



Spike Pattern Structure Influences Synaptic Efficacy Variability under STDP and Synaptic Homeostasis. I: Spike Generating Models on Converging Motifs

Zedong Bi^{1,2*} and Changsong Zhou^{2,3,4,5*}

¹ State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing, China, ² Department of Physics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, ³ Centre for Nonlinear Studies, Beijing-Hong Kong-Singapore Joint Centre for Nonlinear and Complex Systems, Institute of Computational and Theoretical Studies, Hong Kong Baptist University, Kowloon Tong, Hong Kong, ⁴ Beijing Computational Science Research Center, Beijing, China, ⁵ Research Centre, HKBU Institute of Research and Continuing Education, Shenzhen, China

OPEN ACCESS

Edited by:

Guenther Palm,
University of Ulm, Germany

Reviewed by:

Thomas Wennekers,
University of Plymouth, UK
Valeri Makarov,
Universidad Complutense de Madrid,
Spain

*Correspondence:

Zedong Bi
zedong.bi@outlook.com;
Changsong Zhou
cszhou@hkbu.edu.hk

Received: 14 November 2015

Accepted: 01 February 2016

Published: 23 February 2016

Citation:

Bi Z and Zhou CS (2016) Spike Pattern Structure Influences Synaptic Efficacy Variability under STDP and Synaptic Homeostasis. I: Spike Generating Models on Converging Motifs. *Front. Comput. Neurosci.* 10:14. doi: 10.3389/fncom.2016.00014

In neural systems, synaptic plasticity is usually driven by spike trains. Due to the inherent noises of neurons and synapses as well as the randomness of connection details, spike trains typically exhibit variability such as spatial randomness and temporal stochasticity, resulting in variability of synaptic changes under plasticity, which we call *efficacy variability*. How the variability of spike trains influences the efficacy variability of synapses remains unclear. In this paper, we try to understand this influence under pair-wise additive spike-timing dependent plasticity (STDP) when the mean strength of plastic synapses into a neuron is bounded (synaptic homeostasis). Specifically, we systematically study, analytically and numerically, how four aspects of statistical features, i.e., synchronous firing, burstiness/regularity, heterogeneity of rates and heterogeneity of cross-correlations, as well as their interactions influence the efficacy variability in converging motifs (simple networks in which one neuron receives from many other neurons). Neurons (including the post-synaptic neuron) in a converging motif generate spikes according to statistical models with tunable parameters. In this way, we can explicitly control the statistics of the spike patterns, and investigate their influence onto the efficacy variability, without worrying about the feedback from synaptic changes onto the dynamics of the post-synaptic neuron. We separate efficacy variability into two parts: the drift part (DriftV) induced by the heterogeneity of change rates of different synapses, and the diffusion part (DiffV) induced by weight diffusion caused by stochasticity of spike trains. Our main findings are: (1) synchronous firing and burstiness tend to increase DiffV, (2) heterogeneity of rates induces DriftV when potentiation and depression in STDP are not balanced, and (3) heterogeneity of cross-correlations induces DriftV together with heterogeneity of rates. We anticipate our work important for understanding functional processes of neuronal networks (such as memory) and neural development.

Keywords: spike pattern structure, synaptic plasticity, efficacy variability, STDP, synaptic homeostasis, spike generating models

1. INTRODUCTION

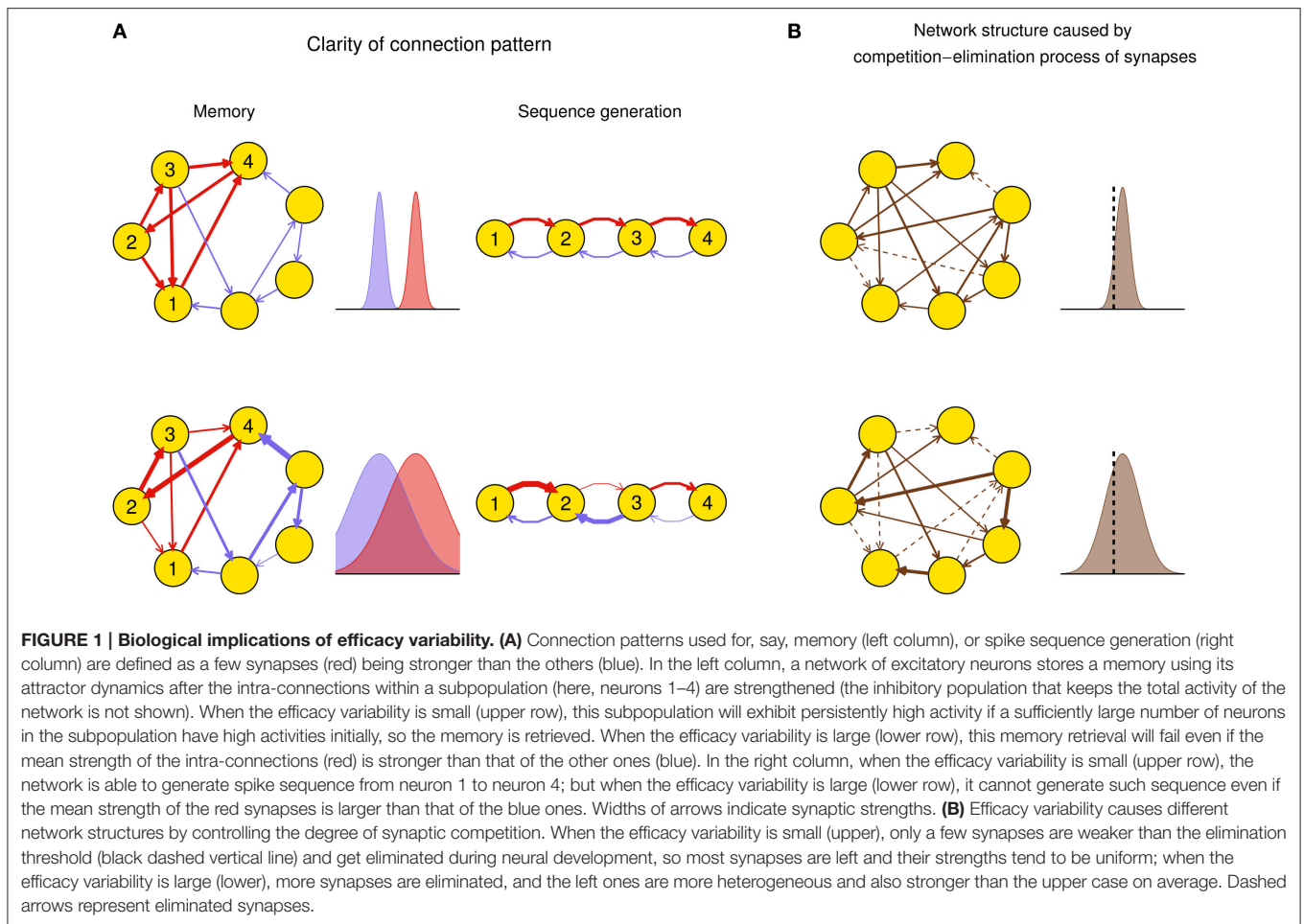
Neuronal spike trains typically exhibit spatial randomness and temporal stochasticity. For example, firing rates are long-tailed distributed in many brain areas (Shafi et al., 2007; O'Connor et al., 2010; Buzsáki and Mizuseki, 2014), spatio-temporal correlations within neuronal population often exhibit rich structures (Funahashi and Inoue, 2000; Kohn and Smith, 2005; Dragoi and Buzsáki, 2006; Schneidman et al., 2006); and two neurons will not emit the same spike train even if they are receiving exactly the same stimuli (Allen and Stevens, 1994; Mainen and Sejnowski, 1995; Shadlen and Newsome, 1998). The spatial randomness may emerge from the randomness of the connection details (Ostojic et al., 2009; Roxin et al., 2011), and the temporal stochasticity may be due to the inner stochasticity of neurons and synapses (Allen and Stevens, 1994; Mainen and Sejnowski, 1995; Shadlen and Newsome, 1998), both of which are inherent properties of neurons, synapses or networks so that the exact spike patterns of the network cannot be fully determined by its inputs. Note that the inputs of a network may exhibit spatial heterogeneity and temporal fluctuation; here, by spatial randomness and temporal stochasticity, we mean the spatial heterogeneity and temporal fluctuation of spike patterns that emerge from the inner randomness and stochasticity of the network. As synaptic plasticity depends on the spike times in the pre- and post-synaptic spike trains (Dan and Poo, 2006; Caporale and Dan, 2008; Markram et al., 2012), these variabilities of spike trains should result in the variability of synaptic changes during plasticity, which we call *efficacy variability* in this paper (See Section 2.1 for more discussions on efficacy variability). When the synapses are fixed, downstream neurons may work under the variability of spike trains by reading out the coded information through spatial and temporal averaging; however, how the variability of the spike trains influences the ability of the neuronal population to facilitate information processing under synaptic plasticity remains poorly understood.

Efficacy variability may have important influence on the function of a network after plasticity. For example, suppose a function of a neuronal network, say memory (Mongillo et al., 2008) or spike sequence generation (Long et al., 2010), requires a connection pattern in which a few synapses (foreground synapses) have stronger efficacies than the others (background synapses). When the efficacy variability is small, both the foreground and background synapses tend to be uniform around their mean values, respectively: thus the connection pattern is clear-cut. However, when the efficacy variability is large, some foreground synapses can be very weak and some background ones can be very strong, which destroys the connection pattern even if the mean strength of the foreground synapses is still larger than that of the background ones (**Figure 1A**). As another example, synaptic competition and elimination is a classical scenario for the formation of neural network structure during development, when synapses compete with each other for strength and those that are too weak will disappear (Cancedda and Poo, 2009). In this case, efficacy variability quantifies the degree of competition. If we suppose that the total synaptic strength before elimination is constrained by, say, synaptic

homeostasis (Watt and Desai, 2010), then when the efficacy variability is small, only a few synapses are below the elimination threshold and get eliminated, and those left also have similar strength; when the efficacy variability is large, a larger portion of synapses get below the elimination threshold, while the remaining ones have a wider efficacy distribution with also a larger mean value than the case of small efficacy variability (**Figure 1B**). This is consistent with the scenario found in the early development of auditory cortex (Clause et al., 2014): if the spontaneous activity of the medial nucleus of the trapezoid body (MNTB) is modified using genetic methods, then its feedforward projection to the lateral superior olive (LSO) becomes denser and weaker, which suggests that the normal pattern induces stronger efficacy variability than the genetically modified one.

Under temporal stochasticity and spatial heterogeneity, spike trains may exhibit a variety of statistical features, which form rich spike pattern structures. Groups of neurons may spurt firing activity (*synchronous firing*) (Kamioka et al., 1996; Buzsáki and Draguhn, 2004; Bartos et al., 2007), the spike train of a single neuron can be bursty or regular (*auto-correlation structure*) (Softky and Koch, 1993; Schwindt and Crill, 1999; Jacob et al., 2012), firing rates of cortical neurons are typically long-tailed distributed *in vivo* (*heterogeneity of rates*) (Shafi et al., 2007; O'Connor et al., 2010; Buzsáki and Mizuseki, 2014), and spike trains of different neurons also reveal rich degrees of interdependence (*heterogeneity of cross-correlations*) (Funahashi and Inoue, 2000; Schneidman et al., 2006; Ostojic et al., 2009; Trousdale et al., 2012, see **Figure 2A**). As synaptic plasticity is driven by spike trains, spike pattern structure must have strong influence on efficacy variability, thereby inducing neuronal networks with sharply different structures even under the same population rate. How spike pattern structures influence the efficacy variability of synapses remains unclear. In this study, we will move a step forward along this direction by studying this influence under a conventional pair-wise additive STDP (Gerstner et al., 1996), which involves potentiation of the synapse when presynaptic spikes precede postsynaptic spikes, and depression for the reverse ordering (**Figure 2B**). Neuronal networks with STDP alone are not stable due to effects of positive resonance which results in runaway excitation; and a neuron may conserve its activity level by adjusting its total input strength, which is called synaptic homeostasis (Turrigiano and Nelson, 2004; Turrigiano, 2011). In this study, we model synaptic homeostasis by supposing that the mean strength of all the synapses input to a neuron is dynamically bounded (see **Figure 2C**, Equation 7 in Section 2).

Synaptic efficacies under STDP can be regarded as particles doing 1-dimensional random walks driven by the stochastic processes of spike trains: a synapse gets depressed at the time when a pre-synaptic spike arrives at the axonal terminal, and gets potentiated at the time when a post-synaptic spike is back-propagated to the dendritic end. A group of particles doing random walks starting from the same point can have displacement variability either because that they have different trial-averaged drift velocities with each other, or because of diffusion. Similarly, the efficacies of two synapses can also get dissimilar because of two reasons. One is that they have different



tendencies to be potentiated or depressed. For example, the first (second) pre-synaptic neuron tends to fire spikes before (after) the first (second) post-synaptic neuron. The other reason is the variability of the synaptic changes caused by the stochasticity of the spike trains, which can be accumulated with time, inducing dissimilarity even if the two synapses have the same tendency to be potentiated or depressed. In general, the total variance (TotalV) can be written as the summation of the variance caused by heterogeneity of learning rates (DriftV, short for “drift variance”) and the variance caused by diffusion (DiffV, short for “diffusion variance”) (see Equation 3 in Section 2)

$$\text{TotalV} = \text{DriftV} + \text{DiffV}. \quad (1)$$

During plasticity, DriftV is usually caused by the spatial heterogeneity of spike trains. For example, in classical Hebbian learning, synapses sharing the same presynaptic neuron may have different learning rates depending on the firing rates of the post-synaptic neurons; if the plasticity is spike-timing dependent, the heterogeneity of cross-correlations can induce different learning rates even if the firing rates are the same. Because of the inner stochasticity of neurons and synapses, even two neurons receiving exactly the same stimuli emit different spike trains,

causing DiffV. In our STDP model, synaptic updates do not depend on current synaptic weights, and the contributions of all spike pairs are added together; what’s more, we do not consider synaptic bounds, and assume that synapses are free of bounds for simplicity. These make $\text{DriftV} \propto t^2$ and $\text{DiffV} \propto t$ during time t of running, so that the synaptic variances caused by drift velocities and diffusion in our model have the same order of time-dependence as those in Brownian motion.

In this paper, we systematically study how the four aspects of spike pattern structures, i.e., synchronous firing, auto-correlation structure, heterogeneity of rates and heterogeneity of cross-correlations as well as their interactions, influence efficacy variability under STDP and synaptic homeostasis using converging motifs (i.e., simple networks in which one neuron receives from many other neurons). See **Figure 2** for the concepts above. The activities of all the neurons in converging motifs (including the non-central neurons and the central neuron, see **Figure 2D**) are generated using statistical models (Section 2.3), which explicitly control different aspects of pattern structure while keeping population rate constant.

