



# Effects of Firing Variability on Network Structures with Spike-Timing-Dependent Plasticity

Bin Min<sup>1,2</sup>, Douglas Zhou<sup>3\*</sup> and David Cai<sup>1,2,3\*</sup>

<sup>1</sup> Center for Neural Science, Courant Institute of Mathematical Sciences, New York University, New York, NY, United States, <sup>2</sup> NYUAD Institute, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, <sup>3</sup> School of Mathematical Sciences, MOE-LSC, Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai, China

Synaptic plasticity is believed to be the biological substrate underlying learning and memory. One of the most widespread forms of synaptic plasticity, spike-timing-dependent plasticity (STDP), uses the spike timing information of presynaptic and postsynaptic neurons to induce synaptic potentiation or depression. An open question is how STDP organizes the connectivity patterns in neuronal circuits. Previous studies have placed much emphasis on the role of firing rate in shaping connectivity patterns. Here, we go beyond the firing rate description to develop a self-consistent linear response theory that incorporates the information of both firing rate and firing variability. By decomposing the pairwise spike correlation into one component associated with local direct connections and the other associated with indirect connections, we identify two distinct regimes regarding the network structures learned through STDP. In one regime, the contribution of the direct-connection correlations dominates over that of the indirect-connection correlations in the learning dynamics; this gives rise to a network structure consistent with the firing rate description. In the other regime, the contribution of the indirect-connection correlations dominates in the learning dynamics, leading to a network structure different from the firing rate description. We demonstrate that the heterogeneity of firing variability across neuronal populations induces a temporally asymmetric structure of indirect-connection correlations. This temporally asymmetric structure underlies the emergence of the second regime. Our study provides a new perspective that emphasizes the role of high-order statistics of spiking activity in the spike-correlation-sensitive learning dynamics.

**Keywords:** STDP, linear response theory, correlation structure, firing variability, synaptic plasticity

## OPEN ACCESS

### Edited by:

Paul Miller,  
Brandeis University, United States

### Reviewed by:

Maxim Volgushev,  
University of Connecticut,  
United States  
Joost Le Feber,  
University of Twente, Netherlands

### \*Correspondence:

Douglas Zhou  
dzdz@situ.edu.cn  
Dedicated to  
David Cai

**Received:** 17 August 2017

**Accepted:** 03 January 2018

**Published:** 23 January 2018

### Citation:

Min B, Zhou D and Cai D (2018)  
Effects of Firing Variability on Network  
Structures with  
Spike-Timing-Dependent Plasticity.  
*Front. Comput. Neurosci.* 12:1.  
doi: 10.3389/fncom.2018.00001

## INTRODUCTION

Spike-timing-dependent plasticity (STDP) (Markram et al., 1997; Bi and Poo, 1998; Caporale and Dan, 2008) is one of the major forms of Hebbian learning. In canonical STDP, a synapse is potentiated (depressed) when the presynaptic spikes precede (follow) the postsynaptic ones. With this temporally asymmetric learning window, STDP can regulate both the mean rate and variability of postsynaptic firing (Song et al., 2000), enforce the competition between convergent synaptic inputs (Song et al., 2000), and support temporal sequence learning (Rao and Sejnowski, 2001).

While STDP of a single neuron with many synapses has been well characterized, issues related to the STDP learning dynamics in recurrent networks are yet to be fully addressed, in spite of recent progress (Morrison et al., 2007; Clopath et al., 2010; Gilson, 2010; Kozloski and Cecchi, 2010; Babadi and Abbott, 2013, 2016; Litwin-Kumar and Doiron, 2014; Ocker et al., 2015; Zenke et al., 2015; Bi and Zhou, 2016; Ravid Tannenbaum and Burak, 2016). The role of the low-order statistical structure (i.e., firing rate) of spiking activity has been intensively investigated in the learning dynamics in recurrent networks (Morrison et al., 2007; Kozloski and Cecchi, 2010; Babadi and Abbott, 2013). Though it is known that, as a correlation-sensitive learning rule, STDP inherently incorporates the high-order statistics of spiking activity (Morrison et al., 2008), the role of the high-order statistics of spiking activity in STDP learning dynamics for recurrent networks remains largely unknown. In this work, we explore the issue of how the high-order firing statistics (for example, firing variability) affect the network connectivity structures learned through STDP in recurrent networks. We demonstrate that the information of firing variability is encoded in a linear response kernel that quantifies how a neuron responds to a weak incoming signal. By developing a self-consistent linear response theory pioneered in Pernice et al. (2012), Trousdale et al. (2012), and Helias et al. (2013), we show that this response kernel, together with the firing rate, determines the pairwise correlation of spiking activity. By decomposing the pairwise correlation into one component associated with local direct connections and the other associated with indirect connections, we can identify two distinct dynamical regimes. In the first regime, the local direct connections dominate over the indirect connections in determining the learned network structure. This regime corresponds to those studied in Babadi and Abbott (2013), Kozloski and Cecchi (2010), and Morrison et al. (2007), where the connections from neurons of higher firing rate to those of lower firing rate are strengthened whereas the connections with the opposite situation are weakened. The other is a regime in which the indirect connections dominate over the direct connections in learning dynamics. In this second regime, the connections from neurons of higher firing rate and higher firing variability to those of lower firing rate and lower firing variability are weakened whereas the connections with the opposite situation (i.e., connections from neurons of lower firing rate and lower firing variability to those of higher firing rate and higher firing variability) are strengthened. We find that there is a temporally asymmetric structure of common-input-induced spike correlations between different neuronal populations. This structure is induced by the heterogeneity of firing variability across neuronal populations, which underlies this counterintuitive dynamical regime. Our study provides a theoretical framework for addressing the question of how the high-order statistics of spiking activity affect network structures learned through STDP. Using this framework, we demonstrate how the introduction of heterogeneity of firing variability across different neuronal populations can give rise to counterintuitive network connectivity structures in a plastic neuronal network.

## METHODS

### Model

The model we used here is the current-based leaky integrate-and-fire (LIF) model with finite synaptic time constant. It is the same as that in Babadi and Abbott (2013).

$$\tau_m \frac{dV_i}{dt} = (V_r - V_i) + I_i, \quad (1)$$

$$\frac{dI_i}{dt} = -\frac{I_i}{\tau_s} + \sum_{j=1}^{N_E, N_I} w_{ij} s_j(t) + \frac{\mu_{\text{ext},i}}{\tau_s} + \frac{\sigma_{\text{ext},i}}{\tau_s} \sqrt{\tau_m} \xi_i(t), \quad (2)$$

where  $\tau_m = 20$  ms is the membrane time constant,  $V_r = -60$  mV is the resting potential,  $\tau_s = 5$  ms is the synaptic time constant,  $w_{ij}$  is the synaptic strength from neuron  $j$  to neuron  $i$  and  $s_j(t)$  is the spike train of neuron  $j$ . The parameters  $\mu_{\text{ext},i}$  and  $\sigma_{\text{ext},i}^2$  are the mean and variance of the external input, respectively, and  $\xi_i$  is the white noise satisfying  $\langle \xi_i \rangle = 0$  and  $\langle \xi_i(t) \xi_j(t') \rangle = \delta_{ij} \delta(t - t')$ . Every time when the membrane potential  $V_i$  crosses the threshold  $V_{\text{th}} = -40$  mV, neuron  $i$  would emit a spike and  $V_i$  would be reset to  $V_r$ .

