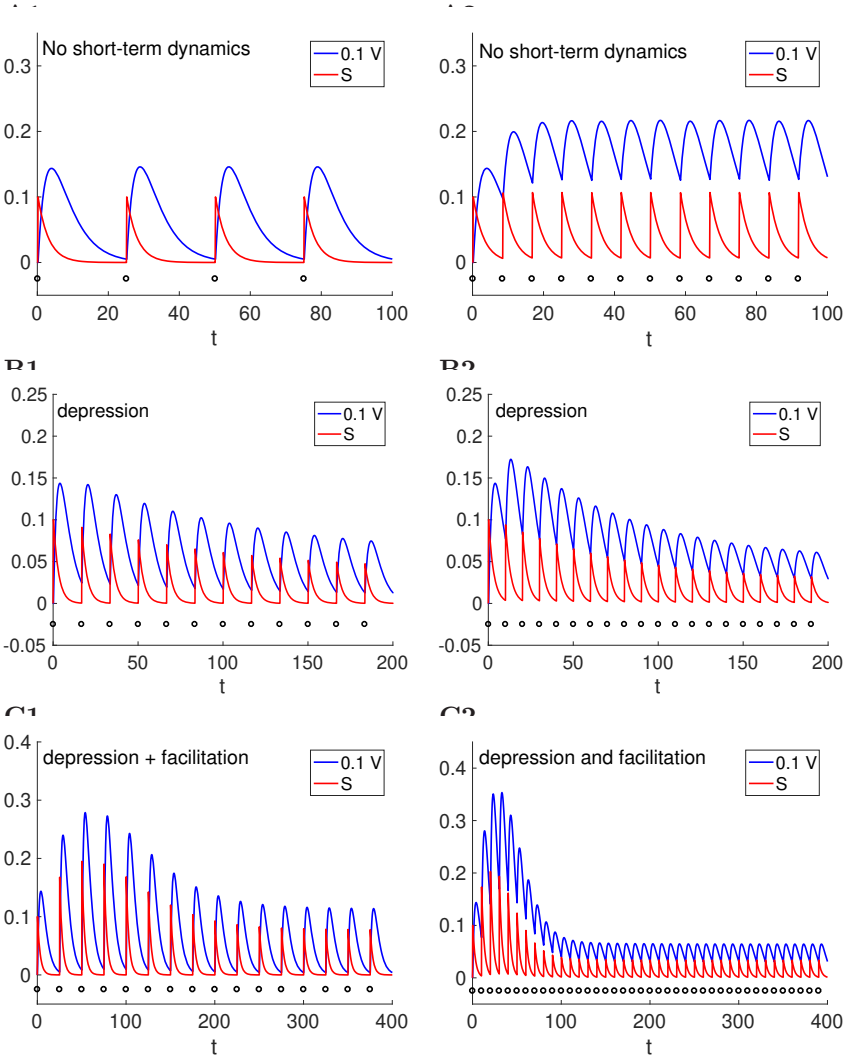

FIG. 3.1. **Subthreshold (membrane potential) neuronal response to spike train inputs with synaptic excitation (AMPA): representative examples for different input frequencies and short-term dynamic properties.** *Black dots indicate the arrival of a presynaptic spike. The postsynaptic voltage response (V) is scaled and raised (by -52.5 mV) for comparison purposes with the synaptic function S.* **A.** *No short-term dynamics.* **B.** *Synaptic depression.* **C.** *Synaptic depression and facilitation. We used the following parameter values:* $C=1$, $E_L=-60$, $G_L=0.2$, $I_{app}=1.5$, $G_{ex}=0.2$, $E_{ex}=0$, $\tau_{dep}=300$, $\tau_{fac}=400$, $U=0.1$.

where the kernel $K(t)$ can be extracted from the voltage response to a single (isolated) input spike, or by adopting a prototypical shape for it –alpha functions or the difference of exponentials– and estimating its parameters from the data. The components of the vector $A=\{A_j\}_{j=1}^N$ measure the size of the response to the input spike at time t_j as compared to the baseline size ($A=0$), independently of summation, and are history-dependent.