The reason why we study converging motifs here is that in a recurrent network, all the synapses input to a neuron form the links of a converging motif, and all the synapses of the

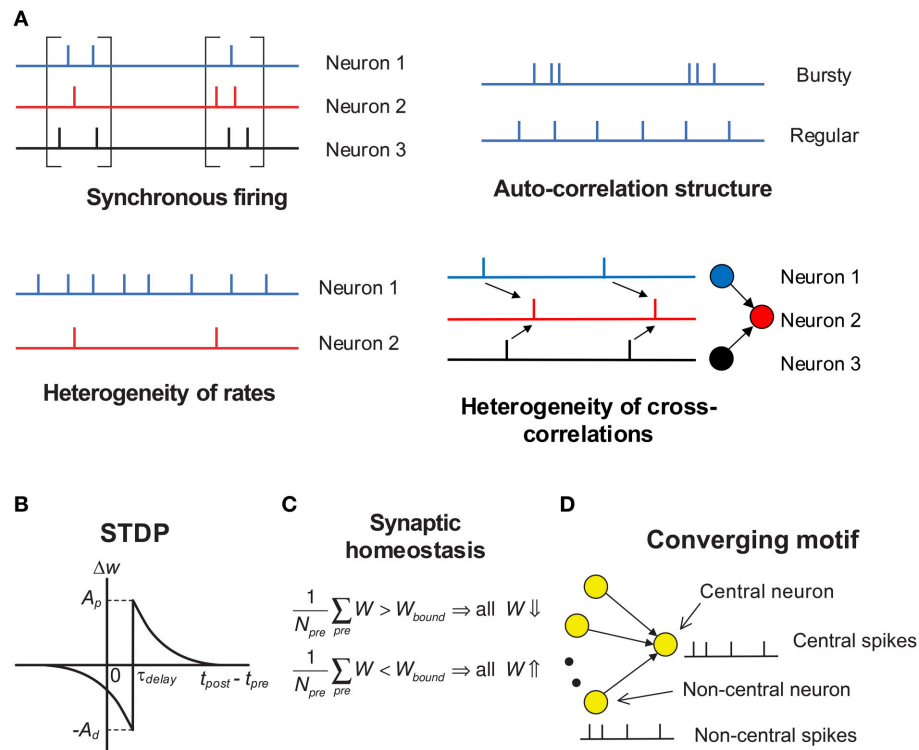


FIGURE 2 | Schematic of key concepts in our modeling work. (A) The four aspects of pattern structure studied in this paper. By “synchronous firing”, we typically mean the spurt of firing activity of a population. For stationary spike trains, “auto-correlation structure” reflects the burstiness/regularity of the spike trains, which is quantified by coefficient of variance (CV) in this paper. Here, by “burstiness”, we typically mean the irregular structure of spike trains, instead of the regular burstiness in the spike patterns of, say, central pattern generator. For spike trains with synchronous firing, we consider three-types of “auto-correlation structure” to reflect the burstiness/regularity features of the spike patterns (see **Figure 8**). By “heterogeneity of rates”, we mean that the time-averaged firing rates are different for different neurons. By “heterogeneity of cross-correlations”, we mean that different pre-synaptic neurons of a neuron tends to fire spikes at different times relative to the spikes of the neuron. For example, in the right-bottom subplot, before a spike of neuron 2, neuron 1 tends to fire before neuron 3. **(B)** The STDP time window used in our work. The synapses are updated according to the time difference between when the post-synaptic spike is back-propagated to the dendritic end and when the pre-synaptic spike arrives at the axonal terminal. τ_{delay} is the difference between the axonal delay and the dendritic delay. In this paper, we generally discuss the case when $\tau_{delay} > 0$ (i.e., the axonal delay is larger than the dendritic delay); and for the convenience of discussion, we set the dendritic delay to be zero, so τ_{delay} is the axonal delay. The case when $\tau_{delay} < 0$ is then discussed in Section 3.8. The STDP updatings of all spike pairs are summed together. **(C)** Synaptic homeostasis. The synapses input to a neuron are subject to a bound on their mean strength: when their mean strength is different from this bound, all the incoming synapses of that neuron will undergo an adjustment. **(D)** Converging motif, on which we conduct all the simulations in this paper. We call the neuron that receives from many neurons in a converging motif to be the *central neuron*, and call the other neurons the *non-central neurons*; we then call the spikes emitted by the central neuron (or non-central neurons) to be *central spikes* (or *non-central spikes*). Synaptic homeostasis is imposed onto the central neuron of a converging motif. Modeling details are presented in Section 2.

recurrent network can be considered as the union of the links in all the converging motifs. Under synaptic homeostasis, the mean synaptic strengths in these converging motifs are almost the same, therefore, the efficacy variability of the recurrent network is approximately to be the mean of the efficacy variability of all the converging motifs in it (see Equation 8 in Section 2). Therefore, converging motifs are basic units to understand the efficacy variability of a recurrent network.

Using statistical models (see Section 2.3), we generate spike patterns with different statistical features (such as coefficients of variance (CV), firing rate distributions, distributions of the spike number that a neuron fires during a synchronous event, and so on), and then study how they influence the efficacy variability of converging motifs. These spike generating models do not aim to generate spike trains with precise pre-determined

spatio-temporal characteristics (see, for example, Krumin and Shoham, 2009; Macke et al., 2009; Gutnisky and Josić, 2010), but instead aim to implement physical intuitions on spatio-temporal characteristics in a simple way. For example, we use Gamma processes with different shape parameters to model bursty or regular spike trains, and use lognormal distributions to model heterogeneity of rates. Note that in our study, the spike trains of the central neuron are not generated by integrating its inputs, but instead are also generated according to statistical models, independent of the synaptic weights in converging motifs. In this way, we can explicitly control the statistical features of both the central and non-central spike trains using statistical models, and study their influences onto the efficacy variability without worrying about the feedback of the synaptic changes onto the spike train of the post-synaptic

neuron as usually happens in biologically more realistic models.

The spike trains of both the central and non-central neurons are generated according to the same model, so that they have similar auto-correlation structure and synchronous firing rate fluctuation. This makes our results particularly suitable to understand the efficacy variability of a recurrent network without hidden feedforward dynamics (Ganguli and Latham, 2009). For example, in a random excitatory-inhibitory network that works in the balanced state (van Vreeswijk and Sompolinsky, 1996, 1998), spike trains exhibit irregularity and stochasticity due to being driven by fluctuation. If the dynamics of the network is unstable to perturbations of population firing rates, then oscillation may emerge (Brunel and Hakim, 1999; Brunel, 2000), resulting in synchronous firing (**Figure 2A**). If its dynamics is unstable to heterogeneous perturbations, then the neurons may fire with strong burstiness (Ostojic, 2014) (auto-correlation structure). Due to the heterogeneity of the input degrees and nonlinear current-rate relationship of neurons, their firing rates are typically long-tailed distributed (Roxin et al., 2011) (heterogeneity of rates). Cross-correlations between neurons may result from connectivity details such as unidirectional connections or input sharing (Ostojic et al., 2009) (heterogeneity of cross-correlations). In this paper, we aim to mimic these spike patterns using statistical models, thereby gaining understanding on their influence onto the efficacy variability of a network. In another paper that will be published soon (Bi and Zhou, in preparation), we will study the influence of spike pattern structures onto the efficacy variability of recurrent networks by implementing sophisticated spike shuffling methods onto the spike patterns self-organized by recurrent LIF networks, thereby examining the results of this paper in a biologically more plausible manner.

The Results part of this paper is organized as follows. We will first study how auto-correlation structure, synchronous firing, heterogeneity of rates and their interactions influence DiffV and DriftV (Sections 3.1–3.6), then discuss the influence of heterogeneity of cross-correlations onto DriftV (Section 3.7). Under STDP, the main effect of heterogeneity of cross-correlations is to change the drift velocities of different synapses by different degrees, thereby influencing DriftV; so we don't seriously consider how heterogeneity of cross-correlations influence DiffV, except for briefly discussing it in Supplementary Materials Section S6. In this paper, we generally discuss the case when $\tau_{delay} > 0$ in STDP (i.e., the axonal delay is larger than the dendritic delay, see **Figure 2B**); and for the convenience of discussion, we set the dendritic delay to be zero, so τ_{delay} becomes the axonal delay. The case when $\tau_{delay} < 0$ is then discussed in Section 3.8. The results of this paper are then summarized in Section 3.9.

2. MATERIALS AND METHODS

2.1. The Definition of Efficacy Variability

In this subsection, we give the exact definition of efficacy variability, explain the meanings of DiffV and DriftV, and

show our general strategy to study efficacy variability using simulations.

Suppose a set of synapses \mathcal{W} in a neuronal network \mathcal{N} . Now we run a plasticity process in \mathcal{N} for several trials (or on several ensembles), and construct a matrix $\Delta\mathbf{W}$, each column of which represents the weight changes of \mathcal{W} in one trial, and different columns represent different trials. Then we define *efficacy variability* of \mathcal{W} during the plasticity process to be the variance of the elements of the matrix $\Delta\mathbf{W}$, i.e., $\text{Var}_{S,T}(\Delta\mathbf{W})$, in which the subscript S represents integrating over row index, i.e., structural index, and T represents integrating over column index, i.e., trial index.

The law of total variance says that if the probability space of Y is decomposed into several subspaces labeled by X , then the variance of Y in the whole space is equal to the summation of the variance of the expectations in these subspaces and the expectation of the variances in these subspaces, i.e.,

$$\text{Var}(Y) = \text{Var}(E(Y|X)) + E(\text{Var}(Y|X)) \quad (2)$$

Using the law of total variance, efficacy variability can be written as

$$\text{Var}_{S,T}(\Delta\mathbf{W}) = \text{Var}_S(E_T(\Delta\mathbf{W})) + E_S(\text{Var}_T(\Delta\mathbf{W})) \quad (3)$$

Here, $E_T(\Delta\mathbf{W})$ represents the trial expectations of the changes of all the synapses in \mathcal{W} ; and $\text{Var}_S(E_T(\Delta\mathbf{W}))$ is the variance of these trial expectations, representing DriftV. $\text{Var}_T(\Delta\mathbf{W})$ represents the trial-to-trial variances caused by diffusion, and $E_S(\text{Var}_T(\Delta\mathbf{W}))$ is the average of these variances over all the synapses, representing DiffV. Equation 3 is the formal writing of Equation 1 in the introduction.

The law of total variance can decompose $\text{Var}_{S,T}(\Delta\mathbf{W})$ in another way:

$$\text{Var}_{S,T}(\Delta\mathbf{W}) = \text{Var}_T(E_S(\Delta\mathbf{W})) + E_T(\text{Var}_S(\Delta\mathbf{W})) \quad (4)$$

Here $\text{Var}_T(E_S(\Delta\mathbf{W}))$ is the trial-to-trial variability of the mean synaptic change of the whole network. But a real biological process only allows a single trial, so this trial-to-trial variability cannot contribute to biological functions except for individual differences. Fortunately, usually $\text{Var}_T(E_S(\Delta\mathbf{W})) \sim \mathcal{O}(1/|\mathcal{W}|)$, with $|\mathcal{W}|$ being the number of synapses in \mathcal{W} . So when $|\mathcal{W}|$ is large enough, Equation 4 becomes $\text{Var}_{S,T}(\Delta\mathbf{W}) \approx E_T(\text{Var}_S(\Delta\mathbf{W}))$, which means that the trial expectation of the efficacy variance (i.e., $E_T(\text{Var}_S(\Delta\mathbf{W}))$) can be used to approximate the efficacy variability $\text{Var}_{S,T}(\Delta\mathbf{W})$. Under this insight, Equation 3 becomes

$$\begin{aligned} E_T(\text{Var}_S(\Delta\mathbf{W})) &\approx \text{Var}_S(E_T(\Delta\mathbf{W})) + E_S(\text{Var}_T(\Delta\mathbf{W})) \\ &= \text{DriftV} + \text{DiffV} \end{aligned} \quad (5)$$

In our simulations, we use $E_T(\text{Var}_S(\Delta\mathbf{W}))$ to quantify efficacy variability. As we mentioned in the introduction, DriftV is usually caused by the spatial heterogeneity of spike trains, and DiffV is by temporal stochasticity. Methodologically, when we analyze DiffV, we set the spatial properties of spike trains (here, firing rates of neurons and cross-correlations between neuronal pairs)

homogeneous, so that $\text{DriftV} = 0$. When we analyze DriftV , we use the fact that $\text{DriftV} \sim \mathcal{O}(t^2)$ and $\text{DiffV} \sim \mathcal{O}(t)$ after time t of running, so the efficacy variability after a sufficiently long time largely reflects DriftV .

2.2. Converging Motifs, STDP and Synaptic Homeostasis

Converging motifs are simple networks in which one neuron receives inputs from many other neurons (**Figure 2D**). We call the neuron that receives from many neurons in a converging motif to be the *central neuron*, and call the other neurons the *non-central neurons*; we then call the spikes emitted by the central neuron (or non-central neurons) to be *central spikes* (or *non-central spikes*). All the simulations in this paper are conducted on converging motifs.

In this paper, the STDP updating caused by a pair of pre- and post-synaptic spike (here, non-central and central spike) at t_{pre} and t_{post} is

$$\Delta w(t_{pre}, t_{post}) = \begin{cases} A_p \exp\left(-\frac{t_{post} - (t_{pre} + \tau_{delay})}{\tau_{STDP}}\right), & t_{post} > t_{pre} + \tau_{delay} \\ -A_d \exp\left(-\frac{(t_{pre} + \tau_{delay}) - t_{post}}{\tau_{STDP}}\right), & t_{post} < t_{pre} + \tau_{delay} \end{cases} \quad (6)$$

with τ_{delay} being the difference between the delay caused by the propagation of pre-synaptic spike along the axon and the delay caused by the back-propagation of the post-synaptic spike along the dendrite. $\tau_{delay} > 0$ (or $\tau_{delay} < 0$) if the axonal delay is larger (or smaller) than the dendritic delay. In this paper, we generally discuss the case when $\tau_{delay} > 0$ (i.e., the axonal delay is larger than the dendritic delay); and for the convenience of discussion, we set the dendritic delay to be zero, so τ_{delay} becomes the axonal delay. The case when $\tau_{delay} < 0$ is then discussed in Section 3.8. The contribution of all pairs of pre- and post-synaptic spikes are added together. $\tau_{STDP} = 20$ ms throughout the paper, and $\tau_{delay} = 1$ ms by default. For simplicity, synaptic weights in this paper are free of boundaries.

Under synaptic homeostasis, the synaptic efficacies are updated every ΔT time according to

$$w_a(t + \Delta T) = w_a(t) + \epsilon(w_{bound} - \frac{1}{N_{in}} \sum_{c=1}^{N_{in}} w_c(t)), \quad (7)$$

with w_a being the synaptic efficacy between the a th non-central neuron and the central neuron, N_{in} being the in-degree of central neuron, w_{bound} being the ground line of synaptic homeostasis, and ϵ being the plasticity rate. Therefore, synaptic homeostasis dynamically bounds the mean strength of the synapses to the central neuron, and it does not change the efficacy variability of a converging motif due to its additive nature.

According to the law of total variance Equation 2, the efficacy variability of a large recurrent network can be written as

$$\text{Var}_{ab}(\Delta w_{ab}) = \text{Var}_a(\text{E}_{b \in \partial a}(\Delta w_{ab})) + \text{E}_a(\text{Var}_{b \in \partial a}(\Delta w_{ab})), \quad (8)$$

with Δw_{ab} being the synaptic change from neuron b to neuron a , and ∂a being all the neurons that input to neuron a in the network. If synaptic homeostasis Equation 7 is imposed onto the synapses input to each neuron in the network, $\text{Var}_a(\text{E}_{b \in \partial a}(\Delta w_{ab})) \approx 0$, especially if the plasticity rate is high (i.e., $\epsilon \approx 1$ in Equation 7). Therefore, the efficacy variability of the recurrent network is approximately the mean of the efficacy variability of all the converging motifs in it. This is the reason why we study converging motifs in this paper.

2.3. Spike Generating Models

Here are the statistical models we use to generate the spike trains of the neurons in a converging motif (**Figure 2D**). The spike trains of the central neuron and the non-central neurons are generated according to the same model, mimicking the dynamics of converging motifs embedded in a recurrent network without hidden feedforward dynamics (Ganguli and Latham, 2009).

2.3.1. Model Auto

This model generates spike trains with stationary firing rate, whose auto-correlation structure (burstiness/regularity) can be controlled.

Spike trains are Gamma processes with inter-spike intervals following the distribution

$$p(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$$

The rate of the Gamma process is β/α , and the coefficient of variance is $1/\sqrt{\alpha}$.

We use α to control the burstiness/regularity of the spike train, while adjusting β to keep the firing rate at 20 Hz. The spike train becomes more bursty when α is smaller, and more regular when α is large.

2.3.2. Model Sync

This model generates synchronous firing patterns whose broadness of the distribution of spike number per neuron per synchronous event can be explicitly controlled.

In this model, a synchronous event lasts for τ_{cross} . The spike number per neuron per synchronous event is distributed the same as the spike number of an underlying Gamma process within an interval of length τ_{cross} . The rate of the Gamma process is p/τ_{cross} and its coefficient of variance is $CV_{SpikeNum}$. So the mean value of the distribution of spike number per neuron per synchronous event is p , and this distribution is narrow if $CV_{SpikeNum}$ is small, and gets broadened when $CV_{SpikeNum}$ increases. If a neuron is to fire M spikes in a synchronous event, then the spike times of these M spikes will be randomly and uniformly chosen within this time interval of duration τ_{cross} . The occurrence of synchronous events is a Poisson process with rate r_0/p , so that the firing rate of a neuron is kept at $r_0 = 20$ Hz when p changes.

2.3.3. Model Long Tail

This model generates long-tailed distributed firing rates for the non-central neurons in a converging motif, the firing rate of the central neuron is always kept at $r_0 = 20$ Hz.

The firing rates of the non-central neurons are lognormal distributed as

$$p(x|m, s) = \frac{1}{sx\sqrt{2\pi}} \exp\left[-\frac{(\ln x - m)^2}{2s^2}\right]$$

The mean of this distribution is at $\exp(m + \frac{s^2}{2})$. Parameter s is used to control the shape, while m is accordingly adjusted to keep the mean at $r_0 = 20$ Hz. This distribution is a δ function when $s = 0$, and gradually becomes long tailed when s increases.

This model generates homogeneous Poisson processes by default, but it can be also combined with other models to introduce heterogeneity of rates into the spike patterns with other aspects of pattern structure.