The network consists of  $N_E$  excitatory neurons and  $N_I$  inhibitory neurons. We use  $N_E = 250$ ,  $N_I = 250$ . The connections between these neurons are of the all-to-all type. We keep the inhibitory-to-excitatory (EI), excitatory-to-inhibitory (IE) and inhibitory-to-inhibitory (II) synaptic strengths constant during the entire simulation while the excitatory-to-excitatory (EE) connections are subject to the STDP learning rule to be introduced in section “The effect of firing rate and firing variability on learning dynamics—numerical evidence”. At the beginning of simulations, the strengths for the EE and IE are drawn at random from a uniform distribution from 0 mV to  $w_{EE}^{\text{max}} = 1$  mV and  $w_{IE}^{\text{max}} = 2$  mV, respectively, and the strengths for the EI and II connections are drawn at random from a uniform distribution from  $w_{EI}^{\text{max}} = -4$  mV and  $w_{II}^{\text{max}} = -4$  to 0 mV, respectively.

Based on the mean and variance of external inputs, we divide the excitatory neurons into three populations in which the  $\alpha$ th ( $\alpha = 1, 2, 3$ ) population receives the external input of the mean  $\mu_{\text{ext},\alpha}^p$  and standard deviation  $\sigma_{\text{ext},\alpha}^p$ . The number of neurons in the  $\alpha$ th ( $\alpha = 1, 2, 3$ ) population is denoted as  $N_\alpha$ .

### Statistical Properties of a Single LIF Neuron

We first describe the statistical properties of a single LIF neuron. The dynamics of a single LIF neuron with finite synaptic time constant is governed by Equation (1) and the following equation:

$$\tau_s \frac{dI}{dt} = -I + \mu + \sigma \sqrt{\tau_m} \xi(t), \quad (3)$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation of the external input respectively, and  $\xi(t)$  is the Gaussian white noise satisfying  $\langle \xi \rangle = 0$  and  $\langle \xi(t) \xi(t') \rangle = \delta(t - t')$ .

When  $\tau_s$  is much smaller than  $\tau_m$ , the mean firing rate  $\bar{r}$  has the following approximation:

$$\bar{r} = F(\mu, \sigma) = \left[ \tau_m \sqrt{\pi} \int_{-\frac{\mu}{\sigma} + \beta}^{\frac{V_{th} - V_r - \mu}{\sigma} + \beta} e^{x^2} (1 + \operatorname{erf}(x)) dx \right]^{-1}, \quad (4)$$

where  $\beta = -\zeta(1/2) \sqrt{\tau_s / (2\tau_m)}$ ,  $\operatorname{erf}(\cdot)$  is the error function and  $\zeta(\cdot)$  is the Riemann-Zeta function (Fourcaud and Brunel, 2002).

We use a response kernel to characterize how the LIF neuron ensemble responds to a perturbation. Specifically, we consider Equation (1) and the following equation:

$$\tau_s \frac{dI}{dt} = -I + \mu + \sigma \sqrt{\tau_m} \xi(t) + u(t), \quad (5)$$

where  $u(t)$  is a perturbation. When  $u(t)$  is weak, the difference between the instantaneous firing rate  $r(t)$  of neuronal ensemble that is governed by Equations (1, 5) and the mean firing rate  $\bar{r}$  of neuronal ensemble that is governed by Equations (1, 3) is small. Therefore, we can approximate the instantaneous firing rate  $r(t)$  by the mean firing rate  $\bar{r}$  plus a linear response term with respect to  $u(t)$ , i.e.,

$$r(t) = \bar{r} + (h * u)(t), \quad (6)$$

where  $h(\tau)$  is the linear response kernel of Equations (1, 3) and the symbol  $*$  stands for temporal convolution.

The analytical results for the linear response kernel are available for the case of zero synaptic time constant (Fourcaud and Brunel, 2002). Recently, for the case of nonzero synaptic time constant, the analytical result was derived (Schuecker et al., 2015). Here, we use the reverse correlation method to compute it (Dayan and Abbott, 2001; Ostojic et al., 2009). Specifically, we inject a weak Gaussian white noise signal  $\xi_s(t)$  with correlation  $\langle \xi_s(t + \tau) \xi_s(t) \rangle = \sigma_s^2 \delta(\tau)$  to the LIF neuron, i.e., adding  $\xi_s(t)$  on the right hand side of Equation (3). Note that this testing white noise signal  $\xi_s(t)$  is different from the external noise input  $\xi(t)$  in Equation (3). By using the spike-triggered average technique, we calculate the response kernel in the following way:

$$h(\tau) = \frac{\frac{1}{N_T} \sum_{i=1}^{N_T} \xi_s(t_i - \tau) \bar{r}}{\sigma_s^2}, \quad (7)$$

where  $\{t_i\}_{i=1}^{N_T}$  is the spike timing of the LIF neuron in the duration  $T$  and the firing rate  $\bar{r}$  is given by Equation (4) (Dayan and Abbott, 2001). In our simulation,  $\sigma_s = 2$  mV and  $T = 10^6$  ms. We use one realization of  $\xi_s(t)$  and perform the average over 1,000 realizations of external noise input  $\xi(t)$ .

## Linear Response Theory

Here, we describe how we obtain  $\bar{r}_i$  and  $h_i(\tau)$  for  $i = 1, \dots, N$  in a self-consistent manner for the network described in Equations (1, 2). By splitting the spike train  $s_j(t)$  in Equation (2) into the

mean firing rate and its fluctuations, we can rewrite Equation (2) in the following form:

$$\begin{aligned} \frac{dI_i}{dt} = & -\frac{I_i}{\tau_s} + \sum_{j=1}^N w_{ij} \bar{r}_j + \frac{\mu_{\text{ext},i}}{\tau_s} + \frac{\sigma_{\text{ext},i}}{\tau_s} \sqrt{\tau_m} \xi_i(t) \\ & + \sum_{j=1}^N w_{ij} (s_j - \bar{r}_j), \end{aligned} \quad (8)$$

where the term  $\sum_{j=1}^N w_{ij} (s_j - \bar{r}_j)$  can be regarded as a perturbation. By enforcing the first-order Volterra approximation of Equations (1, 8), we arrive at Equation (31), where  $\bar{r}_i$  and  $h_i(\tau)$  are the mean firing rate and response kernel, respectively, of Equation (1) and the following equation:

$$\frac{dI_i}{dt} = -\frac{I_i}{\tau_s} + \sum_{j=1}^N w_{ij} \bar{r}_j + \frac{\mu_{\text{ext},i}}{\tau_s} + \frac{\sigma_{\text{ext},i}}{\tau_s} \sqrt{\tau_m} \xi_i(t). \quad (9)$$