We refer to A as a sequence and we view it as a discrete dynamical system evolving along the $t_{spk,in}=\{t_j\}_{j=1}^N$ domain [25]. We illustrate this is Figure 3.2. In the presence

of synaptic depression only (blue curve), A is a decreasing function of $t_{spk,in}$, which eventually saturates. In the presence of both synaptic depression and facilitation (red), A first increases, then decreases and finally saturates. In the presence of facilitation only (not shown) A is an increasing function of $t_{spk,in}$, which eventually also saturates (qualitatively, a mirror image of the blue curve). The transition among the different types of A patterns depends on the relative levels of depression and facilitation, which can be measured in terms of τ_{dep} and τ_{fac}. The details are being studied in [25].

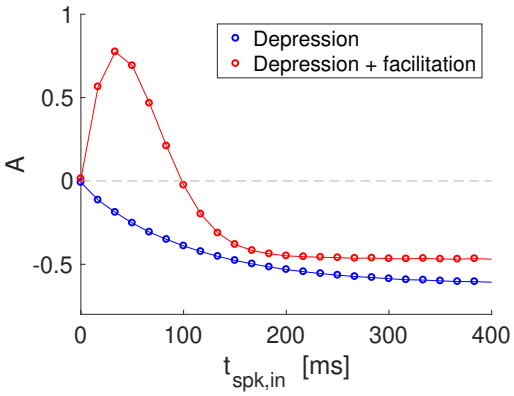

FIG. 3.2. **Components of the vector $A=\{A_j\}_{j=1}^{N}$ in Equation (3.1).** *We used the following parameter values:* $C=1$, $E_L=-60$, $G_L=0.1$, $I_{app}=0$, $G_{ex}=0.5$, $E_{ex}=0$, $\tau_{dep}=300$, $\tau_{fac}=400$, $U=0.1$.

The shapes of A (see Figure 3.2) suggest they can be generated by discrete low-dimensional linear autoregressive maps of the form

$$\hat{A}_k=\sum_{j=1}^{M}\alpha_{k-j}\,\hat{A}_{k-j} \tag{3.2}$$

where $M<N$ and the coefficients α_{k-j} $(j=1,\ldots,M)$ are input-frequency dependent. Because the map requires M initial conditions, M should be small for the map to be of any use. Processes that involve only depression or facilitation can be captured by 1D linear maps ($M=1$), while processes that involve both require up to $M=3$ [25]. We refer to them as MD maps. We illustrate this in Figure 3.3 where we show the error between A and the linear map approximations $\hat{A}$

$$Error=\frac{1}{N}\sum_{K=1}^{N}|A_k-\hat{A}_k| \tag{3.3}$$

as a function of τ_{fac} for fixed values of τ_{dep} using $M=1$ (blue), $M=2$ (red) and $M=3$ (green). The parameter N represents the number of input spike times within an interval that captures the transient behavior of A.

For low enough values of τ_{fac}, a 1D map is enough to capture the evolution of A, consistently with the fact that depression dominates. As τ_{fac} increases, the 1D map (blue) does no longer provide a good approximation, since it cannot capture the non-monotonic behavior of A. The 2D map (red) still provides a good approximation for a small range of intermediate values of τ_{fac}. For higher values of τ_{fac}, only the 3D map (green) provides a good approximation. As expected, the error varies with τ_{dep} (compare

panels A and B), indicating that the relative value of the time constants between the two processess matter.

We emphasize that these map approximations are only valid for uniform distributions of input spikes, and that they depend on the input frequency. Yet they identify the interplay of depression and facilitation, on the basis of the order of the model, that is the number of steps back required to capture the behavior of A. The fact that we obtained qualitatively similar results for different input frequencies suggests that this memory effect (maximum value of M) may persist for non-uniform distribution of presynaptic spike-trains, and that the values of M would identify the different levels of complexity resulting from different ratios τ_{fac}/τ_{dep}.