2.3.4. Model Sync-Auto-LongTail

In this model, the firing rates of the non-central neurons follow log-normal distribution with mean $r_0 = 20$ Hz (with r_0 being the firing rate of the central neuron) and shape parameter s . A synchronous event lasts for τ_{cross} , and the mean spike number per neuron per synchronous event is p . The occurrence of synchronous events is a Gamma process with coefficient of variance CV_{events} and rate r_0/p , so that the population firing rate is kept at r_0 when p changes. The spike train of a neuron with firing rate r_i is generated as follows: we generate a Gamma process with rate pr_i/r_0 and coefficient of variance $CV_{rescale}$, and divide the Gamma spike train into bins of length τ_{cross} . Then we randomly shuffle the order of these bins to destroy possible spike dependency between adjacent bins, and the piece of spike train within the j th bin after shuffling is to be the piece of spike train of the neuron within the j th synchronous event.

In this model, the auto-correlation structure of the occurrence of synchronous events (see **Figure 10**) is controlled by the parameter CV_{events} , while the auto-correlation structure of a spike train in the rescaled time is controlled by the parameter $CV_{rescale}$.

3. RESULTS

3.1. The Scheme to Investigate the Influence of Synchronous Firing and Auto-correlation Structure onto DiffV

When the firing rates of neurons and the cross-correlations between neurons are homogeneous, synchronous firing and auto-correlation structure influence efficacy variability as a DiffV effect. To understand how these two pattern statistics influence DiffV, we write the total change of the a th synapse as

$$\Delta w_a = \sum_{k=p,d} \Delta w_{a,k} = \sum_{k=p,d} \sum_i \sum_j \Delta w_{a,k}(t_i, t_{j|a,i,k}), \quad (9)$$

with $\Delta w_{a,p}$ being the total potentiation value on the synapse, $\Delta w_{a,p}(t_i, t_{j|a,i,p})$ being the potentiation value pairing the i th spike of central neuron and the j th spike of the a th non-central neuron (the index of j depends on a, i , and k , see below), and $\Delta w_{a,d}$ and $\Delta w_{a,d}(t_i, t_{j|a,i,d})$ being the corresponding depression values. The index of i starts from the beginning of the spike train of

the central neuron; but the indexing of j , however, is a little complicated, and depends on a, i and k . In our modeling work, axons have delay τ_{delay} ($\tau_{delay} = 1$ ms by default), and the sign of STDP updatings depends on the time difference between $t_i - \tau_{delay}$ and $t_{j|a,i,k}$. For the potentiation process ($k = p$) of the a th non-central neuron, we let j start from the spike immediately before $t_i - \tau_{delay}$ in the spike train, and go backward along the spike train; for the depression process ($k = d$), we let j start from the spike immediately after $t_i - \tau_{delay}$, and go forward along the spike train (**Figure 3A**). In the following discussions, we sometimes do not explicitly write the dependence of j on a, i and k for simplicity.

Using Equation 9, we can rewrite the variance of efficacy changes per spike like this:

$$\text{Var}_a(\Delta w_a)/\bar{N}_0 = \text{Var}_a\left(\sum_{k=p,d} \sum_i \sum_j \Delta w_{a,k}(t_i, t_j)\right)/\bar{N}_0 = c_{II} \cdot c_I \cdot d \quad (10)$$

with $\bar{N}_0 = r_0 T$ being the expectation of the spike number of the central neuron during simulation time (r_0 is the firing rate of the central neuron, and T is the duration of simulation), and

$$c_{II} = \frac{\text{Var}_a(\sum_k \sum_i \sum_j \Delta w_{a,k}(t_i, t_j))}{\sum_k \text{Var}_a(\sum_i \sum_j \Delta w_{a,k}(t_i, t_j))} \quad (11)$$

$$c_I = \frac{\sum_k \text{Var}_a(\sum_i \sum_j \Delta w_{a,k}(t_i, t_j))}{\sum_k \sum_i \text{Var}_a(\sum_j \Delta w_{a,k}(t_i, t_j))} \quad (12)$$

and

$$d = \sum_k \sum_i \text{Var}_a\left(\sum_j \Delta w_{a,k}(t_i, t_j)\right)/\bar{N}_0 \quad (13)$$

Apparently, d means the mean efficacy variability contributed by a single spike of the central neuron. To understand c_{II} , note that

$$\text{Var}_a(\Delta w_a) = \sum_{k=p,d} \text{Var}_a\left(\sum_i \sum_j \Delta w_{a,k}(t_i, t_j)\right)$$

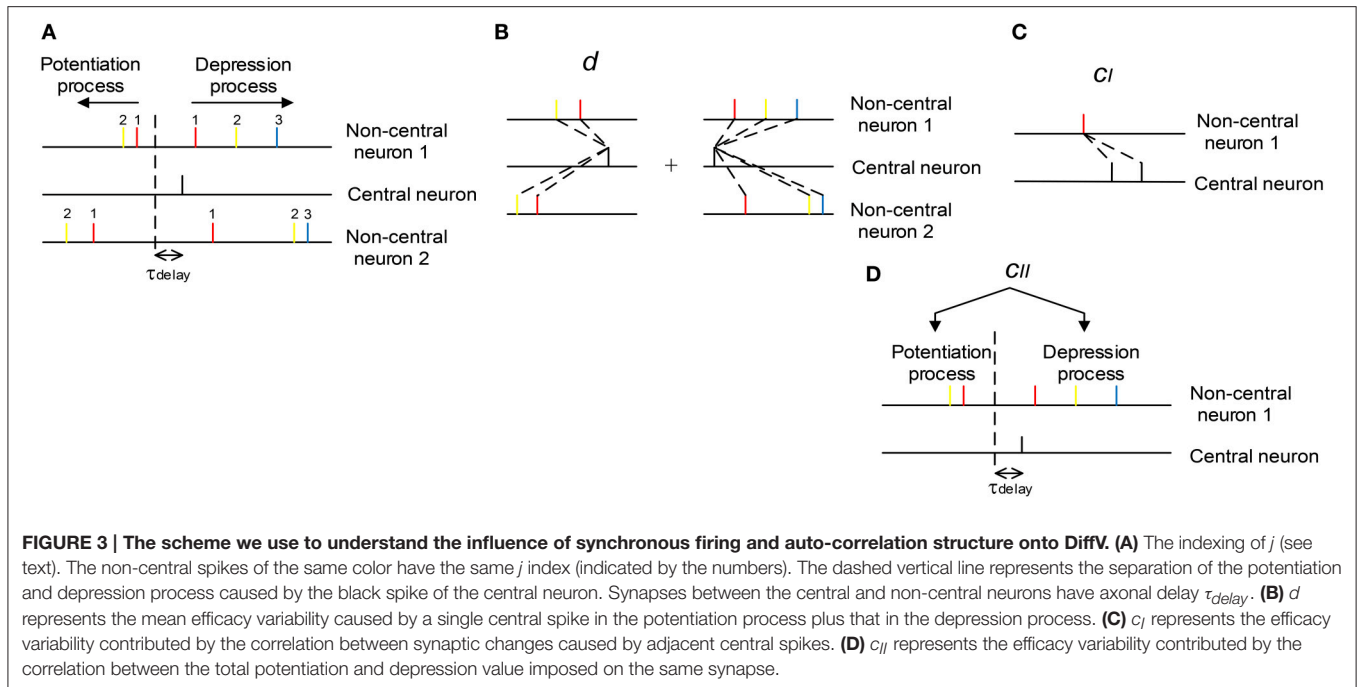
$$+ 2\rho_{PD} \sqrt{\text{Var}_a\left(\sum_i \sum_j \Delta w_{a,p}(t_i, t_j)\right) \cdot \text{Var}_a\left(\sum_i \sum_j \Delta w_{a,d}(t_i, t_j)\right)} \quad (14)$$

with $\rho_{PD} = \text{Corr}_a(\sum_i \sum_j \Delta w_{a,p}(t_i, t_j), \sum_i \sum_j \Delta w_{a,d}(t_i, t_j))$ being the correlation coefficient between the total potentiation and depression value imposed on the same synapse. Therefore,

$$c_{II} = 1 + \rho_{PD} f_{PD} \quad (15)$$

with

$$f_{PD} = \frac{2\sqrt{\text{Var}_a(\sum_i \sum_j \Delta w_{a,p}(t_i, t_j)) \cdot \text{Var}_a(\sum_i \sum_j \Delta w_{a,d}(t_i, t_j))}}{\sum_k \text{Var}_a(\sum_i \sum_j \Delta w_{a,k}(t_i, t_j))} \quad (16)$$



being the coupling factor. Similarly,

$$c_I = 1 + \sum_k \sum_{m < n} \rho_{m,n;k} f_{m,n;k} \quad (17)$$

with $\rho_{m,n;k} = \text{Corr}_a(\sum_j \Delta w_{a,k}(t_m, t_j), \sum_j \Delta w_{a,k}(t_n, t_j))$ being the correlation coefficient between the synaptic changes contributed by the m th and n th spikes of the central neuron, which is particularly non-zero for two central spikes adjacent in time, and

$$f_{m,n;k} = \frac{2\sqrt{\text{Var}_a(\sum_j \Delta w_{a,k}(t_m, t_j)) \cdot \text{Var}_a(\sum_j \Delta w_{a,k}(t_n, t_j))}}{\sum_k \sum_i \text{Var}_a(\sum_j \Delta w_{a,k}(t_i, t_j))} \quad (18)$$

being the coupling factors.

Therefore, we can see that there are three factors that are important to understand the efficacy variability $\text{Var}_a(\Delta w_a)$ (Figures 3B–D):

- (1) The mean efficacy variability contributed by a single spike of the central neuron, which is represented by d .
- (2) The correlation between the synaptic changes caused by adjacent spikes of the central neuron, which contributes to $\text{Var}_a(\Delta w_a)$ through c_I .
- (3) The correlation between the total potentiation and depression value imposed on a synapse, which contributes to $\text{Var}_a(\Delta w_a)$ through c_{II} .

In the following discussions, we will generate spike patterns with different synchronous firing and auto-correlation structure with homogeneous firing rates and cross-correlations using spike-generating models, and investigate how the efficacy variability as well as d , c_I and c_{II} change with model parameters; then we

will discuss the physical pictures underlying these phenomena, aiming to giving the readers an intuitive understanding on how synchronous firing and temporal structure influences DiffV.

3.2. The Influence of Auto-correlation Structure onto DiffV in Stationary Spike Trains

We used Gamma processes (Model Auto in Section 2) to model the auto-correlation structure of stationary spike trains (whose firing rates do not change with time). We changed the shape parameter α of the Gamma processes while conserving the firing rates at $r_0 = 20$ Hz. The coefficient of variance (CV) of a Gamma process is $CV = 1/\sqrt{\alpha}$. When CV gets larger, spike trains are burstier; when CV gets smaller, spikes are more regular. We found that both burstiness and strong regularity in auto-correlation structure increased the efficacy variability, and the efficacy variability gets its minimal value when CV is around $0.3 \sim 0.7$ (Figure 4A), which is the range most neurons lie within *in vivo* (Softky and Koch, 1993).

Spike trains in this model are stationary processes, and we also let $A_p = A_d$ in our model, with A_p and A_d being the strengths of the exponentially decayed STDP windows for potentiation and depression (see Figure 2B). Therefore,

- (1) the variance caused by potentiation and depression processes should be equal. From Equation 16, this means that $f_{PD} = 1$, so that c_{II} can link with ρ_{PD} in a direct way:

$$c_{II} = 1 + \rho_{PD} \quad (19)$$

- (2) $\text{Var}_a(\sum_j \Delta w_{a,k}(t_i, t_j))$ does not change with the i index. From Equations 17 and 18, this means that c_I can link with

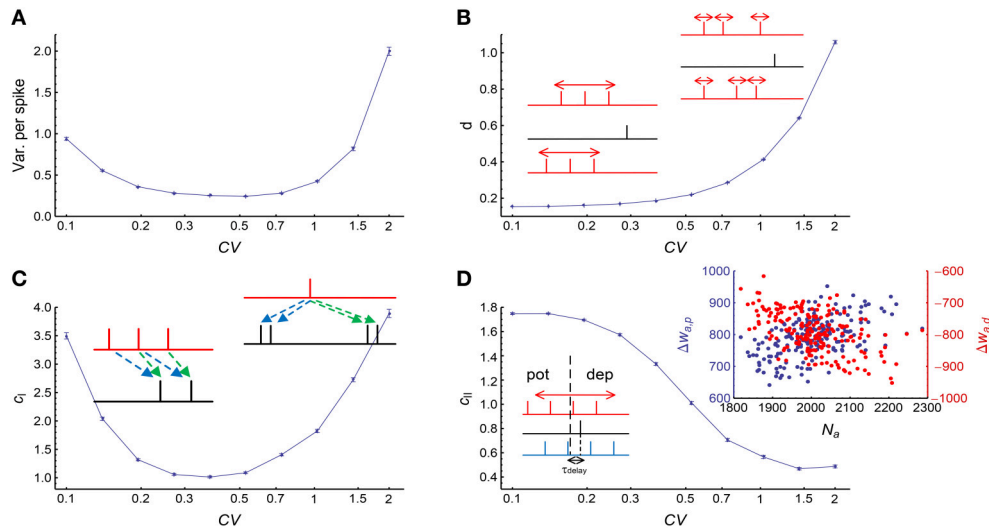


FIGURE 4 | The influence of auto-correlation structure onto DiffV in stationary spike trains. (A) Variance per spike ($E_T[\text{Var}_a(\Delta w_a)]/\bar{N}_0$, see Equation 10) as a function of CV, when spike trains are generated according to Model Auto (see Section 2). **(B)** d as a function of CV. Left inset: when the non-central spike trains are regular, the non-central spikes move left or right almost simultaneously. Right inset: when the non-central spike trains get bursty, each spike can move with a larger freedom, thereby contributing to the increase of DiffV. Black lines in the insets represent the central spike train, and colored (here, red) lines represent the non-central spike trains: the same color scheme is also used in the insets of the following panels. **(C)** c_I as a function of CV. Left inset: when the spike trains are regular, the synaptic updatings caused by adjacent central spikes are correlated because of transient cross-correlation. Right inset: when the central spike train gets bursty, the synaptic updatings caused by adjacent spikes in the same bursting events are correlated. Dashed arrows of the same color in the two insets represent similar synaptic updatings during STDP. **(D)** c_{II} as a function of CV. Left inset: when the non-central spike trains are regular, the spikes that potentiate and depress the synapses move left or right almost simultaneously, which correlates the total potentiation and depression values. In this inset, the red spike train cause weaker potentiation and stronger depression than the blue spike train. Right inset: when the spike trains are bursty (here, $CV = 2$), the spike number of a non-central neuron (horizontal coordinate) is positively correlated with the total potentiation (blue) value and negatively correlated with the total depression (red) value. In **(A–D)**, error bars represent s.e.m., the converging motif has 200 non-central neurons. Parameters for STDP: $A_d = A_p = 1$ (see Figure 2B and Equation 6 in Section 2). All synaptic efficacies were 0 at the beginning, and simulations were run for 100 s biological time, with 32 trials.

$\rho_{m,n;k}$ in a direct way:

$$c_I = 1 + \frac{\sum_k \sum_{m < n} \rho_{m,n;k}}{N_0} \quad (20)$$

with N_0 being the spike number of the central neuron.

Equations 19 and 20 mean that in this model the change of c_I and c_{II} with model parameters directly reflect the change of the correlations.

To understand Figure 4A, we investigated how d , c_I and c_{II} change with CV (Figures 4B–D). We found that d monotonically increases with CV (Figure 4B), c_I gets its minimum when CV is around 0.3 ~ 0.7, and gets large both when CV is too large or too small (Figure 4C), and c_{II} monotonically decreases with CV (Figure 4D). With the help of Equations 19 and 20, we can understand their changes with CV, and thus gain insight on the reason for the change of DiffV with CV.