Here, we can observe that, even with the same mean and variance value of external inputs, different neurons may still have different mean firing rates due to the randomness of synaptic strengths. To simplify the self-consistent derivation of mean firing rate, we make a further assumption that the mean firing rates of single neurons within the same population are approximated by the population-averaged mean firing rate. Denote the population-averaged mean firing rates of the different populations as  $\bar{r}_\alpha^p$ ,  $\alpha = 1, 2, 3, I$ , where  $I$  stands for the inhibitory population. Then,  $\bar{r}_\alpha^p$  is the mean firing rate of the following equations:

$$\tau_m \frac{dV_\alpha}{dt} = (V_r - V_\alpha) + I_\alpha, \quad \text{for } \alpha = 1, 2, 3, I, \quad (10)$$

$$\begin{aligned} \frac{dI_\alpha}{dt} = & -\frac{I_\alpha}{\tau_s} + \left( \sum_{\beta=1}^3 N_\beta w_{\alpha\beta}^p \bar{r}_\beta^p + N_I w_{\alpha I}^p \bar{r}_I^p \right) + \frac{\mu_{\text{ext},\alpha}^p}{\tau_s} \\ & + \frac{\sigma_{\text{ext},\alpha}^p}{\tau_s} \sqrt{\tau_m} \xi_\alpha(t), \end{aligned} \quad (11)$$

where  $w_{\alpha\beta}^p$  stands for the average synaptic strength from Population  $\beta$  to Population  $\alpha$ . Therefore,  $\bar{r}_\alpha^p$  can be determined in the following self-consistent way:

$$\bar{r}_\alpha^p = F^{-1}(\mu_\alpha^p, \sigma_{\text{ext},\alpha}^p), \quad \text{for } \alpha = 1, 2, 3, I, \quad (12)$$

$$\mu_\alpha^p = \tau_s \left( \sum_{\beta=1}^3 N_\beta w_{\alpha\beta}^p \bar{r}_\beta^p + N_I w_{\alpha I}^p \bar{r}_I^p \right) + \mu_{\text{ext},\alpha}^p. \quad (13)$$

## The Power Spectrum for the Hawkes Process

The first derivation of the power spectrum of Hawkes process (i.e., Equation 31) was carried out by Hawkes (1971a), in which the Wiener-Hopf theory was used. Here, we follow the method

in Grytskyy et al. (2013). First, for the spike correlation function  $\mathbf{C}(\tau)$ , the following equations hold:

$$\mathbf{C}(\tau) \equiv \langle \mathbf{s}(t + \tau) \mathbf{s}^T(t) \rangle - \langle \mathbf{s}(t) \rangle \langle \mathbf{s}^T(t) \rangle \quad (14a)$$

$$= \langle (\mathbf{r}(t + \tau) - \bar{\mathbf{r}}) \mathbf{s}^T(t) \rangle + \mathbf{D} \delta(\tau) \quad (14b)$$

$$= \langle [\mathbf{h} * \mathbf{w}(\mathbf{s} - \bar{\mathbf{r}})](t + \tau) \mathbf{s}^T(t) \rangle + \mathbf{D} \delta(\tau) \quad (14c)$$

$$= (\mathbf{h} * \mathbf{w} \mathbf{C})(t) + \mathbf{D} \delta(\tau), \quad (14d)$$

for  $\tau \geq 0$ , where  $\mathbf{s}(t)$  and  $\mathbf{r}(t)$  are the spike train and the instantaneous firing rate of Hawkes process, respectively, and  $\mathbf{D} = \text{diag}\{\bar{r}_1, \dots, \bar{r}_N\}$ . In Equation (14b), we use the fact that  $\langle \mathbf{s}(t + \tau) \mathbf{s}^T(t) \rangle = \langle \mathbf{r}(t + \tau) \mathbf{s}^T(t) \rangle + \mathbf{D} \delta(\tau)$ . In Equation (14c), we insert the Hawkes process (Equation 31). In Equation (14d), we take advantage of the definition of spike correlation. Next, define  $\mathbf{y}(t) = \mathbf{r}_y(t) + \mathbf{x}(t)$ , where

$$\mathbf{r}_y(t) = \bar{\mathbf{r}} + [\mathbf{h} * \mathbf{w}(\mathbf{r}_y - \bar{\mathbf{r}} + \mathbf{x})](t), \quad (15)$$

and  $\mathbf{x}(t)$  is the Gaussian white noise with correlation  $\langle \mathbf{x}(t + \tau) \mathbf{x}^T(t) \rangle = \mathbf{D}_x \delta(\tau)$ . For  $\mathbf{y}(t)$ ,

$$\mathbf{C}_y(\tau) \equiv \langle \mathbf{y}(t + \tau) \mathbf{y}^T(t) - \bar{\mathbf{r}}^T \rangle \quad (16a)$$

$$= \langle [\mathbf{h} * \mathbf{w}(\mathbf{y} - \bar{\mathbf{r}}) + \mathbf{x}](t + \tau) (\mathbf{y}^T(t) - \bar{\mathbf{r}}^T) \rangle \quad (16b)$$

$$= (\mathbf{h} * \mathbf{w} \mathbf{C}_y)(\tau) + \langle \mathbf{x}(t + \tau) (\mathbf{r}_y^T(t) - \bar{\mathbf{r}}^T) \rangle + \langle \mathbf{x}(t + \tau) \mathbf{x}^T(t) \rangle, \quad (16c)$$

$$= (\mathbf{h} * \mathbf{w} \mathbf{C}_y)(\tau) + \mathbf{D}_x \delta(\tau), \quad (16d)$$

for  $\tau \geq 0$ . In Equation (16b), we have made use of the definition of  $\mathbf{y}(t)$ . In Equation (16c), we take advantage of the definition of  $\mathbf{C}_y(t)$ . In Equation (18), we use the fact that  $\langle \mathbf{x}(t + \tau) \mathbf{r}_y^T(t) \rangle = \langle \mathbf{x}(t + \tau) \rangle \langle \mathbf{r}_y^T(t) \rangle = 0$ . By setting  $\mathbf{D}_x = \mathbf{D}$ , we obtain  $\mathbf{C}(\tau) = \mathbf{C}_y(\tau)$ . Therefore, to understand the correlation structure of the Hawkes process, it suffices to study the correlation structure of  $\mathbf{y}(t)$ . Define  $\hat{\mathbf{r}}_y(\omega)$  as the Fourier transform of  $\mathbf{r}_y(t)$ . From Equation (16b), we have

$$\hat{\mathbf{r}}_y = \hat{\mathbf{h}}(\omega) \mathbf{w} (\hat{\mathbf{r}}_y + \hat{\mathbf{x}}). \quad (17)$$

Thus,  $\hat{\mathbf{y}} = \hat{\mathbf{r}}_y + \hat{\mathbf{x}} = (1 - \hat{\mathbf{h}} \mathbf{w})^{-1} \hat{\mathbf{x}}$ . By using the Wiener-Khinchin theorem, we have

$$\begin{aligned} \hat{\mathbf{C}}(\omega) &= \langle \hat{\mathbf{y}}(\omega) \hat{\mathbf{y}}^T(-\omega) \rangle \\ &= (1 - \hat{\mathbf{h}}(\omega) \mathbf{w})^{-1} \mathbf{D} (1 - \mathbf{w}^T \hat{\mathbf{h}}(-\omega))^{-1}. \end{aligned} \quad (18)$$

This completes the derivation of the power spectrum formula as discussed in Equation (32).