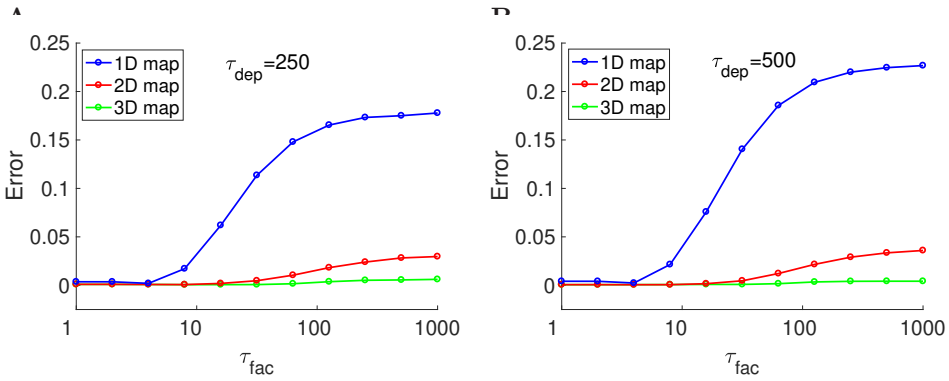

FIG. 3.3. **Approximation error (3.3) between the linear maps and the sequence A for representative parameter values.** *The input frequency (uniform distribution of input spikes) is 60 Hz. We used the following parameter values:* $C=1$, $E_L=-60$, $G_L=0.1$, $I_{app}=0$, $G_{ex}=0.5$, $E_{ex}=0$, $U=0.1$, $M=31$ ($t_{spk}\leq 500$. **A.** $\tau_{dep}=250$. **B.** $\tau_{dep}=500$.

3.1.2. Poisson distributed input spike-trains. The analysis above was restricted to uniformly timed spike-trains. In order to analyze more general input data, we follow the following steps:

(1) We draw two input sequences of spikes $\{t_j\}$ from a Poisson distribution with specified mean frequency, one for training and one for testing.

(2) We model the response $V(t)$ to this input via the model in (2.1) as before, with specified values for τ_{fac} and τ_{dep}.

(3) As in the prior subsection, we post-analyze this response, transforming it into a sequence $\{A_j\}$.

(4) We apply attributable-component analysis to the training sequence $\{t_j, A_j\}$, seeking to determine the functional dependence of A_j on $(A_{j-1},\ldots,A_{j-n_s})$ and $(\Delta_{j-1},\ldots,\Delta_{j-n_s})$, where $\Delta_j = t_j - t_{j-1}$ represents the time-interval between two consecutive input spikes, and n_s represents the number of prior input pairs that the analysis will consider. Thus we seek

$$\bar{A}_j = F(A_{j-1},\Delta_{j-1},\ldots,A_{j-n_s},\Delta_{j-n_s}) \tag{3.4}$$

in the functional form given by (2.9).

(5) We apply the model (3.4), with parameters determined using the training sequence, to the testing sequence, attempting to reconstruct the $\{A_j\}$ from the $\{t_j\}$ alone. Denoting our sought reconstruction $\left\{\tilde{A}_j\right\}$, we set $\left\{\tilde{A}_1,\ldots,\tilde{A}_{n_s}\right\}=0$ – a natural

default, since we cannot determine the first n_s elements of the sequence through a model that requires the prior n_s values –, and then compute for all subsequent values of j

$$\tilde{A}_j = F\left(\tilde{A}_{j-1}, \Delta_{j-1}, \ldots, \tilde{A}_{j-n_s}, \Delta_{j-n_s}\right).$$

Thus we are effectively replacing the dynamical model (2.1) with one derived solely from data, whose only input – additional to the training data sequence itself and the penalization parameter λ – is n_s, indicating the extent of the system's memory.

(6) We measure the accuracy of the reconstruction via the variance reduction it achieves, defined as

$$VR = \frac{\frac{1}{n-n_s}\sum_{j=n_s+1}^{n}\left(A_j - \tilde{A}_j\right)^2}{\frac{1}{n}\sum_{j=1}^{n}\left(A_j - \bar{A}\right)^2}, \quad \bar{A} = \frac{1}{n}\sum_{j=1}^{n} A_j. \tag{3.5}$$

For the reconstruction to work, it is not only necessary that the model (3.4) be accurate and robust, but also that the dynamics of both the biophysical model (2.1) and the data-driven model (3.4) be stable, so that discrepancies in the initialization and in realizations of the noise and local errors do not propagate downstream the time-series.

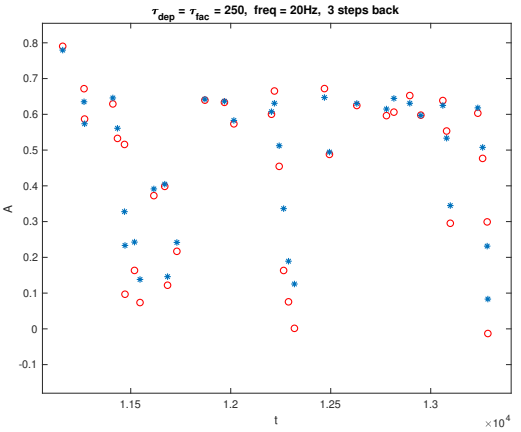

FIG. 3.4. *Out of sample observations (in blue stars) and predictions (in red circles) for A_j, for an input mean frequency of $20Hz$, equal values $\tau = 250$ for the time-scales for facilitation and depression, and $n_s = 3$. The dynamics resulting from the biophysical and the data-driven model are virtually indistinguishable.*