3.2.1. Understanding the Change of d with CV

To understand the change of d with CV, consider the term $\text{Var}_a(\sum_j \Delta w_{a,p}(t_i, t_{j|a,i,p}))$ in Equation 13. When the spike trains of the non-central neurons are regular, $t_{j|a,i,p} \approx t_{1|a,i,p} - (j-1)/r_0$, with r_0 being the firing rate, so the difference of $\sum_j \Delta w_{a,p}(t_i, t_{j|a,i,p})$ for different a s mainly comes from the difference of $t_{1|a,i,p}$ (Figure 4B, left inset). What's more, $\{t_{1|a,i,p}\}_a$ approaches to a uniform distribution within the interval ($t_i -$

$\tau_{delay} - 1/r_0, t_i - \tau_{delay}$) when the spike trains get regular. When the spike trains get burstier, not only the distribution of $t_{1|a,i,p}$ gets broader, but the other spikes $t_{j \neq 1|a,i,p}$ can also move with a larger freedom, which increases the variability of $\sum_j \Delta w_{a,p}(t_i, t_{j|a,i,p})$ (Figure 4B, right inset). Similar arguments also apply to the $k = d$ case. It is easy to show (Supplementary Material Section S1.2) that under strict regularity ($CV = 0$),

$$d_{reg} = (A_d^2 + A_p^2) \frac{\tau_{STDP}}{2\Delta t} \left[\frac{1 + \exp(-\frac{\Delta t}{\tau_{STDP}})}{1 - \exp(-\frac{\Delta t}{\tau_{STDP}})} - \frac{2\tau_{STDP}}{\Delta t} \right], \quad (21)$$

with $\Delta t = 1/r_0$ being the inter-spike interval; and under Poisson process ($CV = 1$),

$$d_{Poi} = (A_d^2 + A_p^2) \frac{\tau_{STDP}}{2\Delta t}. \quad (22)$$

As $\frac{1 + \exp(-\frac{\Delta t}{\tau_{STDP}})}{1 - \exp(-\frac{\Delta t}{\tau_{STDP}})} - \frac{2\tau_{STDP}}{\Delta t} < 1$ (Supplementary Material Section S1.2), $d_{reg} < d_{Poi}$.

3.2.2. Understanding the Change of c_I with CV

c_I does not change monotonically with CV, it is large both when CV is large (i.e., spike trains are bursty) and when CV is very small (i.e., spike trains are very regular), and gets its minimal

value when CV is around 0.3 ~ 0.7 (**Figure 4B**). Now we try to understand the change of c_I with CV.

Suppose there are two spikes of the central neuron at time x and y . Because the spike trains in this model are stationary processes, the correlation between the potentiation (or depression) value caused by these two spikes only depends on $y - x$:

$$\rho_p(y - x) = \text{Corr}_a\left(\sum_{j \in (x, \infty)} \Delta w_{a,p}(x, j), \sum_{j \in (y, \infty)} \Delta w_{a,p}(y, j)\right) \quad (23)$$

$$\rho_d(y - x) = \text{Corr}_a\left(\sum_{j \in (-\infty, x)} \Delta w_{a,d}(x, j), \sum_{j \in (-\infty, y)} \Delta w_{a,d}(y, j)\right) \quad (24)$$

with $\Delta w_{a,k}(x, j)$ being the synaptic updating caused by the interaction of x and a spike of the a th non-central neuron at time j . This $\rho_k(l)$ function depends on the CV of the spike trains of the non-central neurons. Because of the time-reversal invariance of the statistics of the spike trains, $\rho_p(l) = \rho_d(l) \equiv \rho(l)$ in our model. Using $\rho(l)$, c_I (Equation 20) can be rewritten as

$$c_I = 1 + \frac{2}{r_0} \int_0^\infty C_0(l) \rho(l) dl \quad (25)$$

with r_0 being the firing rate of the central neuron, and $C_0(l)$ being the auto-correlation of the central neuron. $\frac{C_0(l)}{r_0}$ means the probability density to find another central spike at time l given an central spike at time 0, which can be written as

$$\frac{C_0(l)}{r_0} = \sum_{i=1}^\infty q_i(l) \quad (26)$$

with $q_i(l)$ being the distribution of the interval between the spike at time 0 and the next i th spike. Spike trains in our model are Gamma processes, so that the inter-spike interval follows $\Gamma(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$. It is easy to show that $q_i(l) = \Gamma(l|\alpha, \beta)$ (<http://mathworld.wolfram.com/GammaDistribution.html>).

From Equation 25, we know that the CV of the non-central spike trains influences c_I through $\rho(l)$, while the CV of the central spike train influences c_I through $q_0(l)$.

To understand the change of c_I with CV, we plot the functions $\rho(l)$ and $C_0(l)/r_0$ under different CV values (**Figure 5**). We find that:

- (1) When spike trains are bursty ($CV > 1$), $\rho(l)$ increases with CV, and $C_0(l)/r_0$ becomes more concentrated near zero (at which point $\rho(l)$ gets its maximal value) when CV increases. Both of these two factors increase the integration $\frac{1}{r_0} \int C_0(l) \rho(l) dl$ in Equation 25, contributing to the increase of c_I when spike trains are bursty.
- (2) When spike trains are very regular (CV is very small), both $\rho(l)$ and $C_0(l)/r_0$ become more concentrated near $l = N/r_0$ (with $r_0 = 20$ Hz being the firing rate and N being a positive integer): in this case, $\frac{1}{r_0} \int C_0(l) \rho(l) dl$ also increases because of the overlap of the peaks of $\rho(l)$ and $C_0(l)/r_0$.

Intuitively, the change of c_I with CV can be understood like this:

- (1) When CV is large, the non-central spike trains tend to be clustered into bursting events, and the time scale τ_{auto} of the bursting events increases with CV (it is easy to show that $\tau_{auto} \sim \frac{CV^2}{r_0}$ when $CV > 1$, see Supplementary Material Section S1.1): this strengthened and broadened auto-correlation makes $\rho(l)$ increase with CV when $CV > 1$. At the same time, more adjacent central spikes are also gathered closer by bursting events, which increases the correlation between the synaptic changes caused by adjacent central spikes (**Figure 4C**, right inset): this is the effect of $C_0(l)/r_0$.
- (2) When CV is very small, the spike trains are very regular. In this case, two adjacent spikes $\{t_i, t_{i+m}\}$ of the central neuron have the relation $t_{i+m} \approx t_i + m/r_0$ if they are well within the time scale τ_{auto} of the oscillating decaying auto-correlation (it is easy to show that $\tau_{auto} \sim \frac{1}{4CV^2 r_0}$ when $CV < 1$, see Supplementary Material Section S1.1); and two spikes $\{t_{j|a,i,k}, t_{j|a,i+m,k}\}$ of the a th non-central neuron also have the relation $t_{j|a,i+m,k} \approx t_{j|a,i,k} + m/r_0$. This means that $t_i - t_{j|a,i,k} \approx t_{i+m} - t_{j|a,i+m,k}$, which makes the synaptic changes caused by t_i and t_{i+m} almost the same (i.e., $\sum_j \Delta w_{a,k}(t_i, t_j) \approx \sum_j \Delta w_{a,k}(t_{i+m}, t_j)$) (**Figure 4C**, left inset). This increases the correlation $\rho_{i,i+m;k}$ caused by adjacent central spikes, thereby increasing c_I .

In this model (Model Auto in Section 2), we suppose that the firing rates of the non-central neurons r_* are the same as the firing rate of the central neuron r_0 . What if $r_* \neq r_0$? **Figure 5** suggests that when spike trains are very regular, $\rho(l)$ peaks at N/r_* (with N being a positive integer), and $C_0(l)/r_0$ peaks at N/r_0 when $CV \ll 1$. Equation 25 indicates that c_I gets its maximum value when the peaks of $\rho(l)$ overlap with the peaks of $C_0(l)/r_0$. This overlap is strong when $r_* = r_0$, but can also occur when $N_1 r_* = N_2 r_0$, with N_1 and N_2 being two positive integers with no common divisor larger than 1. For the simplicity of the following argument, we call the increase of the correlation $\rho_{m,n;k}$ caused by adjacent central spikes under strong regularity and $N_1 r_* = N_2 r_0$ to be the mechanism of *transient cross-correlation*. Because of the decaying oscillation of $\rho(l)$ and $C_0(l)/r_0$, N_1 and N_2 tend to be small integers.

3.2.3. Understanding the Change of c_{II} with CV

We find that $c_{II} > 1$ when spike trains are regular, which means that the potentiation and depression values are positively correlated ($\rho_{PD} > 0$) in this case; and $c_{II} < 1$ when spike trains are bursty, which means that the potentiation and depression values are negatively correlated ($\rho_{PD} < 0$). Now we try to understand the change of c_{II} with CV.

- (1) Under strong regularity, if $t_{j|a,i,k}$ are well within the time scale of τ_{auto} with $t_{1|a,i,p}$, then $t_{j|a,i,p} \approx t_{1|a,i,p} - (j - 1)/r_0$ and $t_{j|a,i,d} \approx t_{1|a,i,p} + j/r_0$. This means that if $t_{1|a,i,p}$ moves farther away from t_i , the potentiation value $\sum_j \Delta w_{a,p}(t_i, t_{j|a,i,p})$ will get weakened, while the depression value $\sum_j \Delta w_{a,d}(t_i, t_{j|a,i,d})$ will get strengthened (**Figure 4D**, left inset), which means that $\sum_j \Delta w_{a,p}(t_i, t_{j|a,i,p})$ and $\sum_j \Delta w_{a,d}(t_i, t_{j|a,i,d})$ positively correlated. This makes

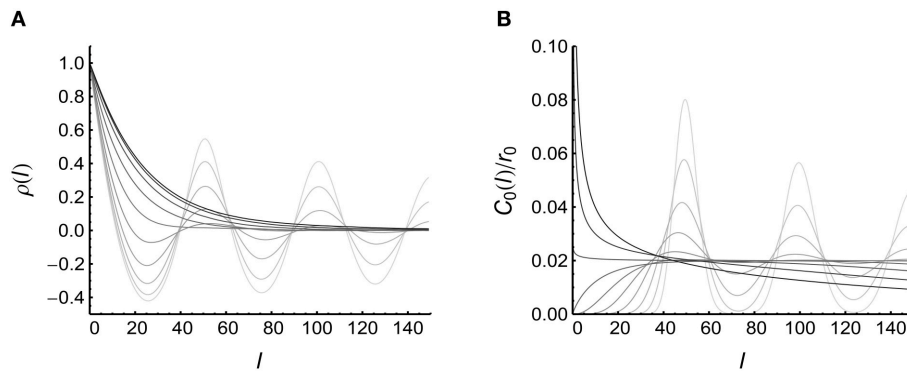


FIGURE 5 | Understanding the influence of auto-correlation structure onto c_I (see Equation 25). (A) $\rho(l)$ when $CV = 0.1, 0.139, 0.195, 0.271, 0.379, 0.528, 0.737, 1.03, 1.43, 2.00$ (represented by lines from light to dark). (B) $C_0(l)/r_0$ when CV takes values the same as in (A). In (A), the results are from 10,000 trials of simulations on a converging motif with 200 non-central neurons. (B) is plotted using Equation 26.

$\sum_i \sum_j \Delta w_{a,p}(t_i, t_{j|a,i,p})$ and $\sum_i \sum_j \Delta w_{a,d}(t_i, t_{j|a,i,d})$ are also positively correlated. By the definition of ρ_{PD} , we know that $\rho_{PD} > 0$ in this case.

- (2) To understand the case when CV is large, note that the Fano factor of a long time scale approaches to CV^2 (Cox, 1962; Tuckwell, 1988; Nawrot et al., 2008), which means that different non-central neurons may emit quite different numbers of spikes during simulation when CV is large. If the a th non-central neuron fires more (less) spikes, then both the potentiation and depression values imposed on the a th synapses tend to be strong (weak) (Figure 4D, right inset). This is the reason why ρ_{PD} is negative when spike trains are bursty.

Comparing Figure 4A with Figures 4B–D, we know that both d and c_I are the reasons for the steep increase of DiffV under large CV , while c_I (i.e., the mechanism of transient cross-correlation) is the reason for the fast increase of DiffV with the decrease of CV under small CV . c_{II} , however, does not significantly contribute to the change of DiffV under large or small CV .

3.2.4. The Interaction of Auto-correlation Structure With Heterogeneity of Rates

Till now, we only considered the case when the spike trains are Gamma processes with the same firing rate. In Supplementary Material Section S2, we will consider the case that the Gamma processes of different non-central neurons have different firing rates, while their mean firing rate is still the same as the firing rate of the central neuron r_0 . Our simulations suggest that heterogeneity of rates does not strongly influence DiffV (as different synapses have different diffusion strengths under heterogeneity of rates, DiffV here means the mean diffusion strength over all the synapses) when the spike trains are bursty, but discounts the increase of DiffV with the decrease of CV under strong regularity because of it destroying transient cross-correlation (Supplementary Figure 1). This makes DiffV tend to monotonically increase with CV .

3.3. The Influence of Synchronous Firing onto DiffV

There are usually two scenarios regarding to synchronous firing: one is spike synchrony, which means that the spikes of neurons are emitted almost simultaneously; the other one is rate synchrony, which means that the firing rates of neurons rise and fall at the same time. The key difference between these two scenarios is the distribution of the spike number of a neuron within a synchronous event. For spike synchrony, a neuron can fire no more than a single spike in a synchronous event, so that this distribution is narrow; but for rate synchrony, different neurons can fire a different number of spikes in a synchronous event, so that this distribution is broad. To understand how DiffV is influenced by synchronous firing under different distributions of spike numbers per neuron per synchronous event, we generated spike trains using Model Sync (Section 2). In this model, a synchronous event lasts for τ_{cross} . The spike number per neuron per synchronous event is distributed the same as the spike number of an underlying Gamma process within an interval of length τ_{cross} . The rate of this Gamma process is p/τ_{cross} and its coefficient of variance is $CV_{SpikeNum}$. So the mean value of the distribution of spike number per neuron per synchronous event is p , and this distribution is narrow if $CV_{SpikeNum}$ is small, and gets broadened when $CV_{SpikeNum}$ increases. If a neuron is to fire M spikes in a synchronous event, then the spike times of these M spikes will be randomly and uniformly chosen within this time interval of duration τ_{cross} . The occurrence of synchronous events is a Poisson process with rate r_0/p , so that the firing rate of a neuron is kept at $r_0 = 20$ Hz when p changes. We focus on the case that $\tau_{cross} \ll \tau_{STDP}$ and $\tau_{cross} \ll p/r_0$ (with p/r_0 being the average inter-event interval) in our study. It is easy to think that if $\tau_{cross} \ll p/r_0$ but $\tau_{cross} > \tau_{STDP}$, STDP updates will tend to zero so that DiffV will decrease with τ_{cross} ; and if $\tau_{cross} > p/r_0$, the spike pattern will approach to asynchronous state.

For spike patterns with synchronous firing, we don't have simple equations like Equations 19 and 20 to directly link $\rho_{m,n;k}$ and ρ_{PD} with c_I and c_{II} . The coupling factors f_{PD} (Equation 16)

and $f_{m,n;k}$ (Equation 18) must be considered. As

$$\sqrt{\text{Var}_a\left(\sum_j \Delta w_{a,k}(t_m, t_j)\right) \cdot \text{Var}_a\left(\sum_j \Delta w_{a,k}(t_n, t_j)\right)} \\ \leq \frac{1}{2} \left(\text{Var}_a\left(\sum_j \Delta w_{a,k}(t_m, t_j)\right) + \text{Var}_a\left(\sum_j \Delta w_{a,k}(t_n, t_j)\right) \right)$$

and the equality is obtained only when $\text{Var}_a(\sum_j \Delta w_{a,k}(t_m, t_j)) = \text{Var}_a(\sum_j \Delta w_{a,k}(t_n, t_j))$, $f_{m,n;k}$ reflects the inhomogeneity of $\text{Var}_a(\sum_j \Delta w_{a,k}(t_m, t_j))$ for different m . And similarly, f_{PD} reflects the difference between $\text{Var}_a(\sum_i \sum_j \Delta w_{a,p}(t_i, t_j))$ and $\text{Var}_a(\sum_i \sum_j \Delta w_{a,d}(t_i, t_j))$ relative to their total values.