## Population-Averaged Spike Correlation

Assume that there are  $N_\alpha$  ( $\alpha = 1, 2, 3, I$ ) neurons in the  $\alpha$ th population. The spike correlation between population  $\alpha$  and population  $\beta$  is defined as follows:

$$C_{\alpha\beta}^p(\tau) = \frac{1}{N_\alpha N_\beta} \sum_{i \in \alpha, j \in \beta} C_{ij}(\tau). \quad (19)$$

The corresponding spike cross correlation is defined as follows:

$$C_{\alpha\beta}^{p,cr}(\tau) = \frac{1}{N_\alpha N_\beta} \sum_{i \in \alpha, j \in \beta, i \neq j} C_{ij}(\tau). \quad (20)$$

Note that  $C_{\alpha\beta}^{p,cr}(\tau)$  is used in the caption of **Figure 4**. Inserting Equation (16a) into Equation (19), we can obtain

$$\begin{aligned} C_{\alpha\beta}^p(\tau) &= \frac{1}{N_\alpha N_\beta} \left\{ \sum_{\gamma} \sum_{i \in \alpha, j \in \beta, k \in \gamma} [h_i * w_{ik} C_{kj}](\tau) \right. \\ &\quad \left. + \sum_{i \in \alpha, j \in \beta} D_i \delta_{ij} \delta(\tau) \right\} \\ &= \sum_{\gamma} h_{\alpha}^p w_{\alpha\gamma}^p N_{\gamma} C_{\gamma\beta}(\tau) + \frac{\bar{r}_{\alpha}^p}{N_{\alpha}} \delta_{\alpha\beta} \delta(\tau), \end{aligned} \quad (21)$$

where  $h_{\alpha}^p(\tau)$  stands for the linear response kernel of Population  $\alpha$ . Similar to the result (Equation 18), the population-averaged power spectrum  $\hat{\mathbf{C}}^p(\cdot) = (\hat{\mathbf{C}}_{\alpha\beta}^p(\cdot))$  has the following expression:

$$\hat{\mathbf{C}}^p(\omega) = (1 - \hat{\mathbf{h}}^p(\omega) \mathbf{w}^p)^{-1} \mathbf{D}^p (1 - \mathbf{w}^{pT} \hat{\mathbf{h}}^p(-\omega))^{-1}, \quad (22)$$

where  $\mathbf{h}^p(\cdot) = \text{diag}\{h_1^p(\cdot), h_2^p(\cdot), h_3^p(\cdot), h_I^p(\cdot)\}$ ,  $\mathbf{w}^p = (w_{\alpha\beta}^p)$  and  $\mathbf{D}^p = \text{diag}\{\bar{r}_1^p/N_1, \bar{r}_2^p/N_2, \bar{r}_3^p/N_3, \bar{r}_I^p/N_I\}$ .

## Decomposition of Cross Correlation Structures

Formally, the power spectrum (Equation 18) can be expanded in the following way:

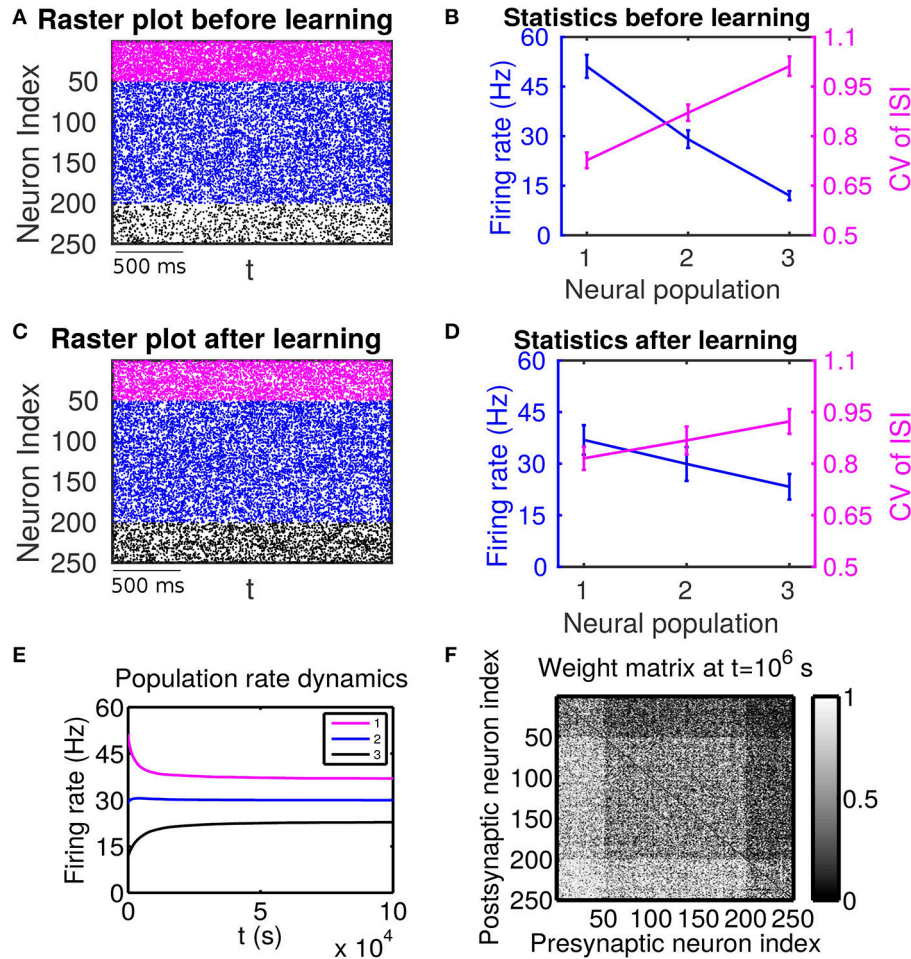
$$\hat{\mathbf{C}}(\omega) = \mathbf{D} + \hat{\mathbf{h}}(\omega) \mathbf{w} \mathbf{D} + \mathbf{D} \mathbf{w}^T \hat{\mathbf{h}}^T(-\omega) + \hat{\mathbf{h}}(\omega) \mathbf{w} \mathbf{D} \mathbf{w}^T \hat{\mathbf{h}}^T(-\omega) + \dots \quad (23)$$

Therefore, the cross correlation  $\mathbf{C}(\tau)$ , which is the inverse Fourier transform of  $\hat{\mathbf{C}}(\omega)$  by the Wiener-Khinchin theorem, can be approximated by Equation (33). The same decomposition can also be carried out for the population-averaged cross correlation as discussed previously.