Figure (3.4) displays the reconstructed $\left\{\tilde{A}_j\right\}$ (in red circles) and the true $\{A_j\}$ (in blue stars) as a function of the spiking times $\{t_j\}$ for one instance of the parameters, for a subset of the testing set small enough that one can distinguish visually the individual predictions. As one can see, the procedure is accurate and robust, and one can indeed replace (2.1) by the data driven model (3.4) as an effective dynamical model.

Does the data-driven model inform us on the biophysical mechanisms underlying the data? Table (3.1) displays the variance reduction VR from (3.5) for a range of combinations of the biophysical parameters τ_{dep} and τ_{fac} and values of n_s ranging between 1 and 4. As one can see, while for no facilitation one or two steps back suffice

τ_{dep}	τ_{fac}	freq.	1 step	2 steps	3 steps	4 steps
250		10	0.0815	0.0631	0.0730	0.1409
500		10	0.0746	0.0440	0.0494	0.0689
250		20	0.0896	0.0766	0.0816	0.1334
500		20	0.0799	0.0788	0.1087	0.2575
250		30	0.0993	0.1095	0.1001	0.6897
500		30	0.1038	0.1399	0.2991	0.8321
250	5	10	0.4476	0.3812	0.4915	0.4029
250	10	10	0.6767	0.6598	0.3775	0.3826
250	100	10	0.2793	0.1918	0.1568	0.1232
250	250	10	1.1011	0.2572	0.1158	0.3712
250	5	20	0.5626	0.2560	0.2584	0.2389
250	10	20	0.7462	0.2633	0.2373	0.2590
250	100	20	1.1920	0.1394	0.0654	0.0742
250	250	20	0.1225	0.0683	0.0389	0.0521
250	5	30	0.7190	0.2496	0.2040	0.1860
250	10	30	0.8236	0.2558	0.1846	0.1712
250	100	30	1.2266	0.1205	0.0925	0.1018
250	250	30	0.0997	0.0463	0.0347	0.0372

TABLE 3.1. *Variance reduction through factor removal for subthreshold dynamics. As the timescales τ for depression and facilitation become comparable, the optimal number of prior steps to include increases from 1-2 to 3-4.*

for capturing the dynamics, for values of τ_{fac} comparable to τ_{dep} three or four steps back are required to maximize out-of-sample variance reduction. Hence, just from looking at the optimal n_s, one could infer at least qualitatively the amount of facilitation present in the process. Figure (3.5) corresponds to a case with no facilitation where just one step back suffices. Since in this case z is only two-dimensional, consisting of the previous value of A and the last time-interval, one can display the prediction $\tilde{A}_j(A_{j-1},\Delta_{j-1})$, as well as its individual components, sorted by the amount of variability that they explain. Even though 5 components were used for this simulation, the figure shows how the first component captures most of the behavior of the function, and the second provides a smaller, yet meaningful correction.

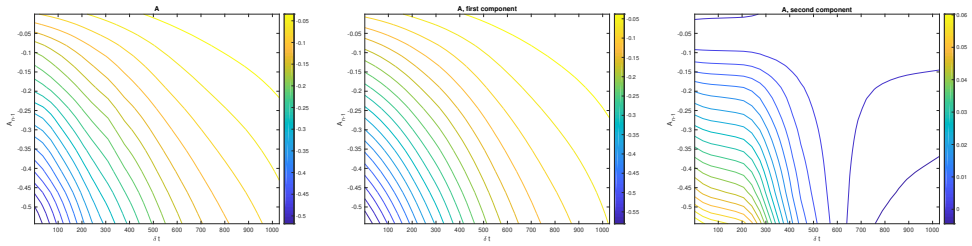

FIG. 3.5. *The predicted function $\tilde{A}_j(A_{j-1},\Delta_{j-1})$ for a case without facilitation, and its two first components, sorted by the amount of variability that they explain. Notice how the first component explains much of the variability of A, while the second corrects mainly its behavior for small values of A_{j-1} and Δ_{j-1}.*

3.2. History-dependent spiking response patterns. Here we focus on the postsynaptic spiking response to presynaptic spike trains in the presence of depression and facilitation. We use a metric analogous to the vector A discussed above, assigning to each presynaptic spike t_j the probability $P(t_j)$ of a spike being generated within some time interval $B_j = [t_j, t_j + \Delta t]$. In order to compute the vector P, we carry out N_{trial} simulations using the same presynaptic input with different realization of white noise, we compute the number of times that a postsynaptic spike is produced within B_j, and we average over the N_{trial} trials. We focus on the case where the cell is silent in the absence of presynaptic input.