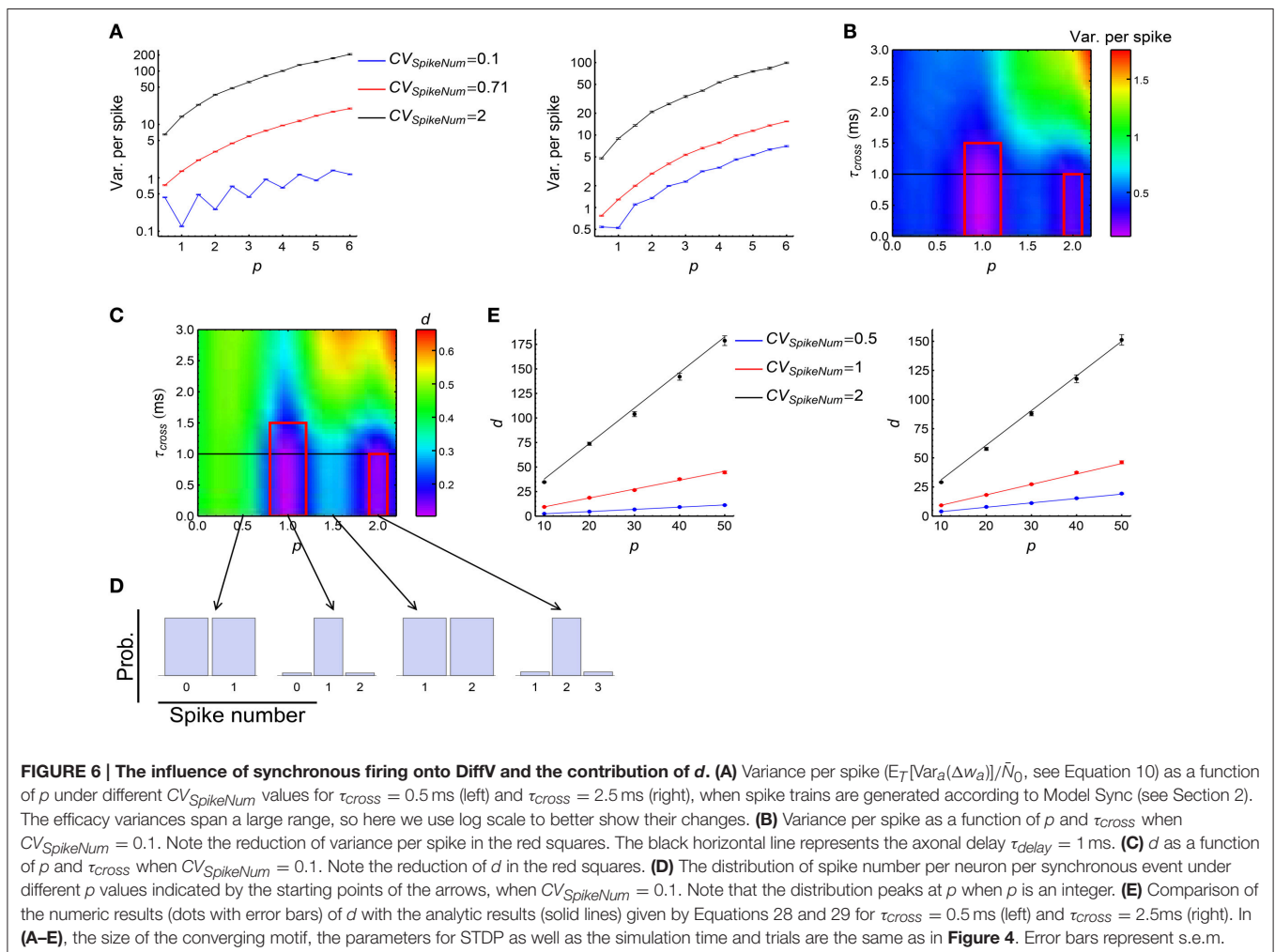
We investigated how DiffV changes with p , τ_{cross} and the $CV_{SpikeNum}$. Our simulations suggest:

- (1) DiffV tends to increase with the strength of synchronous firing (i.e., p) and the broadness of the distribution of the spike number per neuron per synchronous event (i.e., $CV_{SpikeNum}$) (Figure 6A, see Supplementary Figures 2A1,B1,C1 for more information);

- (2) when $CV_{SpikeNum}$ is small enough, DiffV can be significantly decreased if p is close to an integer and $\tau_{cross} < \tau_{delay}$ (the red squares in Figure 6B).

Here we explain the underlying physical pictures of these phenomena based on the observation of the change of d , c_I and c_{II} with model parameters:

- (1) d is the only reason for the second phenomenon above (the red squares in Figure 6C, compare with the changes of c_I and c_{II} in Supplementary Figures 2A3,A4). To understand this, suppose a synchronous event happening during $[t_1, t_2]$ ($t_2 = t_1 + \tau_{cross}$), then the central neuron will receive its afferents during $[t_1 + \tau_{delay}, t_2 + \tau_{delay}]$, with τ_{delay} being the axonal delay. If $\tau_{cross} \leq \tau_{delay}$, the spike time t_0 of the central neuron $t_0 < t_1 + \tau_{delay}$, which means that the central neuron always starts to receive spikes after its own firing in a synchronous event, so that all the spikes it receives depress the corresponding synapses after the synchronous event. If $\tau_{cross} \leq \tau_{delay} \ll \tau_{STDP}$, the depression value Δw contributed by each non-central spike is approximately the same. If p is an integer and $CV_{SpikeNum}$ is small enough, the distribution of the spike number per neuron per synchronous event will



sharply peak at p (Figure 6D). In this case, nearly all the non-central neurons fire exactly p spikes during a synchronous event, so that almost all the synapses are depressed by $p\Delta w$ after this synchronous event. This is the reason why d is so small in this case. Mathematically, it can be shown (Supplementary Materials Section S3.1) that when $\tau_{cross} \leq \tau_{delay} \ll \tau_{STDP}$,

$$d \approx \text{Var}_a(N_a)[A_d^2 \exp(-\frac{2\tau_{delay}}{\tau_{STDP}}) + \frac{r_0\tau_{STDP}}{2p}(A_d^2 + A_p^2)] \quad (27)$$

with E_T represent average over trials, and $\text{Var}_a(N_a)$ being the variance of spike numbers of the non-central neurons in a synchronous event. We can see that d gets reduced if $\text{Var}_a(N_a)$ is small, which is the case when $CV_{SpikeNum}$ is small and p is an integer. When $\tau_{cross} > \tau_{delay}$, however, it is possible that $t_1 + \tau_{delay} < t_0 < t_2 + \tau_{delay}$, so that all the in-coming spikes of the central neuron during $[t_1 + \tau_{delay}, t_0)$ potentiate the corresponding synapses, and all the in-coming spikes during $(t_0, t_2 + \tau_{delay}]$ depress the corresponding synapses. This may split the synapses into different directions, thereby increasing d (note the large d value when $\tau_{cross} > \tau_{delay}$ in Figure 6C).

- (2) d tends to increase with p or $CV_{SpikeNum}$ (Figure 6E, see Supplementary Figures 2A2,B2,C2 for more information), thereby contributing to the tendency of the increase of DiffV with p or $CV_{SpikeNum}$. In this paper, we generally suppose that $\tau_{cross} \ll \tau_{STDP}$ and $\tau_{delay} \ll \tau_{STDP}$. When p large, it can be shown (Supplementary Materials Section S3.1) that if $\tau_{cross} \leq \tau_{delay}$, then

$$d = E_T[\text{Var}_a(\sum_j \Delta w_{a,d}(t_i, t_j))] + E_T[\text{Var}_a(\sum_j \Delta w_{a,p}(t_i, t_j))] \\ \approx pCV_{SpikeNum}^2 A_d^2 \exp(-\frac{2\tau_{delay}}{\tau_{STDP}}) + CV_{SpikeNum}^2 \frac{r_0\tau_{STDP}}{2}(A_d^2 + A_p^2), \quad (28)$$

with E_T representing trial average; and if $\tau_{cross} > \tau_{delay}$, then

$$d \approx p[CV_{SpikeNum}^2 A_d^2 \frac{\tau_{delay}}{\tau_{cross}} \exp(-\frac{\tau_{cross} + \tau_{delay}}{\tau_{STDP}}) + A_d^2 \exp(-\frac{\tau_{cross} + \tau_{delay}}{2\tau_{STDP}}) \cdot \mathcal{A} + A_p^2 \exp(-\frac{\tau_{cross} - \tau_{delay}}{2\tau_{STDP}}) \cdot \mathcal{B}] + CV_{SpikeNum}^2 \frac{r_0\tau_{STDP}}{2}(A_d^2 + A_p^2) \quad (29)$$

with

$$\mathcal{A} = (\frac{1}{6} + \frac{1}{3}CV_{SpikeNum}^2) - \frac{1}{2}(\frac{\tau_{delay}}{\tau_{cross}})^2 - \frac{1}{3}(CV_{SpikeNum}^2 - 1)(\frac{\tau_{delay}}{\tau_{cross}})^3$$

$$\mathcal{B} = (1 - \frac{\tau_{delay}}{\tau_{cross}})^2(\frac{1}{2} + \frac{1}{3}(CV_{SpikeNum}^2 - 1)(1 - \frac{\tau_{delay}}{\tau_{cross}})).$$

We compare the analytic results with numeric results in Figure 6E, we can see the tendency that d increases with both

p and $CV_{SpikeNum}$. The first terms of Equations 28 and 29 represent the contribution of the synchronous event S_0 that the central spike t_i belongs to, and the last terms represent the contributions of the other synchronous events. p only exists in the first terms of Equations 28 and 29, which suggests that p increases d through increasing the synchrony strength of S_0 . $CV_{SpikeNum}$ exists in both terms, which suggests that the broadness of the spike number distribution in both S_0 and the other synchronous events contribute to d .

- (3) c_I tends to increase with both p and $CV_{SpikeNum}$ (Figure 7A, see more in Supplementary Figures 2A3, B3, C3), thereby contributing to the tendency of the increase of DiffV with p or $CV_{SpikeNum}$. From Equation 17, we know that the change of c_I may either caused by the correlation $\rho_{m,n;k}$ or by the coupling factor $f_{m,n;k}$. To estimate their contributions to c_I , we define

$$\rho_{l,k} = \frac{\sum_m \rho_{m,(m+l);k}}{\sum_m 1}, \quad l = 1, 2, \dots; k = p, d, \quad (30)$$

and

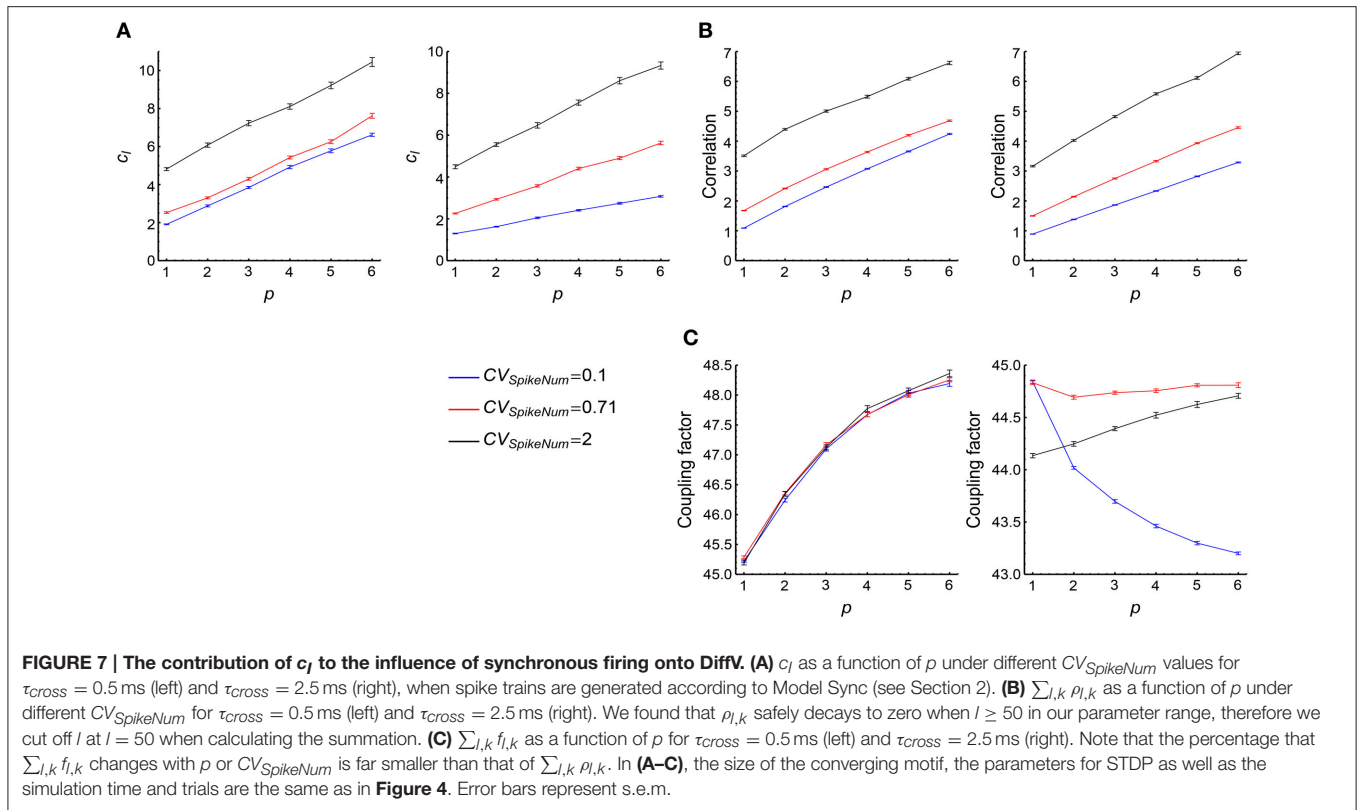
$$f_{l,k} = \sum_m f_{m,(m+l);k} \quad l = 1, 2, \dots; k = p, d. \quad (31)$$

We find that the change of $\sum_k \sum_l f_{l,k}$ with model parameters is far slower than the change of $\sum_k \sum_l \rho_{l,k}$ (Figures 7B,C), which suggests that the coupling factors $f_{m,n;k}$ do not significantly contribute to the change of c_I , and the correlation $\rho_{m,n;k}$ between synaptic updatings caused by adjacent central spikes is the main reason for the change of c_I . It is easy to think that if two spikes of the central neuron are in the same synchronous event, then the correlation of the synaptic updatings caused by them tends to be large; but if they are in different synchronous events, then the correlation of the synaptic changes caused by them tends to be small. To understand the increase of c_I with p and $CV_{SpikeNum}$, let us consider the following toy model. Suppose that $\rho_{mn} = a_0$ if m and n are in the same synchronous event (this assumption is particularly correct when $\tau_{cross} < \tau_{delay}$), and $\rho_{mn} = 0$ if they are in different synchronous events. Suppose that the central neuron fires N_s spikes in the s th synchronous event, then the total value of the correlation coefficients will be

$$\sum_{m < n} \rho_{mn} = a_0 \sum_s \frac{N_s(N_s - 1)}{2} = a_0 M \int_{-\infty}^{\infty} q(N_s) \frac{N_s(N_s - 1)}{2} dN_s$$

with M being the number of synchronous events, and $q(N_s)$ being the distribution of the spike number per synchronous event. If p is sufficiently large, $q(N_s) \sim \mathcal{N}(p, \sigma^2)$ with $\sigma^2 = p \cdot CV_{SpikeNum}^2$. In this case

$$\sum_{m < n} \rho_{mn} \approx \frac{a_0 M p(-1 + p + CV_{SpikeNum}^2)}{2} = \frac{a_0 N_0(-1 + p + CV_{SpikeNum}^2)}{2} \quad (32)$$



with N_0 being the total spike number of the central neuron. From Equation 32, we can see that $\sum_{m < n} \rho_{mn}$ tend to increase with both p and $CV_{SpikeNum}$, which explains why c_I tend to increase with p and $CV_{SpikeNum}$. Intuitively, p gathers more spikes into the same synchronous event; and $CV_{SpikeNum}$, which induces heterogeneity of spike numbers in a synchronous event, concentrates spikes into fewer synchronous events. Both of them increase the number of spike pairs that lie in the same synchronous event, thereby increasing $\sum_{m < n} \rho_{mn}$.

- (4) We found that when p is large, the change of c_{II} with p is small (Supplementary Figures 2A4, B4, C4), which suggests that the contribution of c_{II} to the increase of DiffV with p is small compared to the other factors. The change of c_{II} with $CV_{SpikeNum}$ is not strong when $\tau_{cross} \leq \tau_{delay}$; and c_{II} tends to decrease with $CV_{SpikeNum}$ when $\tau_{cross} > \tau_{delay}$ (Supplementary Figures 2A4, B4, C4), which negatively contributes to the increase of DiffV with $CV_{SpikeNum}$. These suggest that c_{II} is not an important factor to understand the change of DiffV under synchronous firing. We move more discussions on c_{II} into Supplementary Materials Section S3.2.

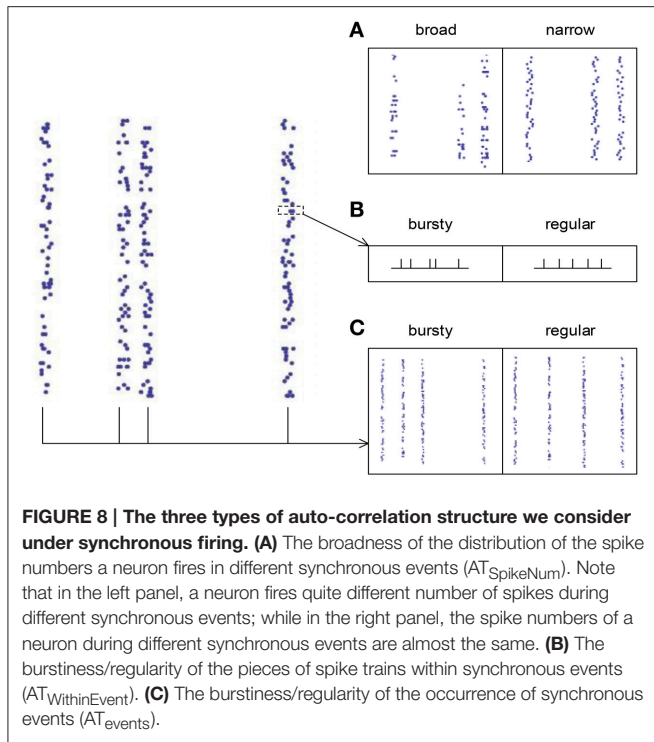
From the discussion above, we know that both the mean (controlled by p) and variance (controlled by $CV_{SpikeNum}$) of the spike number per neuron per synchronous event are important factors to understand the change of DiffV under synchronous firing. They contribute to DiffV mainly through d and c_I .