## RESULTS

We first present the numerical evidence of the complexity of mutual interactions between plastic network structures and spiking activity, as it relates to the effect of high-order statistical structures implied in the STDP learning dynamics. In particular, we emphasize the importance of firing variability of spiking activity for the network structures learned through STDP. This fact is particularly emphasized by the sharp contrast of learned network structures between two scenarios (Case I in **Figure 1** and Case II **Figure 2**); they share the same trend for firing rate change across different neuronal populations, but have the opposite trend for their firing variability change. We then develop a self-consistent linear response theory to account for the effects of both firing rate and firing variability on learning dynamics and explain the role observed in our numerical study of high-order statistics in learning dynamics.





**FIGURE 1 |** The effect of firing rate and variability on STDP learning dynamics—Case I. All excitatory neurons receive external inputs with an identical variance (i.e.,  $\sigma_{\text{ext},\alpha}^p = 15.8$  mV for  $\alpha = 1, 2, 3$ ). Based on the values of mean of external inputs, the excitatory neurons are divided into three populations, in which Population 1 (including neurons labeled by 1–50, purple), Population 2 (including neurons labeled by 51–200, blue) and Population 3 (including neurons labeled by 201–250, black) receive the largest ( $\mu_{\text{ext},1}^p = 40$  mV), intermediate ( $\mu_{\text{ext},2}^p = 30$  mV) and smallest ( $\mu_{\text{ext},3}^p = 20$  mV) mean, respectively. The initial values of excitatory-to-excitatory connection strengths are drawn from a uniform distribution ranging from 0 to 1 mV. **(A,C)** Raster plots before learning and after learning, respectively. **(B,D)** Firing rates and coefficients of variation (CVs) of the different populations before learning and after  $10^6$  s of learning, respectively. Error bars indicate the standard deviation of corresponding values within each population. **(E)** Firing rate dynamics of the different populations during learning. **(F)** Weight matrix after  $10^6$  s of learning. The largest (i.e., brightest) and the smallest (i.e., darkest) value of the gray bar is 1 and 0 mV, respectively.

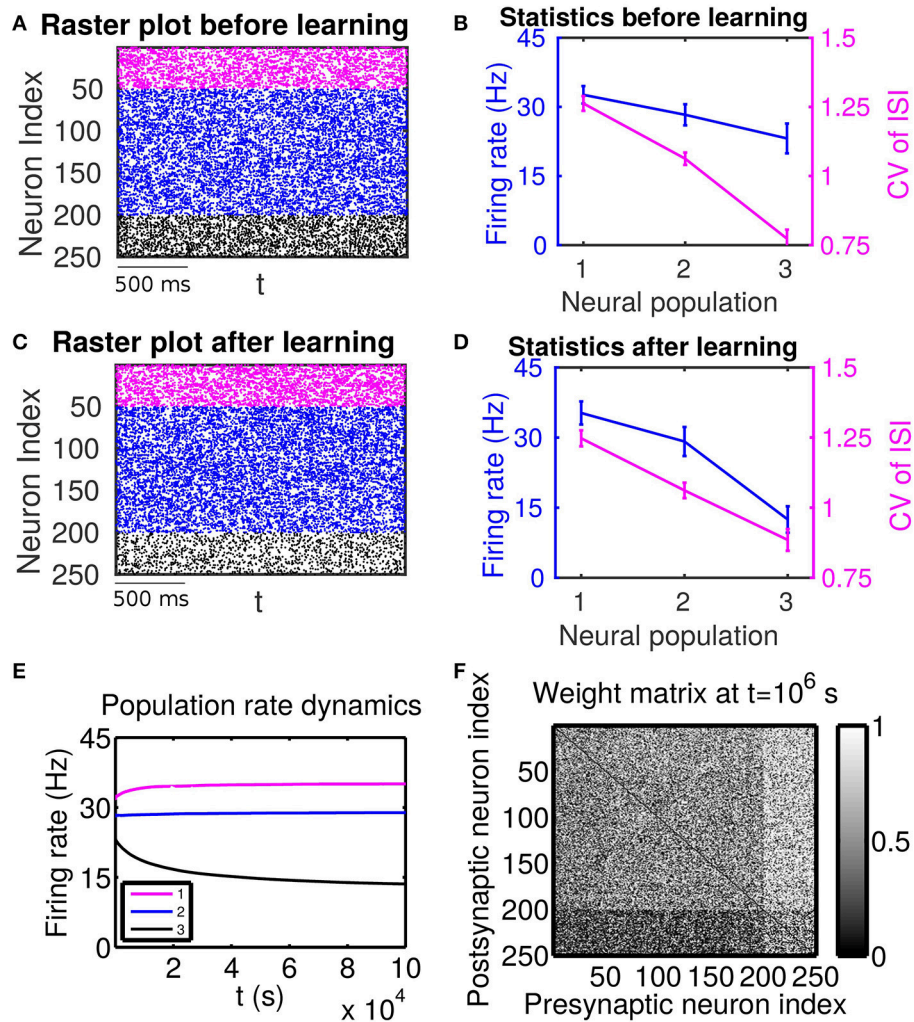
## The Effect of Firing Rate and Firing Variability on Learning Dynamics—Numerical Evidence

For our numerical experiments, we use the current-based leaky integrate-and-fire (LIF) neuron model with finite synaptic time constant (Babadi and Abbott, 2013). The network consists of 250 excitatory neurons and 250 inhibitory neurons (See section Methods for details). Based on the values of mean and variance of external inputs, we divide the excitatory neurons into three populations—Population 1 (denoted as  $P_1$ , including neurons indexed from 1 to 50), Population 2 (denoted as  $P_2$ , including neurons indexed from 51 to 200) and Population 3 (denoted as  $P_3$ , including neurons indexed from 201 to 250). The initial values of excitatory-to-excitatory (EE) connection strengths are

drawn at random from a uniform distribution ranging from 0 to  $w_{EE}^{\text{max}} = 1$  mV. During the simulation, the EE connections are subject to the canonical all-to-all pairwise additive STDP learning rule (Song et al., 2000; Babadi and Abbott, 2013). Specifically, let  $\tau = t_i - t_j$ , where  $t_i$  and  $t_j$  are a pair of spike times of excitatory neuron  $i$  and excitatory neuron  $j$ , respectively. This pair of spike times will affect a change of the synaptic strength  $w_{ij}$  from neuron  $j$  to neuron  $i$  as follows:  $w_{ij} \rightarrow w_{ij} + L(\tau)$ , where the STDP learning window  $L(\tau)$  is given by

$$L(\tau) = \begin{cases} A_+ e^{-\tau/\tau_+} & \text{for } \tau > 0, \\ -A_- e^{-|\tau|/\tau_-} & \text{for } \tau \leq 0. \end{cases} \quad (24)$$

