Comments on *Triplets of Spikes in a Model of Spike Timing-Dependent Plasticity (Pfister and Gerstner, 2006)*

Oliver Y. Chén

Department of Biostatistics
Johns Hopkins University

Abstract

In their 2006 paper, the authors propose a triplet learning rule. Within the framework, All-to-All and Nearest-spikes interactions are investigated, and both visual cortex and hippocampal culture data sets are studied. As a review, we provide some comments organized as follows: first, we attempt to decode several critical equations used in the paper where little information is provided; second, we address some technical and biological questions; and finally, we provide a software program in order to study the visual and quantitative properties of different interactions and STDP learning window included in this paper.

Keywords STDP, Equation decoding, Technical questions, Biological questions.

1 Introduction

Standard STDP models have expressed the weight change as a function of pairs of presynaptic and postsynaptic spike. However, (1) those paired-based STDP models cannot account for the dependence on the repetition frequency of the pairs of spike; and (2) STDP models cannot reproduce recent triplet andquadruplet experiments. To tackle these problems: the authors first review experimental protocols performed in visual cortex and hippocampal culture, and show

why the classical pair-based STDP models fail to reproduce those experimental data; the authors then assume that synaptic plasticity is governed by a suitable combination of pairs and triplets of spikes, and show that the results with the same experimental protocols can be well reproduced. Finally, the authors show that their triplet learning rule relates to that of the BCM theory. Comparisons are given between All-to-All and Nearest-spikes interactions for both pair-based models as well as for triplet-based models, with a preference given for All-to-All interaction. Data analysis is conducted based upon visual cortical and hippocampal culture data sets.

2 Comments

2.1 Equations Decoding

2.1.1 Equation (1) and (2)

For equations (1) and (2), no explicit mathematical forms of the detectors of presynaptic events r_1 and r_2, nor the postsynaptic events o_1 and o_2, are provided. To reproduce such events so as to match the properties in equations (1) and (2), and the trend delineated in Figure 1 B, we consider the following form.

For $i \in \{1, 2\}$,

$$\mathcal{X}_i(t) = e^{-t/\boldsymbol{\tau}} + \mathcal{X}_{i,0} \tag{1}$$

where $\mathcal{X}_i = \{r_i, o_i\}$, and $\boldsymbol{\tau} = \{\tau_+, \tau_x, \tau_-, \tau_y\}$.

Equation (3) and (4) consider weight change as a linear combination of the values of the postsynaptic variable o_1 (resp. presynaptic variable r_1), with the slope being a linear combina-

tion of amplitudes, time t, and the second presynaptic detector r_2 (resp. the second postsynaptic variable o_2). For the choice of choosing (1) the time being first order proportional to o_1 and r_1, and (2) the linear combination of slope, we raise up the following questions:

- The choice of (1) and (2) as linear are not addressed in the paper;
- Is it possible to choose nonlinear combinations;
- What is the biological or neurological meaning of choosing a linear combination.

2.1.2 Equation Involving Δw

In the pairing protocal, the weight change is considered as a function of the frequency ρ for a fixed time Δt. However, the explicit function is not given. Can this function be choosen as the one given in equation (5) in [1]. In particular:

$$\Delta w = \kappa \frac{\partial \mathcal{L}}{\partial w} \tag{2}$$

where $\rho \equiv \kappa$ is the learning rate, and $\mathcal{L}$ is some objective function of t.

Throughout the paper, the authors consider Δw as a function of t; can it be (a) extended to a bivariate function with spikes at some time t, or even further; and (b) be simplified to improve computation efficiency?

For (a), consider:

$$\Delta w = \Delta w(x, t) \tag{3}$$

where x indicates the geometrical position of the synapse on the dendritic tree or the distance between the synapse and the soma, and t.

Specifically, we propose using the bivariate *P*-splines (Eilers and Marx, 2003) to model the function $\Delta w(x,t)$, i.e. $\Delta w(x,t) = \sum_{1\le\kappa\le c_t, 1\le\ell\le c_x} B_\kappa(t)\breve{B}_\ell(x)\gamma_{\kappa\ell} = \{\breve{\mathbf{B}}(x) \otimes \mathbf{B}(t)\}'\boldsymbol{\gamma}$. In

order to mimic the weight change as it in equation (3) and (4) in the paper, we propose:

$$\Delta w(x,t) \to \Delta w(x,t) - f^{\text{pre}}(x,t)$$

if $t = t^{\text{pre}}$, where $f^{\text{pre}}(x,t)$ is some smooth bivariate function.