3.4. The Influence of the Interaction of Auto-correlation Structure And Synchronous Firing onto DiffV

The spike pattern of a real neuronal population possesses both synchronous firing and auto-correlation structure, so it is desirable to know how these two pattern structures interact to influence DiffV. This is a complicated problem, because auto-correlation structure comes into spike patterns with synchronous firing in at least three ways (**Figure 8**):

- (1) The broadness of the distribution of the spike numbers a neuron fires in different synchronous events ($AT_{SpikeNum}$). If the distribution gets broad, a neuron may burst several spikes in some synchronous events, while keep silent in some others.
- (2) The burstiness/regularity of the pieces of spike trains within synchronous events ($AT_{WithinEvent}$).
- (3) The burstiness/regularity of the occurrence of synchronous events (AT_{events}).

There may be other auto-correlation structures under synchronous firing. For example, in biological systems, the amplitudes of synchronous events may exhibit strong variability (Gireesh and Pleniz, 2008; Petermann et al., 2009), which further increases the complexity of the problem. However, we argue that this varying-amplitude situation may be included into the constant-amplitude scenario using AT_{events} , after noting that strong synchronous events can be regarded as the burstiness of many small synchronous events. For simplicity, we will



not consider this varying-amplitude situation in our following discussions.

We have already discussed the effect of AT_{SpikeNum} in the previous subsection, and we will perform detailed discussions on how $AT_{\text{WithinEvent}}$ and AT_{events} influence DiffV through d , c_I and c_{II} in Supplementary Material Section S4. Here we summarize our main results:

- (1) The broadness of the distribution of the spike numbers a neuron fires in different synchronous events increases DiffV through increasing d and c_I . When $\tau_{\text{cross}} \leq \tau_{\text{delay}}$, it does not significantly change c_{II} ; when $\tau_{\text{cross}} > \tau_{\text{delay}}$, it decreases c_{II} , which negatively contributes to the increase of DiffV (see the previous subsection).
- (2) The burstiness of the piece of spike train within a synchronous event significantly increases DiffV when $\tau_{\text{cross}} > \tau_{\text{delay}}$. It does this through increasing d and c_{II} , it does not significantly influence c_I (Supplementary Material Section S4.1, Supplementary Figure 3).
- (3) The burstiness of the occurrence of synchronous events increases DiffV through increasing d and c_I . It decreases c_{II} , which negatively contributes to the increase of DiffV (Supplementary Material Section S4.2, Supplementary Figure 7).

In Supplementary Material Section S4 and Supplementary Figures 3–10, we explain the underlying mechanisms of these influences using both numeric and analytic approaches. We find that broader distribution of spike number per neuron per synchronous event, burstier spike trains within synchronous events, and burstier occurrence of synchronous events tend to increase DiffV. These results can be concluded

into a rule of thumb: the burstiness of spike trains tends to increase DiffV, while the regularity tends to decrease DiffV.

From Section 3.2, we know that when the spike trains are stationary processes, DiffV increases with the decrease of CV if $CV \ll 1$ (Figure 4A), because of transient cross-correlation. However, here we find that after synchronous firing is added into the spike pattern, DiffV usually monotonically decreases with the decrease of CV. To understand this, note that in our model, we suppose that the pieces of spike trains that belong to different synchronous events are independent of each other. In this case, transient cross-correlation may be fragile under synchronous firing, because the synaptic changes caused by two central spikes t_m and t_n (i.e., $\sum_j \Delta w_a(t_m, t_j)$ and $\sum_j \Delta w_a(t_n, t_j)$) are hard to be correlated if t_m and t_n belong to different synchronous events, even if m and n are nearby by index.

3.5. The Influence of Heterogeneity of Rates onto DriftV

Under heterogeneity of rates or heterogeneity of cross-correlations, different synapses may drift with different velocities, thereby inducing DriftV. Heterogeneity of rates and heterogeneity of cross-correlations may also influence the diffusion of synapses. However, as $\text{DriftV} \propto t^2$ and $\text{DiffV} \propto t$, DriftV will dominate in a long run. Therefore, we will focus on their influence onto DriftV in the main text, their influence on DiffV will be briefly discussed in the Supplementary Materials Section S6.

Suppose that the activities of the central and non-central neurons are time-modulated simultaneously, so that the firing rate of the central neuron $r_0(t) = \langle r_0(t) \rangle x(t)$ while the firing rate of the a th non-central neuron $r_a(t) = \langle r_a(t) \rangle x(t)$, with $\langle \cdot \rangle$ representing time average, and $x(t)$ representing a function with $\langle x(t) \rangle = 1$. Then, the change of the a th synapse per unit time is

$$\begin{aligned} v_a &= \frac{dE_T(\Delta w_a)}{dt} = \int_0^\infty d\tau (H(\tau)r_0(t)r_a(t - \tau_{\text{delay}} - \tau) \\ &\quad + H(-\tau)r_0(t - \tau)r_a(t - \tau_{\text{delay}})) \\ &= \langle r_0(t) \rangle \langle r_a(t) \rangle \int_{-\infty}^\infty d\tau H(\tau)C(-\tau_{\text{delay}} - \tau) \end{aligned} \quad (33)$$

with $H(\tau)$ being the STDP time window and $C(-\tau_{\text{delay}} - \tau) = \langle x(t)x(t - \tau_{\text{delay}} - \tau) \rangle$ being the auto-correlation of $x(t)$. If we average the equation above over a , we will have

$$E_a(v_a) = \langle r_0(t) \rangle E_a(\langle r_a(t) \rangle) \cdot \int_{-\infty}^\infty d\tau H(\tau)C(-\tau_{\text{delay}} - \tau) \quad (34)$$

The two equations above together give

$$v_a = E_a(v_a) \cdot \frac{\langle r_a(t) \rangle}{E_a(\langle r_a(t) \rangle)} \quad (35)$$

Therefore,

$$\text{Var}_a(v_a) = \frac{[E_a(v_a)]^2}{[E_a(\langle r_a(t) \rangle)]^2} \cdot \text{Var}_a(\langle r_a(t) \rangle) \quad (36)$$

This equation suggests that the heterogeneity of rates $\langle r_a(t) \rangle$ can induce DriftV; but if the population rate $E_a(\langle r_a(t) \rangle)$ is kept constant, DriftV also depends on $|E_a(v_a)|$, which quantifies the imbalance of the strengths of the potentiation and depression process (*P-D imbalance*). If $E_a(v_a) = 0$, which means that STDP cannot change the mean synaptic strength (*P-D balance*), then $\text{DriftV} = 0$; if $|E_a(v_a)|$ is large, which means that P-D imbalance is strong, then DriftV is also large. From Equation 34, we know that if the firing rates $\langle r_0(t) \rangle$ and $E_a(\langle r_a(t) \rangle)$ are kept constant, then P-D imbalance is controlled by the time-modulated pattern $x(t)$, which will be changed if synchronous firing is added into the spike pattern. Intuitively, during a synchronous event, the axonal delay in the converging motif tends to make the central neuron receive spikes from the non-central neurons *after* its own spikes, which depresses the synapses under STDP (**Figure 9A**, a similar phenomenon has been observed in Lubenov and Siapas, 2008). Under heterogeneity of rates, this makes DriftV become accordingly changed (**Figure 9B**). In summary, heterogeneity of rates makes use of P-D imbalance to induce DriftV; and synchronous firing can influence P-D imbalance, thereby influencing DriftV under heterogeneity of rates.

3.6. The Interaction of Heterogeneity of Rates, Synchronous Firing And Auto-correlation Structure

From the arguments above, we know that if spike trains are stationary processes (with $x(t) = 1$ for any t), then auto-correlation structure cannot influence DriftV under heterogeneity of rates (because auto-correlation structure cannot change $x(t)$). Its influences onto DiffV in this case are discussed in Supplementary Materials Section S3. Interesting things happen when synchronous firing also comes into the picture. As we mentioned in Section 3.4, under synchronous firing, auto-correlation structure may come into the spike pattern in three ways (**Figure 8**):

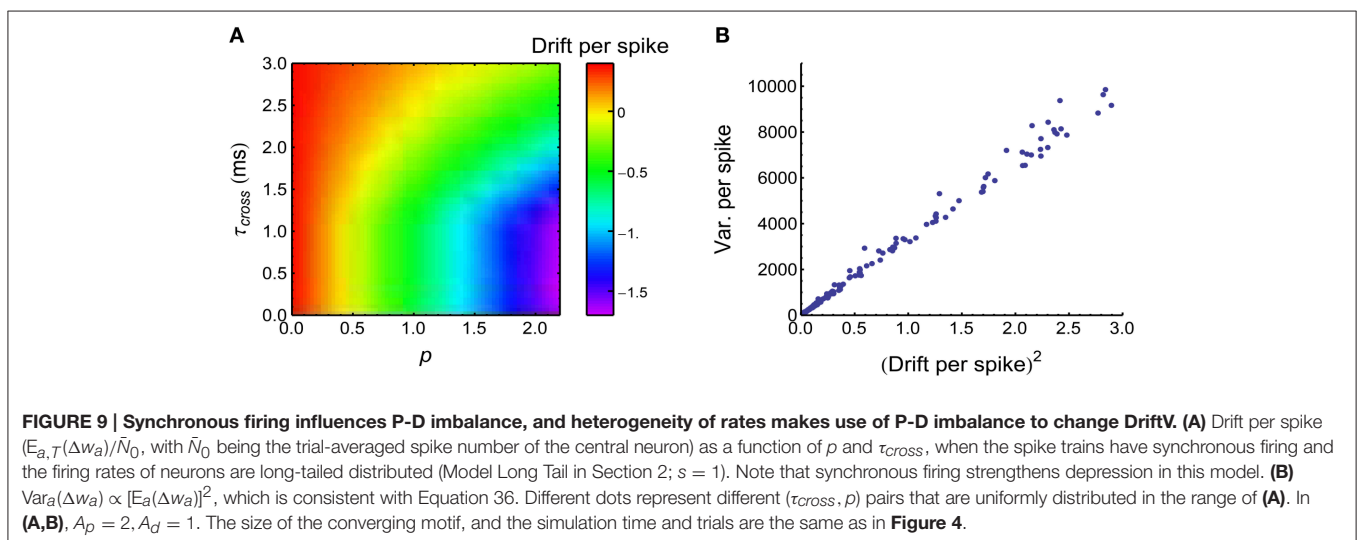
- (1) The broadness of the distribution of spike number per neuron per synchronous event (AT_{SpikeNum}).
- (2) The burstiness/regularity of the pieces of spike trains within synchronous events ($AT_{\text{WithinEvent}}$).
- (3) The burstiness/regularity of the occurrence of synchronous events (AT_{events}).

These three auto-correlation structures can be classified into two kinds: AT_{SpikeNum} and $AT_{\text{WithinEvent}}$ do not change $x(t)$, so that they do not change P-D imbalance, nor DriftV under heterogeneity of rates; AT_{events} , however, changes $x(t)$, thereby changing P-D imbalance, and DriftV under heterogeneity of rates. An idea to separate these two kinds of auto-correlation structure is to define rescaled time using the accumulative function of population firing rate (Pillow, 2009)

$$\Lambda(t) = \int_0^t r(s) ds, \quad (37)$$

which stretches the inter-spike intervals in proportion to the firing rate (**Figure 10**). Auto-correlation structure then comes into the picture in two ways: the auto-correlation structure of the spikes in the rescaled time, which can be quantified by the CV value CV_{rescale} , and the auto-correlation structure of the occurrence of synchronous events, which can be quantified by CV_{events} . Apparently, CV_{rescale} has no influence on $x(t)$, thereby nor DriftV under heterogeneity of rates; but CV_{events} influences $x(t)$, thereby DriftV under heterogeneity of rates. Combining the discussion on the influence of temporal structure onto DiffV in Section 3.4, we come to the conclusion that under synchronous firing and heterogeneity of rates,

- (1) The temporal factors that increase either CV_{rescale} or CV_{events} increase DiffV.
- (2) The temporal factors that influence CV_{rescale} do not change P-D imbalance, thereby nor DriftV under heterogeneity of rates.



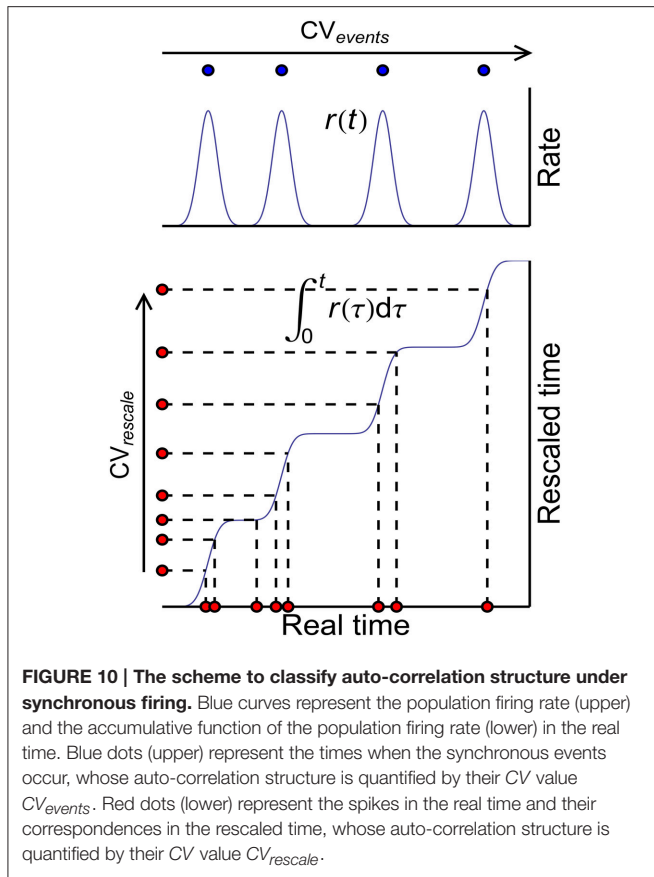


FIGURE 10 | The scheme to classify auto-correlation structure under synchronous firing. Blue curves represent the population firing rate (upper) and the accumulative function of the population firing rate (lower) in the real time. Blue dots (upper) represent the times when the synchronous events occur, whose auto-correlation structure is quantified by their CV value CV_{events} . Red dots (lower) represent the spikes in the real time and their correspondences in the rescaled time, whose auto-correlation structure is quantified by their CV value $CV_{rescale}$.

- (3) The temporal factors that influences CV_{events} change P-D imbalance, thereby DriftV under heterogeneity of rates.

Indeed, we find that both $AT_{SpikeNum}$ and $AT_{WithinEvent}$ increase $CV_{rescale}$ and don't influence P-D imbalance (Supplementary Figures 11A,B); while AT_{events} increases CV_{events} and may change P-D imbalance especially when $CV_{events} > 1$ (Supplementary Figures 11C,D). The influence of CV_{events} onto P-D imbalance seems complicated, and depends on the values of τ_{delay} , A_p and A_d (see Supplementary Figure 12, the physical mechanisms will be explained in Supplementary Material Section S5.2):

- (1) Suppose during a synchronous event S_0 , the central neuron fires at time t_0 . Because of the axonal delay τ_{delay} , there is usually a time interval between t_0 and when the spikes from non-central neurons arrive at the axonal terminal, and the typical length of this interval is τ_{delay} . If synchronous events are *not* close to each other, so that no non-central spikes from synchronous events *other than* S_0 arrive at the central neuron during this interval (i.e., different synchronous events do not *overlap* with each other), then $E_{a,T}(\Delta w_a)$ will increase (or decrease) with CV_{events} if $A_p \exp(\tau_{delay}/\tau_{STDP}) > A_d \exp(-\tau_{delay}/\tau_{STDP})$ (or $A_p \exp(\tau_{delay}/\tau_{STDP}) < A_d \exp(-\tau_{delay}/\tau_{STDP})$). In this paper, typically $\tau_{delay} \ll \tau_{STDP}$, so these conditions become $A_p > A_d$ or $A_p < A_d$.