Similarly, for the synaptic strength  $w_{ji}$  from neuron  $i$  to neuron  $j$ , the update is  $w_{ji} \rightarrow w_{ji} + L(-\tau)$ . As in Song et al. (2000)



**FIGURE 2 |** The effect of firing rate and variability on STDP learning dynamics—Case II. Different populations receive external inputs with different means as well as different variances. Population 1 (including neurons labeled by 1–50, purple) receives the smallest mean ( $\mu_{\text{ext},1}^D = 27.5$  mV) but the largest variance ( $\sigma_{\text{ext},1}^D = 31.6$  mV), Population 2 (including neurons labeled by 51–200, blue) receives an intermediate mean ( $\mu_{\text{ext},2}^D = 30$  mV) and variance ( $\sigma_{\text{ext},2}^D = 22.4$  mV), and Population 3 (including neurons labeled by 201–250, black) receives the largest mean ( $\mu_{\text{ext},3}^D = 32.5$  mV) but the smallest variance ( $\sigma_{\text{ext},3}^D = 11.2$  mV). The initial values of excitatory-to-excitatory connection strengths are drawn from a uniform distribution ranging from 0 to 1 mV. **(A,C)** Raster plots before learning and after learning, respectively. **(B,D)** Firing rates and coefficients of variation (CVs) of the different populations before learning and after  $10^6$  s of learning, respectively. Error bars indicate the standard deviation of corresponding values within each population. **(E)** Firing rate dynamics of the different populations during learning. **(F)** Weight matrix after  $10^6$  s of learning. The largest (i.e., brightest) and the smallest (i.e., darkest) value of the gray bar is 1 and 0 mV, respectively.

and Babadi and Abbott (2013), we enforce the hard bound to the value of  $w_{ij}$  and  $w_{ji}$ , i.e., set  $w_{ij}$  ( $w_{ji}$ ) to  $w_{EE}^{\text{max}}$  if  $w_{ij}$  ( $w_{ji}$ )  $>$   $w_{EE}^{\text{max}}$  and 0 if  $w_{ij}$  ( $w_{ji}$ )  $<$  0. In this work, we will restrict ourselves to the balanced case in which  $A_+ = A_- = 0.005 w_{EE}^{\text{max}}$  and  $\tau_+ = \tau_- = 20$  ms. We characterize spiking activity in the network by its firing rate and coefficient of variation (CV), where CV is the ratio of the standard deviation of interspike interval (ISI) to the mean of ISI. As is well known, for a given neuron under fixed input conditions, CV is often used as a measure of firing variability — the larger CV, the more irregular the spiking activity.

If these populations receive external inputs of different means but an identical variance, we observe that, as the mean

external input increases, the firing rate increases whereas the CV decreases. **Figure 1B** illustrates the corresponding firing statistics of these populations. For this case, as shown in **Figure 1F**, the connections from the populations of higher firing rate to those of lower firing rate are strengthened after learning whereas the connections with the opposite situation (i.e., connections from those of lower firing rate to those of higher firing rate) are weakened. This result is consistent with results of the previous studies (Morrison et al., 2007; Kozloski and Cecchi, 2010; Babadi and Abbott, 2013). Here, STDP is found to act as a firing rate equalizer, i.e., decreasing the rate of the neuronal populations of higher firing rate and increasing the rate of the neuronal populations of lower firing rate. **Figure 1E** displays an example

of this behavior. Furthermore, by comparing the firing statistics before learning (**Figure 1B**) with that after learning (**Figure 1D**), we observe that STDP also acts as a firing variability equalizer, i.e., decreasing the CV of higher-CV neuronal populations and increasing the CV of lower-CV neuronal populations. For ease of discussion below, this case will be termed as Case I.

Next, we consider the case in which both the mean and variance of external inputs vary across these populations. We adjust both the mean and variance of external inputs (see the caption of **Figure 2** for the specific values) such that Population 1 has the highest firing rate and the highest CV, Population 2 has an intermediate firing rate and CV, and Population 3 has the lowest firing rate and the lowest CV. The corresponding firing statistics of these populations before learning is displayed in **Figure 2B**. From **Figure 2F**, it can be seen clearly that after learning the connections from Population 1, which has the highest firing rate, to Population 3, which has the lowest firing rate, are weakened. We note that this phenomenon still exists even when the projections from the excitatory populations to the inhibitory populations become plastic (Supplementary Material). Furthermore, by comparing **Figure 2B** with **Figure 2D**, we observe that STDP tends to increase the rate of the population of the highest firing rate and decrease the rate of the population of the lowest firing rate. That is, STDP is no longer acting as a firing rate equalizer. However, it still acts as a weak equalizer for firing variability. This phenomenon cannot be explained solely by the firing rate theory (Morrison et al., 2007; Kozloski and Cecchi, 2010; Babadi and Abbott, 2013), in which the connections from the population of higher firing rate to the population of lower firing rate should be strengthened and STDP should act as a firing rate equalizer. In the following, we develop a linear response theory to characterize the importance of high-order firing statistics in learning dynamics as illustrated in **Figure 2**. For ease of discussion below, the case in **Figure 2** will be termed as Case II.

## STDP Learning Dynamics

To characterize the learning dynamics, we consider two mutually-connected neurons in a recurrent network, say neuron  $i$  and neuron  $j$ . The STDP learning rule introduced in the last section can be simply written in the following weight dynamics:

$$\frac{dw_{ij}}{dt} = A_+ x_j^+(t) s_i(t) - A_- x_i^-(t) s_j(t), \quad (25)$$

$$\frac{dx_j^+}{dt} = -\frac{x_j^+}{\tau_+} + s_j(t), \quad (26)$$

$$\frac{dx_i^-}{dt} = -\frac{x_i^-}{\tau_-} + s_i(t), \quad (27)$$

where  $s_i(t)$  and  $s_j(t)$  respectively stand for the spike trains of neuron  $i$  and neuron  $j$ . Expressing  $x_j^+$  and  $x_i^-$  as the integral of  $s_j(t)$  and  $s_i(t)$ , respectively, we can obtain

$$\frac{dw_{ij}}{dt} = \int_{-\infty}^{\infty} L(\tau) s_j(t - \tau) s_i(t) d\tau. \quad (28)$$