3.2.1. Uniformly distributed presynaptic spike-trains. Our results for a representative set of parameter values and uniform spiking input frequencies are presented in Figure 3.6. The parameter values in all panels are the same with the exception of the spiking input frequency that increases from A to D.

The input spikes in Figure 3.6-A are well separated and the resulting P-patterns capture the relative strengths of depression and facilitation. For $\tau_{fac} = 10$, depression dominates, while for $\tau_{fac} = 500$, facilitation dominates.

The P-patterns in Figure 3.6-B capture the interaction between depression and facilitation much in the same way as the A-pattern in Figure 3.2 (red). This effect is observed when τ_{fac} is not too small relative to τ_{dep}. For $\tau_{fac} = 100$ and above (red, green and cyan curves), P first increases and then decreases.

For higher spiking input frequencies the effect is felt even for lower values of τ_{fac} (compare Figure 3.6-B and -C), but the "bump" is more compressed in time, as expected from the higher spiking input frequencies. The P-patterns for the different values of τ_{fac} are less well separated than in Figure 3.6-C than in the previous panels. This persists for even higher spiking input frequencies (Figure 3.6-D), but the P-patterns for the different values of τ_{fac} are almost indistinguishable.

As for the A-sequences discussed in the previous section, the P-sequences capture the interplay of depression and facilitation on the basis of the amount of steps back needed to capture the behavior of P. Despite the fact that this is only valid for uniform distributions of input spikes, the similarity of the results discussed here with these of the previous section (for A) and the qualitative similarity of the P distributions for different input frequencies, suggest that relatively low values of M would identify the different levels of complexity resulting from different ratios τ_{fac}/τ_{dep}.

3.2.2. Poisson distributed input spike-trains. When the input spike train is not uniform but follows a Poisson process instead, we follow the same procedure than in the subthreshold case, but with the P_j replacing the A_j, that is:

(1) We draw two input sequences of spikes $\{t_j\}$ from a Poisson distribution with specified mean frequency, one for training and one for testing.

(2) We model the response $V(t)$ to this input via the model in (2.1), but recording only the postsynaptic spike-times. We do this repeatedly for the same input sequences of strikes but different realizations of white noise.

(3) We post-analyze this response, transforming all realizations into a single sequence $\{P_j\}$.

(4) We apply attributable-component analysis to the training sequence $\{t_j, P_j\}$, seeking a relation of the form

$$\bar{P}_j = F(P_{j-1}, \Delta_{j-1}, \ldots, P_{j-n_s}, \Delta_{j-n_s}). \tag{3.6}$$