Similarly,

$$\Delta w(x,t) \to \Delta w(x,t) + f^{\text{post}}(x,t)$$

if $t = t^{\text{post}}$, where $f^{\text{post}}(x,t)$ is some smooth bivariate function.

See Appendix for technique detail.

For (b), whence we consider a more complicated model by introducing more parameters, for every moment of spike t, we may simplify the functional structure and improve computation efficiency via approximating the functional term using for instance a second order Taylor expansion in a neighbourhood of t?

2.2 Other Technical Comments

- "Even if our triplet learning rule captures most of the triplet experiments, the fit is not perfect." Would a non-linear fit be better? Or in general, what is the next step in order to obtain a better fit?

- Equation (3) and (4) seem to not consider an error term. In other words, we propose, for instance, regarding equation (3), for every data point i, consider Δw_i as $\bar{\Delta} + \epsilon_i$, where $\bar{\Delta} := o_1(t)[A_2^- + A_3^- r_2(t-\epsilon)]$, and ϵ_i is the error term following, for example, a known normal law.

2.3 Biological Comments

- What are the benefits of using triplet rule?

- What are the limitations of using triplet rule?
- Are there any ways to capture all the aspects of the pairing, triplet, and quadruplet experiments?
- Are these models the same in all neuronal pathways? Or do different neurotransmitters change the results?
- How can we include the aspects of the dependence of the models on biophysical quantities such as the Ca^{2+} concentration?
- What if the synapse is affected by more than two neurons (pre-synaptic and post-synaptic)?
- The findings of this study are limited to data from hippocampus and visual cortex. What about for data from cerebellum? In other words, can we come up with a more general model - so we do not have to tune for every different data set.
- What are the differences of classical pair-based STDP learning rules and triplet rule?

3 Software

In R:

library(shiny)

runGitHub("oliverychen/pfister")

Note:

(a) For Nearest-spike and All-to-All interactions: the visualization is based upon equation (1), where we set $\mathcal{X}_{i,0} \equiv 0$;

(b) For STDP learning window, we consider the weight change as a function of t using the first

three orders, and an approximate Wiener process where

$$w(t) = 10(\xi_0 t + \sqrt{2}\sum_{i=1}^{500} \xi_i \frac{sin\pi it}{\pi i})$$

where 10 is used to match the magnitude of the other three functions, and 500 is chosen instead of ∞ to reduce computing time.

See attached screenshots for demonstration.

4 Acknowledgement

We would like to thank Dr. Vahid Eslami, M.D., Post-Doctoral Research Fellow at the Department of Neurology of the Johns Hopkins School of Medicine, for his generous help regarding the biological comments.

5 Reference

[1] Optimal Hebbian Learning: a Probabilistic Point of View;

[2] Triplets of Spikes in a Model of Spike Timing-Dependent Plasticity

6 Appendix

6.1 Bivariate Weight Change

Let $\{B_1(t), \ldots, B_{c_t}(t)\}$ be the collection of univariate B-spline basis functions along t, where c_t is the number of interior knots plus the order (degree plus 1) of the B-splines. Similarly, let $\{\breve{B}_1(x), \ldots, \breve{B}_{c_x}(x)\}$ be the collection of B-spline basis functions along x. Then we model $\beta_1(t, x)$ by $\sum_{1\leq\kappa\leq c_t, 1\leq\ell\leq c_x} B_\kappa(t)\breve{B}_\ell(x)\gamma_{\kappa\ell}$, where $\mathbf{\Gamma} = (\gamma_{\kappa\ell})_{1\leq\kappa\leq c_t, 1\leq\ell\leq c_x}$ is a $c_t \times c_x$ coefficient matrix. Let $\mathbf{B}(t) = (B_1(t), \ldots, B_{c_t}(t))'$ be a vector of length c_t and $\breve{\mathbf{B}}(x) = (\breve{B}_1(x), \ldots, \breve{B}_{c_x}(x))'$