By assuming that the change of  $w_{ij}(t)$  is sufficiently slow (Ocker et al., 2015),  $s_j(t - \tau) s_i(t)$  can be approximated by  $C_{ij}(\tau) + \langle s_i(t) \rangle \langle s_j(t) \rangle$  where  $C_{ij}(\tau)$ , defined as  $\langle s_i(t + \tau) s_j(t) \rangle - \langle s_i(t) \rangle \langle s_j(t) \rangle$ , is the temporal spike cross correlation between neuron  $i$  and neuron  $j$  in the recurrent network with a fixed synaptic strength matrix. Combining the dynamics of  $w_{ij}$  in Equation (25), we arrive at the following equations:

$$\frac{dw_{ij}}{dt} = \int_{-\infty}^{\infty} L(\tau) C_{ij}(\tau) d\tau, \quad (29)$$

$$\frac{dw_{ji}}{dt} = \int_{-\infty}^{\infty} L(\tau) C_{ji}(\tau) d\tau. \quad (30)$$

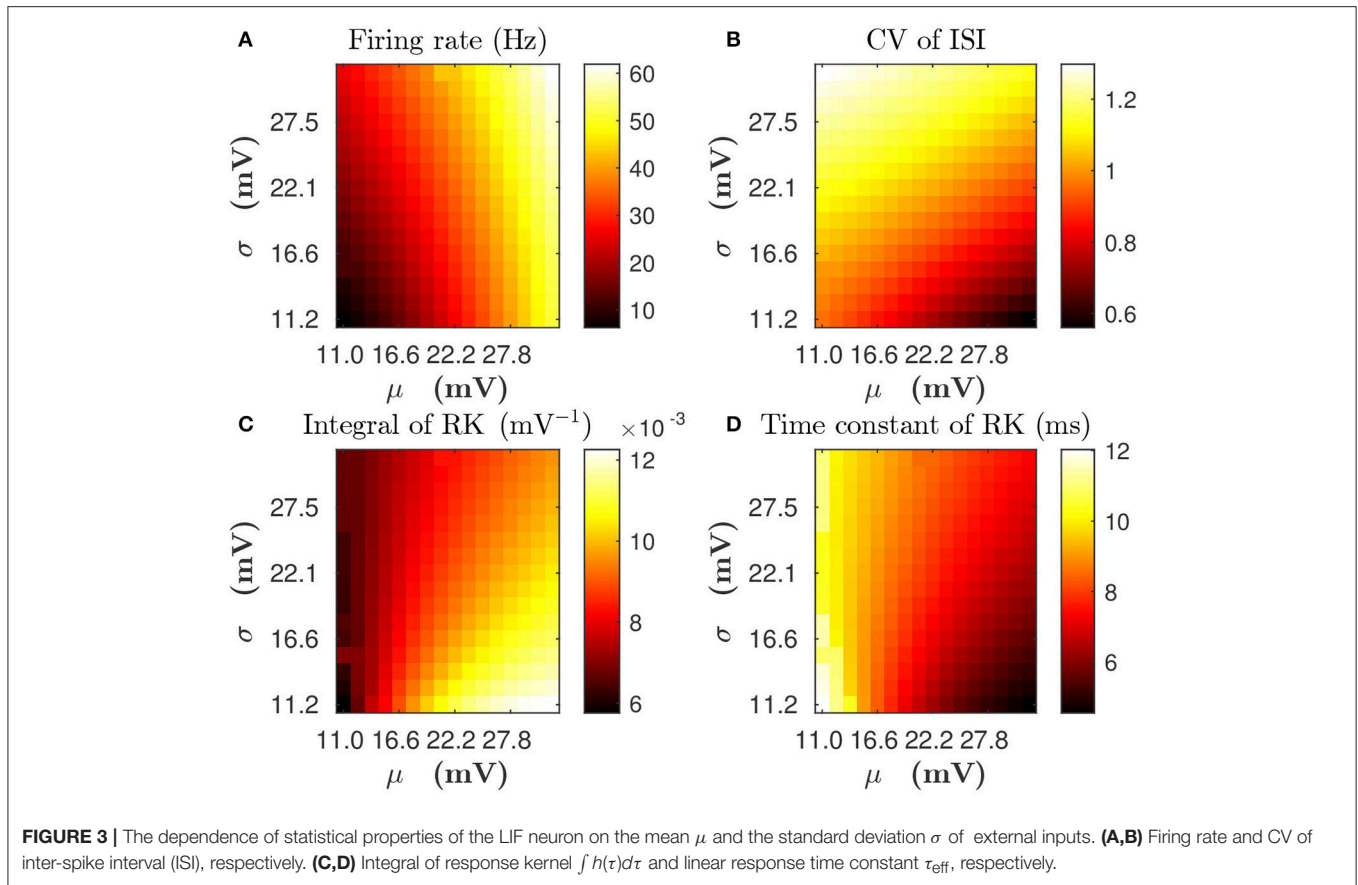
Note that unlike the scenario studied in Babadi and Abbott (2013) and Ocker et al. (2015), there is no offset term in Equations (29, 30). This is due to the balanced condition  $A_+ = A_-$  and  $\tau_+ = \tau_-$  in our STDP learning rule. Therefore, to understand the learning dynamics (Equations 29, 30), it is facilitative to study the correlation structure in recurrent networks with static couplings.

In the following sections, we will follow the linear response theory proposed in Pernice et al. (2012), Trousdale et al. (2012), and Helias et al. (2013) and develop a self-consistent framework for the learning dynamics (Equations 29, 30). To this end, we first characterize how the firing variability information is encoded in the linear response kernel for the single LIF neuron case in section Incorporation of Firing Variability in Linear Response Kernel. Then, in section Hawkes Process as an Approximation of an LIF Network Dynamics, we approximate the LIF network dynamics with a stochastic point process called Hawkes process in which the instantaneous firing rate depends on the spiking history of interacting neurons (Hawkes, 1971b). With this Hawkes process approximation, in section Spike Correlation Structure of Hawkes Process, we discuss the issue of how firing statistics of spiking activity and network connectivity structure shape the spike correlation structure of the LIF network dynamics. In the end, we combine the correlation structure analysis with the STDP learning dynamics (Equations 29, 30) to arrive at the final learning dynamics (Equations 35, 36), which explain the observed numerical results, in section Learning Dynamics with Spike Correlation Decomposition.

## Incorporation of Firing Variability in Linear Response Kernel

For a single LIF neuron described by Equations (1, 3), the dependence of mean firing rate and CV on input parameters (mean input value  $\mu$  and standard deviation value  $\sigma$ ) is illustrated in **Figures 3A,B**, respectively. By using the reverse correlation method (Dayan and Abbott, 2001; Ostojic et al., 2009) (see section Methods for details), We also computed the response kernel  $h(\tau)$  that characterizes how the LIF neuron ensemble responds to a perturbation. We found that response kernel  $h(\tau)$  exhibits an exponential decay profile,  $A \exp\{-\tau/\tau_{\text{eff}}\} \Theta(\tau)$ , where  $\Theta(\tau)$  is the Heaviside function, provided that fluctuations of the external input are sufficiently high, i.e., comparable to the difference between firing threshold and reset voltage  $V_{\text{th}} - V_r$  (for example,  $\sigma$  is about 10 ~ 30 mV). We can use two quantities — the integral of response kernel  $\int h(\tau) d\tau$ ,





which is  $\tau_{\text{eff}}A$  after the substitution of the exponential form of  $h(\tau) = A \exp\{-\tau/\tau_{\text{eff}}\}\Theta(\tau)$ , and the decay time constant  $\tau_{\text{eff}}$  — to characterize  $h(\tau)$ . **Figures 3C,D** illustrate the dependence of  $\int h(\tau)d\tau$  and  $\tau_{\text{eff}}$  on  $\mu$  and  $\sigma$ , respectively. In **Figure 3**, we can observe that, for any fixed mean firing rate, when the standard deviation  $\sigma$  increases, both CV of ISI and  $\tau_{\text{eff}}$  increase whereas  $\int h(\tau)d\tau$  decreases. Therefore, if there are two neurons with the same mean firing rate but different firing CVs, both the decay time constant and the integral of response kernel of these two neurons will differ from each other. In this sense, the information of firing variability is incorporated in the linear response kernel.