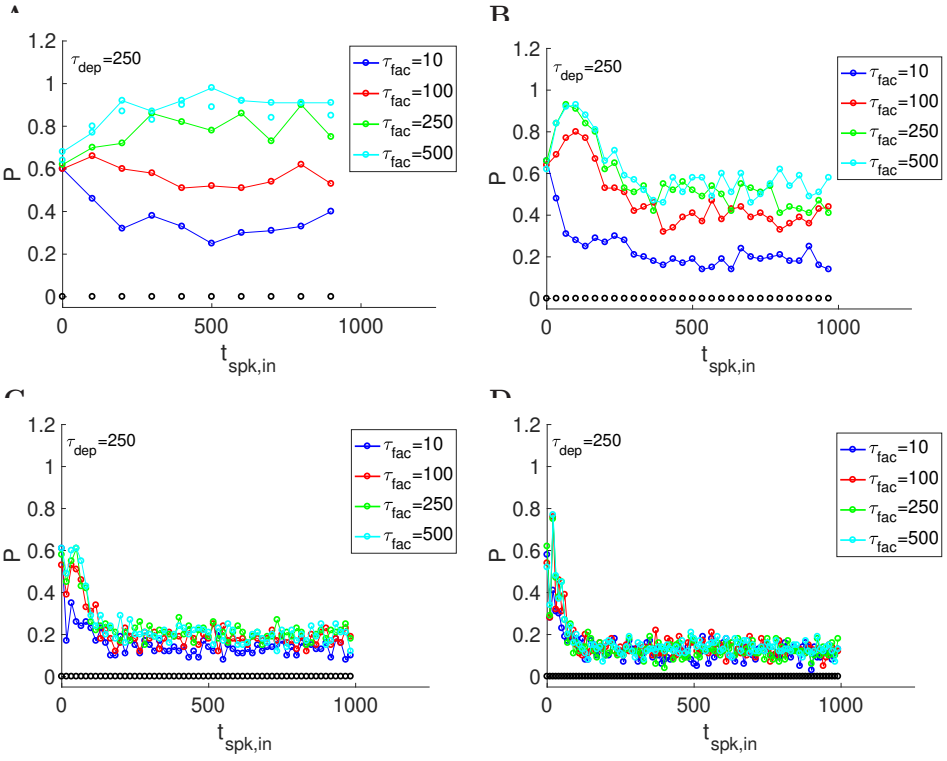

FIG. 3.6. **Spiking probability response to presynaptic input spikes with uniform frequency.** *The input frequency (uniform distribution of input spikes) increases from panel A to D.* **A.** *Input frequency: 10 Hz.* **B.** *Input frequency: 30 Hz.* **C.** *Input frequency: 60 Hz.* **D.** *Input frequency: 100 Hz. We used the following parameter values:* $C=1$, $E_L=-60$, $G_L=0.1$, $I_{app}=0.8$, $D=0.1$, $G_{ex}=0.2$, $E_{ex}=0$, $U=0.1$. **A.** $\tau_{dep}=250$. **B.** $\tau_{dep}=500$.

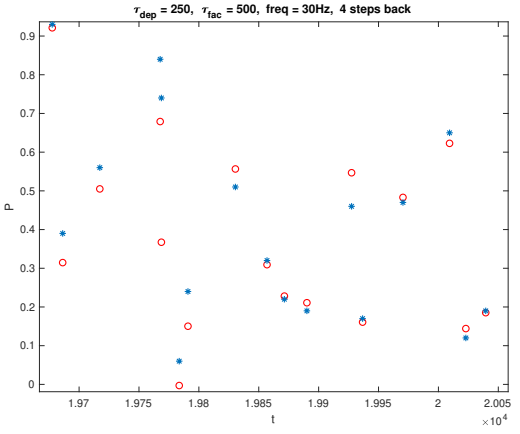

FIG. 3.7. *Out of sample observations (in blue stars) and predictions (in red circles) for the spiking probability* P_j*, for an input mean frequency of* $30Hz$*, facilitation and depression time-scales* $\tau_{fac}=500$ *and* $\tau_{dep}=250$ *, and* $n_s=4$*. Even though the dynamics at suprathreshold are much more noisy than at subthreshold, the data-driven model still mimics the biophysical model very accurately.*

τ_{dep}	τ_{fac}	freq.	1 step	2 steps	3 steps	4 steps
250	10	10	0.9097	0.8685	0.8033	1.3277
250	100	10	0.7812	0.6956	0.7313	0.9046
250	250	10	0.4297	0.2987	0.3267	0.5829
250	500	10	0.3751	0.2693	0.2216	0.6559
250	10	30	0.7466	0.6152	0.6013	0.8475
250	100	30	0.8822	0.6745	0.6708	0.6851
250	250	30	0.2769	0.2210	0.1947	0.1919
250	500	30	0.1703	0.1384	0.1260	0.1290
250	10	60	0.6296	0.5392	0.5303	0.6310
250	100	60	0.7138	0.4429	0.4460	0.4600
250	250	60	0.3407	0.2395	0.2393	0.2368
250	500	60	0.1701	0.1269	0.1257	0.1237

TABLE 3.2. *Out of sample variance reduction through factor removal for spiking dynamics. At high frequencies – less dominated by noise – the optimal number of prior steps to consider increases as the time-scales of depression and facilitation become comparable.*

(5) We apply the model (3.6), with parameters determined using the training sequence, to the testing sequence, so as to reconstruct the $\{P_j\}$ from the $\{t_j\}$ alone.

(6) As before, we measure the accuracy of the reconstruction via the variance reduction it achieves.