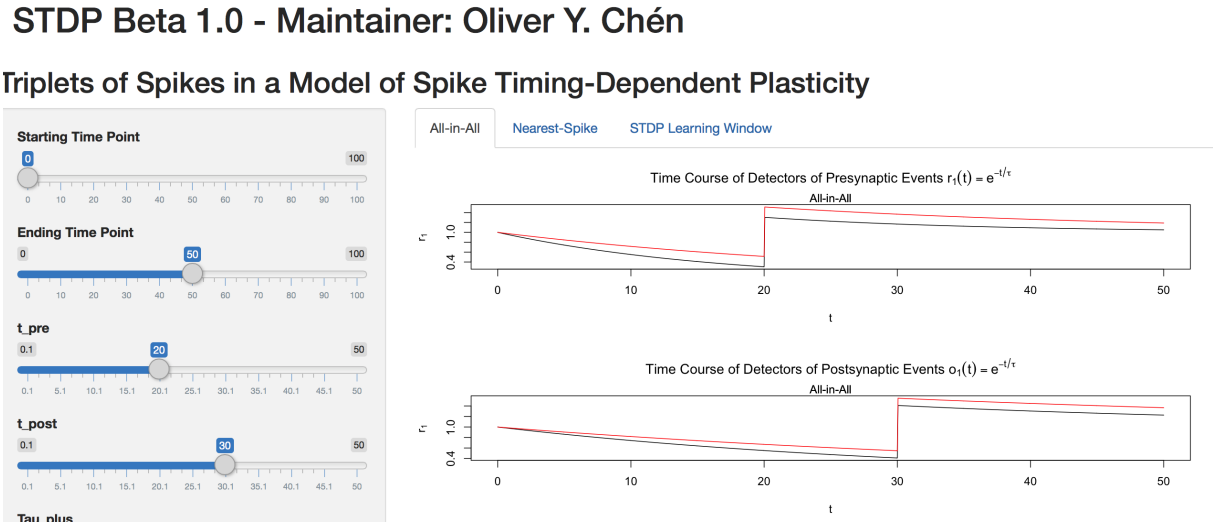
be a vector of length c_x. Then we have

$$\beta_1(t,x) = \sum_{1\le\kappa\le c_t, 1\le\ell\le c_x} B_\kappa(t)\breve{B}_\ell(x)\gamma_{\kappa\ell} = \{\breve{\mathbf{B}}(x)\otimes\mathbf{B}(t)\}'\boldsymbol{\gamma},$$

where $\otimes$ is the Kronecker product of matrices.

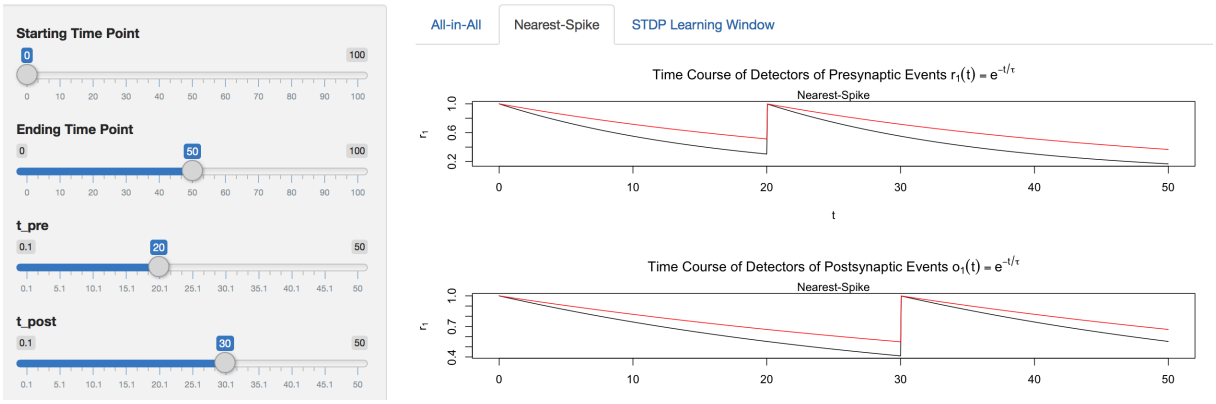
6.2 R Program Representation

	Function	Pair-based model	Full Triplet model	Minimal Triplet (Visual Cortex)	Minimal Triplet (hippocampal)
A_2^+	Amplitude of weight change due to pre-post pair	√	√		√
A_2^-	Amplitude of weight change due to post-pre pair	√	√	√	√
A_3^+	Amplitude of potentiation		√	√	√
A_3^-	Amplitude of depression		√		
τ_+	Time constant	√	√	√	√
τ_-	Time constant	√	√	√	√
τ_x	Time constant		√	√	√
τ_y	Time constant		√	√	√



STDP Beta 1.0 - Maintainer: Oliver Y. Chén

Triplets of Spikes in a Model of Spike Timing-Dependent Plasticity



STDP Beta 1.0 - Maintainer: Oliver Y. Chén

Triplets of Spikes in a Model of Spike Timing-Dependent Plasticity