- (2) If synchronous events are allowed to overlap with each other, then CV_{events} will increase the chance of this overlapping when it is large (typically when $CV_{events} > 1$). In this case, $E_{a,T}(\Delta w_a)$ will decrease (or increase) with CV_{events} if $\tau_{delay} > 0$ (or $\tau_{delay} < 0$). See the meaning of $\tau_{delay} < 0$ in **Figure 2B**.

To get a better understanding on how heterogeneity of rates, synchronous firing and auto-correlation structure interact with each other, we used Model Sync-Auto-LongTail (see Section 2) to generate spikes. In this model, for simplicity, the auto-correlation structure in the rescaled time and the auto-correlation structure of the occurrence of the synchronous events are respectively controlled by a single parameter (i.e., $CV_{rescale}$ and CV_{events}). We generated spike patterns using this model, and investigated how the efficacy variability changes with model parameters (**Figure 11**). The results validate our arguments above:

- (1) $CV_{rescale}$ hardly influences P-D imbalance (**Figure 11A**, lower panel).
- (2) When the strength of firing events is adjusted so that the potentiation and depression almost balance each other (so that $DriftV \approx 0$), the efficacy variability significantly increases with $CV_{rescale}$ (**Figure 11A**, upper panel). However, when the potentiation and depression are imbalanced, the efficacy variability becomes hardly influenced by $CV_{rescale}$ (**Figure 11A**, upper panel). This is because that $CV_{rescale}$ can only increase DiffV, and can hardly influence DriftV. As $DiffV \propto t$ while $DriftV \propto t^2$, the contribution of $CV_{rescale}$ to efficacy variability after a long run is only obvious when $DriftV \approx 0$, which is realized at P-D balance.
- (3) The influence of CV_{events} onto efficacy variability is non-monotonically complicated (**Figure 11B**, upper panel). The reason is that CV_{events} not only influences DiffV, but also influences DriftV in a complicated way (**Figure 11B**, lower panel), and DriftV is also influenced by p . Therefore, potentiation and depression are balanced at different p values for different CV_{events} values.

3.7. The Influence of Heterogeneity of Cross-correlations onto DriftV

Heterogeneity of cross-correlations mainly influences efficacy variability in DriftV manner. Suppose the firing rate of the central neuron is $r_0(t) = r_0 \cdot x_0(t)$ and the firing rate of the a th non-central neuron is $r_a(t) = r_a \cdot x_a(t)$, with r_0 and r_a being their time-averaged firing rates and $\langle x_0(t) \rangle = \langle x_a(t) \rangle = 1$ (here $\langle \cdot \rangle$ represent time average), then similarly as Equation 33, we have

$$v_a = r_0 r_a \int_{-\infty}^{\infty} d\tau H(\tau) C_a(-\tau_{delay} - \tau),$$

with $C_a(\tau) = \langle x_0(t)x_a(t+\tau) \rangle$ being the *unit cross-correlation* between the central neuron and the a th non-central neuron. Therefore,

$$\text{Var}_a(v_a) = r_0^2 \text{Var}_a[r_a \int_{-\infty}^{\infty} d\tau H(\tau) C_a(-\tau_{delay} - \tau)], \quad (38)$$

so the heterogeneity of $C_a(\tau)$ influences DriftV together with the heterogeneity of r_a .

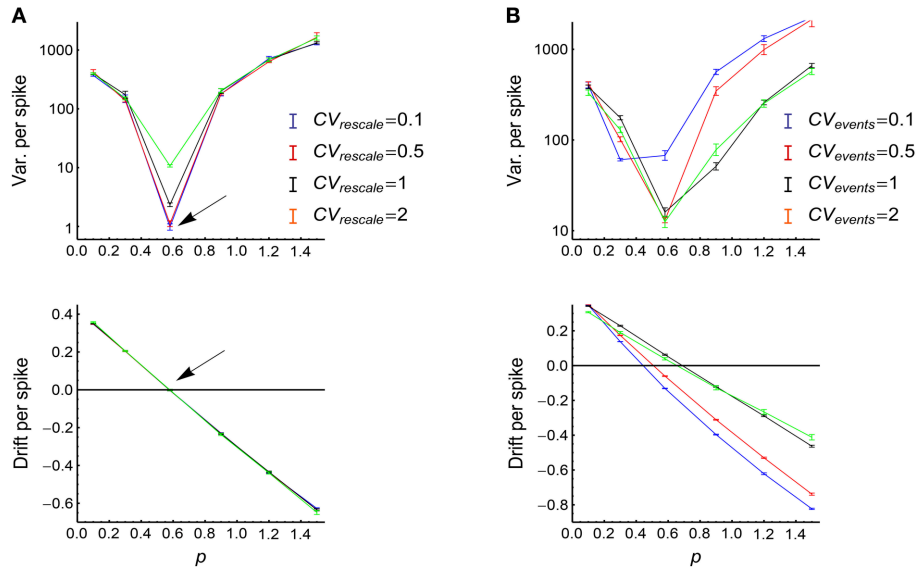


FIGURE 11 | The influence of auto-correlation structure onto the efficacy variability under synchronous firing and heterogeneity of rates. (A) Upper: Variance per spike as a function of p at different $CV_{rescale}$, keeping $CV_{events} = 0.71$ and $s = 1$ (see Model Sync-Auto-LongTail in Section 2). Lower: the corresponding mean efficacy changes per spike, which represent P-D imbalance. The arrows indicate the p value at which the mean efficacy change is almost zero (indicated by the horizontal black line), so DriftV ≈ 0 at this point due to P-D balance. Error bars represent s.e.m. in normal scale and relative errors corresponding to s.e.m. in log scale. **(B)** The same as **(A)** except that different lines represent different CV_{events} values, keeping $CV_{rescale} = 0.71$. In **(A,B)**, $\tau_{cross} = 2\text{ms}$, $A_p = 2$, $A_d = 1$, $\tau_{delay} = 1\text{ms}$. The size of the converging motif, as well as the simulation time and trials are the same as in **Figure 4**.

For simplicity, we did not investigate how the heterogeneity of $C_a(\tau)$ interact with other pattern structures. Under STDP, the main effect of the heterogeneity of $C_a(\tau)$ is to influence DriftV with its interaction with the heterogeneity of firing rates through Equation 38, which will not change even when the other pattern statistics are considered. The influence of the heterogeneity of $C_a(\tau)$ onto DriftV can be understood by comparing Equation 38 with

$$r_0^2 E_a \left[\int_{-\infty}^{\infty} d\tau H(\tau) C_a(-\tau_{delay} - \tau) \right]^2 \text{Var}_a(r_a^2) = \frac{E_a(v_a)^2}{E_a(r_a)^2} \text{Var}_a(r_a),$$

which quantifies the DriftV caused by heterogeneity of rates under P-D imbalance.

3.8. The Case When $\tau_{delay} < 0$

In the discussions above, we generally suppose that $\tau_{delay} \geq 0$, which means that the axonal delay for pre-synaptic spikes is no smaller than the dendritic delay for post-synaptic back-propagated spikes. Here we discuss the case when $\tau_{delay} < 0$, i.e., the dendritic delay is larger than the axonal delay.

To understand the relationship between the case $\tau_{delay} > 0$ and the case $\tau_{delay} < 0$, let us first consider a synapse from neuron a to neuron b in the following two cases. Case I: suppose neuron a (neuron b) fires a spike at time t_a (t_b), and $A_p = A$, $A_d = B$ in the STDP time window Equation 6. In this case, whether the synapse is potentiated or depressed depends on the time difference between t_b and $t_a + \tau_{delay}$ (i.e., the sign of $\Delta t_1 \equiv t_b - (t_a + \tau_{delay})$). Case II: now we let the spike delay become $-\tau_{delay}$, and $A_p = B$, $A_d = A$ in Equation 6, and we also reverse the spike

trains (just like showing the two neurons a backward movie), then neuron a will fire at time $T - t_a$ (with T being the total duration of the spike pattern), and neuron b will fire at time $T - t_b$. In this case, whether the synapse is potentiated or depressed depends on the sign of $\Delta t_2 \equiv (T - t_b) - ((T - t_a) - \tau_{delay}) = -\Delta t_1$. If $\Delta t_1 > 0$, then the synapse in Case I will be potentiated by $A \exp(-\Delta t_1 / \tau_{STDP})$, and the synapse in Case II will be depressed by $-A \exp(-\Delta t_1 / \tau_{STDP})$; if $\Delta t_1 < 0$, then the synapse in Case I will be depressed by $-B \exp(\Delta t_1 / \tau_{STDP})$, and the synapse in Case II will be potentiated by $B \exp(\Delta t_1 / \tau_{STDP})$. After applying this argument to every pair of spikes in the spike trains of the central neuron and the a th non-central neuron in a converging motif, we have that

$$\Delta w_{a,p}(\text{Case I}) = -\Delta w_{a,d}(\text{Case II}), \quad (39)$$

$$\Delta w_{a,d}(\text{Case I}) = -\Delta w_{a,p}(\text{Case II}). \quad (40)$$

Therefore, the total synaptic change $\Delta w_a = \Delta w_{a,p} + \Delta w_{a,d}$ satisfies

$$\Delta w_a(\text{Case I}) = -\Delta w_a(\text{Case II}), \quad (41)$$

which means that

$$E_a[\Delta w_a(\text{Case I})] = -E_a[\Delta w_a(\text{Case II})], \quad (42)$$

but

$$\text{Var}_a[\Delta w_a(\text{Case I})] = \text{Var}_a[\Delta w_a(\text{Case II})]. \quad (43)$$

The spike trains generated by our statistical models (Section 2.3) are all statistically time-reversal invariant (i.e., people cannot tell whether a spike pattern is forward or backward by analyzing its statistics). Therefore, from Equation 43, our results on how auto-correlation structure, synchronous firing and heterogeneity of rates influence the efficacy variability are all valid for $\tau_{delay} < 0$, after replacing τ_{delay} by $|\tau_{delay}|$. The influence of heterogeneity of cross-correlations can be obtained by directly substituting $\tau_{delay} < 0$ into Equation 38.

3.9. A Summary on the Efficacy Variability in Converging Motifs

Here we summarize our results in the previous subsections on the efficacy variability in converging motifs.

3.9.1. DiffV

- (1) In stationary spike trains (i.e., spike trains with constant trial-averaged firing rate over time) with *homogeneity* of rates, DiffV gets large either when the spike trains get bursty or strongly regular, and DiffV is smallest when CV is within the range $0.3 \sim 0.7$ (Section 3.2). The only reason why DiffV is large under strong regularity and homogeneity of rates is transient cross-correlation. In stationary spike trains with *heterogeneity* of rates, DiffV tends to monotonically increase with CV (Supplementary Material Section S2), because the transient cross-correlation under strong regularity is destroyed by heterogeneity of rates.
- (2) In stationary spike trains, heterogeneity of rates does not strongly influence DiffV when the spike trains are bursty, but discounts the increase of DiffV with the decrease of CV under strong regularity (Supplementary Material Section S3).
- (3) Synchronous firing usually increases DiffV, except when $\tau_{cross} \leq |\tau_{delay}|$ and each non-central neuron fires almost the same number of spikes during a synchronous event (Section 3.3).
- (4) Under synchronous firing, auto-correlation structure may have a number of types, such as the broadness of the distribution of the spike numbers a neuron fires in different synchronous events ($AT_{SpikeNum}$), the burstiness/regularity of the pieces of spike trains within synchronous events ($AT_{WithinEvent}$), and the burstiness/regularity of the occurrence of synchronous events (AT_{events}). We find that broader distribution of spike number per neuron per synchronous event, burstier spike trains within synchronous events, and burstier occurrence of synchronous events tend to increase DiffV (Section 3.4). These findings can be concluded into a rule of thumb: the burstiness of spike trains tends to increase DiffV, while the regularity tends to decrease DiffV.

3.9.2. DriftV

- (1) Auto-correlation structure in stationary spike trains does not influence DriftV.
- (2) Heterogeneity of rates makes use of P-D imbalance to induce DriftV; and synchronous firing can change P-D imbalance, thereby changing DriftV under heterogeneity of rates (Section 3.5).

- (3) Under synchronous firing, auto-correlation structure can be classified into two classes using rescaled-time transform (**Figure 10**): the factors that influence $CV_{rescale}$ (here, $AT_{SpikeNum}$ and $AT_{WithinEvent}$) and the factors that influence CV_{events} (here, AT_{events}). The factors that influence $CV_{rescale}$ do not change P-D imbalance, thereby nor DriftV under heterogeneity of rates; while the factors that influence CV_{events} change P-D imbalance, thereby DriftV under heterogeneity of rates (Section 3.6). The influence of CV_{events} onto P-D imbalance is complicated (see Supplementary Materials Section S6.2 for detailed discussions); but if synchronous events are not too close to each other, the synapses will get stronger (or weaker) with the increase of CV_{events} if $A_p \exp(\tau_{delay}/\tau_{STDP}) > A_d \exp(-\tau_{delay}/\tau_{STDP})$ (or $A_p \exp(\tau_{delay}/\tau_{STDP}) < A_d \exp(-\tau_{delay}/\tau_{STDP})$).
- (4) Heterogeneity of cross-correlations can induce DriftV together with heterogeneity of rates (Section 3.7).

4. DISCUSSION

In this paper, we systematically studied the influences of four aspects of pattern structures (i.e., synchronous firing, auto-correlation structure, heterogeneity of rates and heterogeneity of cross-correlations, as well as their interactions) onto efficacy variability under STDP, using spike generating models on converging motifs. We separated efficacy variability into two parts: the variability induced by the heterogeneity of change rates of different synapses (DriftV), and the variability induced by weight diffusion caused by stochasticity of spike trains during learning (DiffV). Our main findings are that both synchronous firing and burstiness in auto-correlation structure tend to increase DiffV, heterogeneity of rates induces DriftV when potentiation and depression in STDP are not balanced (P-D imbalance), and heterogeneity of cross-correlations induces DriftV together with heterogeneity of rates.

The reason why we focused on these four aspects of pattern structures is that under STDP they are the only four that mainly influence the lowest order of DiffV and DriftV: for DiffV we only investigated the mean diffusion strength over synapses (i.e., $E_a(\text{Var}_T(\Delta w_a))$, with a being the index of non-central neurons, and T representing integration over trials), neglecting the possible heterogeneity of diffusion strengths for different synapses; and for DriftV we did not consider correlations among drift velocities of synapses. Strictly speaking, heterogeneity of rates and heterogeneity of cross-correlations may not only make different synapses drift at different velocities, but also make them have different diffusion strengths. However, as $\text{DriftV} \propto t^2$ while $\text{DiffV} \propto t$, the heterogeneity of diffusion strengths caused by them is only important when $\text{DriftV} \approx 0$. For simplicity, we did not seriously consider their influences onto the heterogeneity of diffusion strengths in this study, but only briefly discussed it in Supplementary Materials Section S6.

In this paper, the spike trains of the central neurons in converging motifs are generated according to purely statistical models, instead of integrating their inputs. By doing this, we are able to focus to investigate the influence of spike pattern statistics to the efficacy variability, without worrying about

the feedback from synaptic changes to the spike trains of the central neurons. As we mentioned in the introduction, our spike trains are used to mimic the spike patterns of recurrent networks in which the dynamics of the neurons are almost the same. Although biological neurons *do* fire spikes by integrating their inputs, we presume that if we replace the central neurons in converging motifs with a biologically more plausible model (such as integrate-and-fire neurons), the difference between the obtained results from the results in this paper will be limited. For example, if the spike patterns of the pre-synaptic neurons have synchronous firing, then the post-synaptic neuron will also be likely to fire immediately after the synchronous events of the pre-synaptic neurons. The time delay between the firing of the pre- and post-synaptic neurons can be absorbed into the STDP parameter τ_{delay} in our model. If the pre-synaptic neurons fire burstily (regularly), the post-synaptic neurons may also tend to fire burstily (regularly), because of its stronger (regular) input fluctuation. Worsely, if we replace the central neurons with integrate-and-fire neurons, the post-synaptic spike trains will depend on neuronal model parameters, and their statistics will also continuously change during the plasticity of the synapses. These will make the results less convincing.

In our STDP model, synaptic updatings do not depend on the current weights, and the contributions of all spike pairs are added together. We chose this STDP model largely because of its simplicity, especially for analytic treatment. This form of STDP has been used in the literature to study, for example, the generation of functional maps (Song and Abbott, 2001; Widloski and Fiete, 2014) and synfire chains (Jun and Jin, 2007; Fiete et al., 2010) during development, the organization of feedforward or recurrent connections to enhance the processing of correlation or temporal information (Kistler and van Hemmen, 2000; Gilson et al., 2009a,b), the generation of a negative sensory image (Roberts and Bell, 2000), and the control of synchrony (Lubenov and Siapas, 2008) and the evolution of connectivity (Babadi and Abbott, 2013; Ocker et al., 2015) within a recurrent network. Our work provides an angle to review these works, e.g., to understand why these models work in spite of the noises of plasticity, to investigate how the functional performances of STDP may be influenced by the second-order statistics of spike patterns.