Figure (3.7) displays an example of the performance achieved by the algorithm on a testing set, and Table (3.2) shows the variance reduction achieved for various combinations of facilitation and depression, though models looking back (1) to (4) steps. Even though the dynamics is clearly much more noisy than in the subthreshold scenario, particularly for low frequencies, when the input frequencies are higher one can still see the transition in the optimal out-of-sample fit from 2 to 4 steps back as the facilitation and depression time-scales get closer to each other.

4. Discussion

We set out to explore the construction of data-driven models of spike-driven processes via attributable component analysis [1] in the particular context of the postsynaptic response of individual neurons to presynaptic spikes in the presence of STP. Using this minimal model formulation, data generated from this model using subthreshold linear dynamics (only passive subthreshold ionic currents) in the postsynaptic cell, and simple dynamical systems ideas, we provide a proof of concept that the attributable component methodology provides an effective non-parametric tool for reproducing both the sub and suprathreshold dynamic response of a cell to excitatory input spikes from another cells, and that the optimal number of spikes in the immediate past used by the model provides an indicator of the relative levels of facilitation and depression mechanisms operating at the synaptic level.

We followed a complementary dual strategy that allowed us to interpret the results of the non-parametric attributable components analysis for Poisson-distributed input spike trains in terms of discrete linear dynamics applied to uniformly distributed spike trains. Specifically, for both Poisson- and uniformly-distributed spiking inputs we (i) generated artificial data using the model, and (ii) computed the two metrics capturing the subthreshold (A_j) and spiking (P_j) responses to the incoming presynaptic spikes. For Poisson-distributed spiking inputs, we additionally (iii) robustly reproduced these responses using the attributable component methodology. For uniformly-distributed

inputs, on the other hand, we (iii) described the transient patterns for both $\{A_j\}$ and $\{P_j\}$ for various representative values of the input frequency, and (iv) estimated the corresponding linear autoregressive maps.

For Poisson-distributed inputs we showed that the optimal number of steps in the past the model considers is not uniform across data sets, and relates to the relative amounts of depression and facilitation used to generate the data. This is consistent with the order of the linear autoregressive maps computed for uniform distributed inputs for $\{A_j\}$. Because the order of discrete maps is an indication of the interactions between the participating gating variables and these variables directly relate to the two STP processes, we conclude the effective memory of the system for non-uniform input spikes is also reflecting these interactions. We extend this conclusion to the postsynaptic spiking responses using $\{P_j\}$. Since the discrete map approach requires the uniform distribution of input spike trains and cannot be straightforwardly extended to non-uniform distributions, the combination with the attributable component approach becomes synergistic.

This article considered the simplest combination of network architecture and postsynaptic intrinsic properties that allows us to capture the effects of presynaptic input spikes on the postsynaptic patterns in the presence of STP. Future work should consider models that include (i) the presence of active ionic currents, particularly currents that produce spike-frequency adaptations (e.g., slow potassium), whose effect can be similar to synaptic depression, (ii) the effects of synaptic inhibition, particularly in the presence of hyperpolarization-activated currents (e.g., mixed-cation sodium/potassium and T-type calcium), (iii) the presence of multiple depression and facilitation process with different relaxation scales, (iv) other types of "hidden variables" such as GABA_B and plasticity in electrical gap junctions, and (v) the adaptation of the methods presented here to these situations. Even though we have assumed the postsynaptic cell to be silent in the absence of presynaptic spikes, we expect that, with minimal corrections, the results will remain valid in situations where the cells spike autonomously.

The data used in this paper was generated with a biophysical (conductance-based) model, a necessary step to calibrate the methods and understand their potential to explain experimental data both *in vitro* and *in vivo*.

Acknowledgments. This work was partially supported by the National Science Foundation grants DMS-1608077 (HGR) and DMS-1715753 (EGT), by NIH MH060605 and by the Office of Naval Research (EGT). HGR is thankful to the Courant Institute of Mathematical Sciences at NYU. The authors are grateful to Farzan Nadim, Dirk Bucher and Sam McKenzie for useful discussions.

REFERENCES

[1] E.G. Tabak and G. Trigila, *Conditional expectation estimation through attributable components*, Inf. Inference, 7(4):727–754, 2018. 1, 2.2, 2.2, 2.2, 4

[2] P. Dayan and L.F. Abbott, *Theoretical Neuroscience*, The MIT Press, Cambridge, Massachusetts, 2001. 1, 2.1.1

[3] G.B. Ermentrout and D. Terman, *Mathematical Foundations of Neuroscience*, Springer, 2010. 1, 2.1.1

[4] D. Johnston and S.M.-S. Wu, *Foundations of Cellular Neurophysiology*, The MIT Press, Cambridge, Massachusetts, 1995. 1

[5] C. Stevens and Y. Wang, *Facilitation and depression at single central synapses*, Neuron, 14:795–802, 1995. 1

[6] H. Markram and M. Tsodyks, *Redistribution of synaptic efficacy between neocortical pyramidal neurons*, Nature, 382:807–810, 1996. 1

[7] D.O. Hebb, *The Organization of Behavior: A Neuropsychological Theory*, Wiley, New York, 1949. 1

[8] W. Bialek, F. Rieke, R.R. de Ruyter van Steveninck, and D. Warland, *Reading a neural code*, Science, 252:1854–1857, 1991. 1

[9] L.F. Abbott, *Decoding neuronal firing and modeling neural networks*, Q. Rev. Biophys., 27:291–331, 1994. 1

[10] D. Warland, M. Landolfa, J.P. Miller, and W. Bialek, *Reading between the spikes in the cercal filiform hair receptors of the cricket*, in F.H. Eeckman (ed.), Analysis and Modeling of Neural Systems, Springer, 327–333, 1991. 1

[11] D.F. English, S. McKenzie, T. Evans, K. Kim, E. Yoon, and G. Buzsáki, *Pyramidal cell-interneuron circuit architecture and dynamics in hippocampal networks*, Neuron, 96:505–520, 2017. 1

[12] K. Senn, J.C. Jorge-Rivera, E. Marder, and L. F. Abbott, *Decoding synapses*, J. Neurosci., 16:6307–6318, 1996. 1, 3.1.1

[13] E. Stern, K. García-Crescioni, M.W. Miller, C.S. Peskin, and V. Brezina, *A method for decoding the neurophysiological spike-response transform*, J. Neurosci. Meth., 184:337–356, 2009. 1

[14] D. Song, V.Z. Marmarelis, and T.W. Berger, *Parametric and non-parametric modeling of short-term synaptic plasticity. Part I: Computational study*, J. Comp. Neurosci., 26:1–19, 2009. 1

[15] D. Song, Z. Wang, V.Z. Marmarelis, and T.W. Berger, *Parametric and non-parametric modeling of short-term synaptic plasticity. Part II: Experimental study*, J. Comp. Neurosci., 26:21–37, 2009. 1

[16] R.S. Zucker and W.G. Regher, *Short-term synaptic plasticity*, Ann. Rev. Phys., 64:355–405, 2002. 1

[17] L.F. Abbott, J.A. Varela, K. Sen, and S.B. Nelson, *Synaptic depression and cortical gain control*, Science, 275:221–224, 1997. 1

[18] A. Thomson, *Molecular frequency filters at central synapses*, Prog. Neurobiol., 62:159–196, 2000. 1

[19] J.S. Dittman, A.C. Kreitzer, and W.G. Regehr, *Interplay between facilitation, depression, and residual calcium at three presynaptic terminals*, J. Neurosci., 20:1374–1385, 2000. 1

[20] J.A. Varela, K. Sen, J. Gibson, J. Fost, L.F. Abbott, and S.B. Nelson, *A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex*, J. Neurosci., 17:7926–7940, 1997. 1

[21] J.D. Hunter and J.G. Milton, *Amplitude and frequency dependence of spike timing: Implications for dynamic regulation*, J. Neurophysiol., 90:387–394, 2003. 1

[22] H. Markram, Y. Wang, and M. Tsodyks, *Differential signaling via the same axon of neocortical pyramidal neurons*, Proc. Natl. Acad. Sci. USA, 95:5323–5328, 1998. 1, 2.1.3

[23] M. Tsodyks, A. Uziel, and H. Markram, *Synchrony generation in recurrent networks with frequency-dependent synapses*, J. Neurosci., 20:1–5, 2000. 1

[24] A. Ghanbari, A. Malyshev, M. Volgushev, and I.H. Stevenson, *Estimating short-term synaptic plasticity from pre- and postsynaptic spiking*, PLoS Comput. Bio., 13:e1005738, 2017. 1

[25] H.G. Rotstein, D. Bucher, and F. Nadim, *Decoding short-term plasticity*, Unpublished Results, 2018. 3.1.1, 3.1.1