There has been literature about the influence of the second-order spike pattern statistics onto the synaptic evolution under STDP. The theoretical works using phenomenological STDP models like ours usually focused on the generation of input selectivity when a post-synaptic neuron or a recurrent network is driven by two pools of inputs with different cross-correlations (Gütig et al., 2003; Gilson et al., 2009a,b,c, 2010). There have been also analytical discussions about the increase of the efficacy variability under Poisson spike trains (Kempster et al., 1999; Gilson et al., 2009a). Comparing to these works, our work has two aspects of contributions. Firstly, we provide an angle to investigate the second-order statistics of spike patterns in a physically intuitive way. To do this, we separated the second-order statistics into four parts (i.e., synchronous firing, auto-correlation structure, heterogeneity of

rates and heterogeneity of cross-correlations, see **Figure 2A**), and considered three types of auto-correlation structure (see **Figure 8**) under spike patterns with synchronous events. All these statistics have clear physical meanings, which helps to build an qualitative understanding on spike patterns. Secondly, we provide a theoretical framework to understand how each of these second-order statistics influences efficacy variability. DiffV under heterogeneity of rates and heterogeneity of cross-correlations can be calculated using the trial-averaged synaptic change rates (Equations 36, 38), which has been done in a large body of literature (e.g., Kempster et al., 1999; Meffi et al., 2006; Gilson et al., 2009a,b,c). To understand the influence onto DiffV, we investigated the change of three parameters (i.e., d , c_I and c_{II}) with pattern structures, each of which has clear physical meaning (see **Figure 3**). This framework can also be applied to study the influence of spike pattern structures onto efficacy variability under other STDP rules. Therefore, we believe our work will facilitate computational researchers to understand the simulation results of their models, and also provide a new angle for experimentalists to interpret their observations.

The STDP time window may have a variety of forms of complex realization, depending on synaptic types, spike patterns and even locations of synapses on dendrites (Caporale and Dan, 2008), and may also depend on the current synaptic weight (Bi and Poo, 1998). STDP may also have a variety of spike pairing schemes (Burkitt et al., 2004; Morrison et al., 2008). We used a STDP model with all-to-all pairing scheme in this paper; and we presume that our results may be more sensitive to the spike pairing scheme than to the shape of STDP time window. Synchronous firing gathers more pre- and post-synaptic spikes closer, which strengthens the STDP interactions between them, so that the variability of spike trains will be more likely to transform into the variability of synaptic weights. As a result, the increase of DiffV with synchronous firing seems to be universal as long as STDP accomplishes large synaptic changes for spike pairs that are adjacent in time. Stronger burstiness induces stronger variability of spike times and also stronger correlations between synaptic updatings caused by adjacent spikes (note the increased d and c_I in **Figure 4**), which may also increase DiffV under other STDP time windows. However, in nearest neighbor spike pairing schemes (Burkitt et al., 2004), a pre-synaptic (post-synaptic) spike may shield other spikes immediately before (after) it from STDP interactions, thereby reducing efficacy variability if spikes are gathered too close by synchronous firing or bursts.

Although the influence of spike patterns varies across systems, the concept of *efficacy variability* should be of general importance. The stochasticity of synapses and neuronal responses as well as the emergent heterogeneity of rates and cross-correlations in network dynamics together make efficacy variability an unavoidable nature of plasticity. Therefore, it is of great meaning to understand how animals make use of efficacy variability and get around of it in future researches. We believe that the concept of efficacy variability not only provides a new perspective to understand the function of plasticity, but is also a new angle to review our current knowledge on learning.

AUTHOR CONTRIBUTIONS

ZB and CZ conceived the idea and designed the research. ZB performed the research. ZB and CZ wrote the paper.

ACKNOWLEDGMENTS

This work was partially supported by Hong Kong Baptist University (HKBU) Strategic Development Fund, NSFC-RGC Joint Research Scheme (Grant No. HKUST/NSFC/12-13/01) and

NSFC (Grant No. 11275027). ZB thanks his PhD supervisor Hai-Jun Zhou for persistent support to his study on computational neuroscience. ZB also thanks David Hansel and Carl van Vreeswijk for helpful discussions during his visit in CNRS UMR 8119 in Université Paris Descartes in 2013.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fncom.2016.00014>

REFERENCES

- Allen, C., and Stevens, C. F. (1994). An evaluation of causes for unreliability of synaptic transmission. *Proc. Natl. Acad. Sci. U.S.A.* 91, 10380–10383. doi: 10.1073/pnas.91.22.10380
- Babadi, B., and Abbott, L. F. (2013). Pairwise analysis can account for network structures arising from spike-timing dependent plasticity. *PLoS Comput. Biol.* 9:e1002906. doi: 10.1371/journal.pcbi.1002906
- Bartos, M., Vida, I., and Jonas, P. (2007). Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat. Rev. Neurosci.* 8, 45–56. doi: 10.1038/nrn2044
- Bi, G. Q., and Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464–10472.
- Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.* 8, 183–208. doi: 10.1023/A:1008925309027
- Brunel, N., and Hakim, V. (1999). Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Comput.* 11, 1621–1671. doi: 10.1162/089976699300016179
- Burkitt, A. N., Meffin, H., and Grayden, D. B. (2004). Spike-timing-dependent plasticity: the relationship to rate-based learning for models with weight dynamics determined by a stable fixed point. *Neural Comput.* 16, 885–940. doi: 10.1162/089976604773135041
- Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745
- Buzsáki, G., and Mizuseki, K. (2014). The log-dynamic brain: how skewed distributions affect network operations. *Nat. Rev. Neurosci.* 15, 264–278. doi: 10.1038/nrn3687
- Cancedda, L., and Poo, M. M. (2009). “Synapse formation and elimination: competition and the role of activity,” in *Encyclopedia of Neuroscience*, ed L. R. Squire (Oxford: Academic Press), 697–703.
- Caporale, N., and Dan, Y. (2008). Spike timing-dependent plasticity: a Hebbian learning rule. *Annu. Rev. Neurosci.* 31, 25–46. doi: 10.1146/annurev.neuro.31.060407.125639
- Clause, A., Kim, G., Sonntag, M., Weisz, C. J., Vetter, D. E., Rübsamen, R., et al. (2014). The precise temporal pattern of prehearing spontaneous activity is necessary for tonotopic map refinement. *Neuron* 82, 822–835. doi: 10.1016/j.neuron.2014.04.001
- Cox, D. R. (1962). *Renewal Theory*. London: Chapman and Hall.
- Dan, Y., and Poo, M. M. (2006). Spike timing-dependent plasticity: from synapse to perception. *Physiol. Rev.* 86, 1033–1048. doi: 10.1152/physrev.00030.2005
- Dragoi, G., and Buzsáki, G. (2006). Temporal encoding of place sequences by hippocampal cell assemblies. *Neuron* 50, 145–157. doi: 10.1016/j.neuron.2006.02.023
- Fiete, I. R., Senn, W., Wang, C., and Hahnloser, R. H. R. (2010). Spike time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron* 65, 563–576. doi: 10.1016/j.neuron.2010.02.003
- Funahashi, S., and Inoue, M. (2000). Neuronal interactions related to working memory processes in the primate prefrontal cortex revealed by cross-correlation analysis. *Cereb. Cortex* 10, 535–551. doi: 10.1093/cercor/10.6.535
- Ganguli, S., and Latham, P. (2009). Feedforward to the past: the relation between neuronal connectivity, amplification, and short-term memory. *Neuron* 61, 499–501. doi: 10.1016/j.neuron.2009.02.006
- Gerstner, W., Kempter, R., van Hemmen, J. L., and Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383, 76–78. doi: 10.1038/383076a0
- Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2009a). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks. I. Input selectivity-strengthening correlated input pathways. *Biol. Cybern.* 101, 81–102. doi: 10.1007/s00422-009-0319-4
- Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2009b). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks. II. Input selectivity-symmetry breaking. *Biol. Cybern.* 101, 103–114. doi: 10.1007/s00422-009-0320-y
- Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2009c). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks IV: structuring synaptic pathways among recurrent connections. *Biol. Cybern.* 101, 427–444. doi: 10.1007/s00422-009-0346-1
- Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2010). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks V: self-organization schemes and weight dependence. *Biol. Cybern.* 103, 365–386. doi: 10.1007/s00422-010-0405-7
- Gireesh, E. D., and Plenz, D. (2008). Neuronal avalanches organize as nested theta- and beta/gamma oscillations during development of cortical layer 2/3. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7576–7581. doi: 10.1073/pnas.0800537105
- Gütig, R., Aharonov, R., Rotter, S., and Sompolinsky, H. (2003). Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. *J. Neurosci.* 23, 3697–3714. Available online at: <http://www.jneurosci.org/content/23/9/3697.full>
- Gutnisky, D. A., and Josić, K. (2010). Generation of spatiotemporally correlated spike trains and local field potentials using a multivariate autoregressive process. *J. Neurophysiol.* 103, 2912–2930. doi: 10.1152/jn.00518.2009
- Jacob, V., Petreanu, L., Wright, N., Svoboda, K., and Fox, K. (2012). Regular spiking and intrinsic bursting pyramidal cells show orthogonal forms of experience-dependent plasticity in layer V of barrel cortex. *Neuron* 73, 391–404. doi: 10.1016/j.neuron.2011.11.034
- Jun, J. K., and Jin, D. (2007). Development of neural circuitry for precise temporal sequences through spontaneous activity, axon remodeling, and synaptic plasticity. *PLoS ONE* 2:e723. doi: 10.1371/journal.pone.0000723
- Kamioka, H., Maeda, E., Jimbo, Y., Robinson, H. P. C., and Kawana, A. (1996). Spontaneous periodic synchronized bursting during formation of mature patterns of connections in cortical cultures. *Neurosci. Lett.* 206, 109–112. doi: 10.1016/S0304-3940(96)12448-4
- Kempter, R., Gerstner, W., and van Hemmen, J. L. (1999). Hebbian learning and spiking neurons. *Phys. Rev. E* 59, 4498–4515. doi: 10.1103/PhysRevE.59.4498
- Kistler, W. M., and van Hemmen, J. L. (2000). Modeling synaptic plasticity in conjunction with the timing of preand postsynaptic action potentials. *Neural Comput.* 12, 385–405. doi: 10.1162/089976600300015844

- Kohn, A., and Smith, M. A. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J. Neurosci.* 25, 3661–3673. doi: 10.1523/JNEUROSCI.5106-04.2005
- Krumin, M., and Shoham, S. (2009). Generation of spike trains with controlled auto- and cross-correlation functions. *Neural Comput.* 21, 1642–1664. doi: 10.1162/neco.2009.08-08-847
- Long, M. A., Jin, D. Z., and Fee, M. S. (2010). Support for a synaptic chain model of neuronal sequence generation. *Nature* 468, 394–399. doi: 10.1038/nature09514
- Lubenov, E. V., and Siapas, A. G. (2008). Decoupling through synchrony in neuronal circuits with propagation delays. *Neuron* 58, 118–131. doi: 10.1016/j.neuron.2008.01.036
- Macke, J. H., Berens, P., Ecker, A. S., Tolias, A. S., and Bethge, M. (2009). Generating spike trains with specified correlation coefficients. *Neural Comput.* 21, 397–423. doi: 10.1162/neco.2008.02-08-713
- Mainen, Z. F., and Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science* 268, 1503–1506. doi: 10.1126/science.7770778
- Markram, H., Gerstner, W., and Sjöström, P. J. (2012). Spike-timing-dependent plasticity: a comprehensive overview. *Front. Synaptic Neurosci.* 4:2. doi: 10.3389/fnsyn.2012.00002
- Meffi, H., Besson, J., Burkitt, A. N., and Grayden, D. B. (2006). Learning the structure of correlated synaptic subgroups using stable and competitive spike-timing-dependent plasticity. *Phys. Rev. E* 73:041911. doi: 10.1103/PhysRevE.73.041911
- Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science* 319, 1543–1546. doi: 10.1126/science.1150769
- Morrison, A., Diesmann, M., and Gerstner, W. (2008). Phenomenological models of synaptic plasticity based on spike timing. *Biol. Cybern.* 98, 459–478. doi: 10.1007/s00422-008-0233-1
- Nawrot, M. P., Boucsein, C., Molina, V. R., Riehle, A., Aertsen, A., and Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. *J. Neurosci. Methods* 169, 374–390. doi: 10.1016/j.jneumeth.2007.10.013
- Ocker, G. K., Litwin-Kumar, A., and Doiron, B. (2015). Self-organization of microcircuits in networks of spiking neurons with plastic synapses. *PLoS Comput. Biol.* 11:e1004458. doi: 10.1371/journal.pcbi.1004458
- O'Connor, D. H., Peron, S. P., Huber, D., and Svoboda, K. (2010). Neural activity in barrel cortex underlying vibrissa-based object localization in mice. *Neuron* 67, 1048–1061. doi: 10.1016/j.neuron.2010.08.026
- Ostojic, S. (2014). Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. *Nat. Neurosci.* 17, 594–600. doi: 10.1038/nn.3658
- Ostojic, S., Brunel, N., and Hakim, V. (2009). How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *J. Neurosci.* 29, 10234–10253. doi: 10.1523/JNEUROSCI.1275-09.2009
- Petermann, T., Thiagarajana, T. C., Lebedev, M. A., Nicolelis, M. A. L., Chialvo, D. R., and Plenz, D. (2009). Spontaneous cortical activity in awake monkeys composed of neuronal avalanches. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15921–15926. doi: 10.1073/pnas.0904089106
- Pillow, J. W. (2009). “Time-rescaling methods for the estimation and assessment of non-Poisson neural encoding models,” in *Advances in Neural Information Processing System 22*, eds Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta (Cambridge: MIT Press), 1473–1481.
- Roberts, P. D., and Bell, C. C. (2000). Computational consequences of temporally asymmetric learning rules. II. Sensory image cancellation. *J. Comput. Neurosci.* 9, 67–83. doi: 10.1023/A:1008938428112
- Roxin, A., Brunel, N., Hansel, D., Mongillo, G., and van Vreeswijk, C. (2011). On the distribution of firing rates in networks of cortical neurons. *J. Neurosci.* 31, 16217–16226. doi: 10.1523/JNEUROSCI.1677-11.2011
- Schneidman, E., Berry, M. J. II, Sege, R., and Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440, 1007–1012. doi: 10.1038/nature04701
- Schwindt, P., and Crill, W. (1999). Mechanisms underlying burst and regular spiking evoked by dendritic depolarization in layer 5 cortical pyramidal neurons. *J. Neurophysiol.* 81, 1341–1354.
- Shadlen, M. N., and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* 18, 3870–3896.
- Shafi, M., Zhou, Y., Quintana, J., Chow, C., Fuster, J., and Bodner, M. (2007). Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience* 146, 1082–1108. doi: 10.1016/j.neuroscience.2006.12.072
- Softky, W. R., and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* 13, 334–350.
- Song, S., and Abbott, L. F. (2001). Cortical development and remapping through spike timing-dependent plasticity. *Neuron* 32, 339–350. doi: 10.1016/S0896-6273(01)00451-2
- Trousdale, J., Hu, Y., Shea-Brown, E., and Josić, K. (2012). Impact of network structure and cellular response on spike time correlations. *PLoS Comput. Biol.* 8:e1002408. doi: 10.1371/journal.pcbi.1002408
- Tuckwell, H. C. (1988). *Introduction to Theoretical Neurobiology, Vol. 2*. Cambridge: Cambridge University Press.
- Turrigiano, G. (2011). Too many cooks? Intrinsic and synaptic homeostatic mechanisms in cortical circuit refinement. *Annu. Rev. Neurosci.* 34, 89–103. doi: 10.1146/annurev-neuro-060909-153238
- Turrigiano, G. G., and Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nat. Rev. Neurosci.* 5, 97–107. doi: 10.1038/nrn1327
- van Vreeswijk, C., and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274, 1724–1726. doi: 10.1126/science.274.5293.1724
- van Vreeswijk, C., and Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Comput.* 10, 1321–1371. doi: 10.1162/089976698300017214
- Watt, A. J., and Desai, N. S. (2010). Homeostatic plasticity and STDP: keeping a neuron's cool in a fluctuating world. *Front. Synaptic Neurosci.* 2:5. doi: 10.3389/fnsyn.2010.00005
- Widloski, J., and Fiete, I. R. (2014). A model of grid cell development through spatial exploration and spike time-dependent plasticity. *Neuron* 83, 481–495. doi: 10.1016/j.neuron.2014.06.018

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Bi and Zhou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.