## Hawkes Process as an Approximation of an LIF Network Dynamics

In general, the response of a neuron to recurrent spike inputs is nonlinear. Here, we investigate the asynchronous state, as shown in **Figures 1A, 2A**, in which spikes from different neurons arrive at the synapses of the target neuron at different times. By considering that the magnitude of a postsynaptic potential elicited by the firing event of a single presynaptic neuron is in general small, i.e., the synaptic strength is weak, it is reasonable to treat the asynchronous recurrent inputs in a linear manner (Pernice et al., 2012; Trousdale et al., 2012; Helias et al., 2013). Specifically, by regarding  $\sum_{j=1}^N w_{ij}(s_j - \bar{r}_j)$  in the LIF neuronal

network as a perturbation, we arrive at the following effective description:

$$r_i(t) = \bar{r}_i + [h_i * \sum_{j=1}^N w_{ij}(s_j - \bar{r}_j)](t), \quad (31)$$

where  $\bar{r}_i$  and  $h_i(\tau)$  are the mean firing rate and response kernel of the  $i$ th neuron that is driven by both the external input and the mean-driven recurrent input (see section Methods for details), respectively, and  $N$  is the number of neurons in the network. Note that (i) the mean firing rate  $\bar{r}_i$  is determined by the self-consistent Equations (12, 13); and (ii) the response kernel  $h_i(\tau)$  is computed based on the single LIF neuron Equations (1, 3) with the total mean input value  $\mu_i^p$  in Equation (13) and the standard deviation value  $\sigma_{\text{ext},i}^p$  in Equation (12). Equation (31) can be understood in the following way. On the one hand, given spike trains  $s_j(t)$  for  $j = 1, \dots, N$ , Equation (31) provides a way to evolve the instantaneous firing rate of the  $i$ th neuron  $r_i(t)$ . On the other hand, from Equation (31), the spike train  $s_i(t)$  can be regarded as being generated by a jump process with the instantaneous firing rate  $r_i(t)$ . Therefore, we have obtained a self-consistent stochastic jump process that produces both the instantaneous firing rate  $r_i(t)$  and the stochastic spike train  $s_i(t)$ . This stochastic process Equation (31) is a Hawkes process (Hawkes, 1971b), which has been extensively used in



the neuroscience literature (Pernice et al., 2012; Trousdale et al., 2012; Helias et al., 2013). From the discussion above, such a Hawkes process can be viewed as the first-order (i.e., linear) Volterra approximation of the LIF network dynamics.

## Spike Correlation Structure of Hawkes Process

For the Hawkes process (Equation 31), its spike correlation has an explicit analytical form (Hawkes, 1971a). Specifically, the power spectrum  $\hat{C}(\omega)$ , by the Wiener-Khinchin theorem, is the Fourier transform of spike correlation  $C(\tau) = (C_{ij}(\tau))$  and has the following form (Hawkes, 1971a) (see section Methods for a derivation):

$$\hat{C}(\omega) = (1 - \hat{\mathbf{h}}(\omega)\mathbf{w})^{-1}\mathbf{D}(1 - \mathbf{w}^T\hat{\mathbf{h}}^T(-\omega))^{-1}, \quad (32)$$

where  $\hat{\mathbf{h}}(\omega)$  is the Fourier transform of the response kernel  $\mathbf{h}(\tau) = \text{diag}\{h_1(\tau), \dots, h_N(\tau)\}$ ,  $\mathbf{w}$  is the synaptic strength matrix, and  $\mathbf{D} = \text{diag}\{\bar{r}_1, \dots, \bar{r}_N\}$ .

To examine the validity of our linear response theory, we compute the population-averaged cross correlation (see section Methods for details). We find that our theoretical prediction of the population-averaged cross correlation is in very good agreement with the cross correlation measured from our LIF neuronal network simulation. This agreement is clearly seen in **Figure 4**, demonstrating that our linear response theory can well capture the spike correlation structure of the LIF network dynamics.

We note that Equation (32) also provides a starting point to study how network connectivity affects the spike correlation structure since we can decompose the connections into direct connections, common inputs, and other higher-order connections. In particular, the spike correlation has the following decomposition (see section Methods for a derivation):

$$C(\tau) = \mathbf{D}\delta(\tau) + \mathbf{h}(\tau)\mathbf{w}\mathbf{D} + \mathbf{D}\mathbf{w}^T\mathbf{h}_-^T(\tau) + (\mathbf{h}\mathbf{w}\mathbf{D}\mathbf{w}^T * \mathbf{h}_-^T)(\tau) + \dots, \quad (33)$$

where  $\mathbf{h}_-(\tau) \equiv \mathbf{h}(-\tau)$ . Note that the similar decomposition has been derived in Pernice et al. (2012), Trousdale et al. (2012), Hu et al. (2013), and Ocker et al. (2015).

The first term on the right hand side (RHS) of Equation (33) can be obtained by setting the connection strength  $\mathbf{w}$  in the Hawkes process (Equation 31) to zero. There is no cross correlation in this term.

For the second term on the RHS of Equation (33), the corresponding component with respect to the correlation between neuron  $i$  and neuron  $j$  is  $h_i(\tau)w_{ij}\bar{r}_j$ . This involves firing rate  $\bar{r}_j$  of presynaptic neuron  $j$ , the connection strength  $w_{ij}$  from neuron  $j$  to neuron  $i$  and the linear response kernel  $h_i(\tau)$  of postsynaptic neuron  $i$ . Therefore, this component corresponds to the correlation induced by the direct connection from neuron  $j$  to neuron  $i$ . Similarly, the component  $h_j(-\tau)w_{ji}\bar{r}_i$  (the third term on the RHS of Equation 33), corresponds to the correlation induced by the direct connection from neuron  $i$  to neuron  $j$ . Combining these two terms, we obtain the spike correlation induced by direct connections between neuron  $i$  and neuron  $j$ . For ease of discussion below, we denote these terms as  $C_{ij}^{\text{di}}(t)$ .

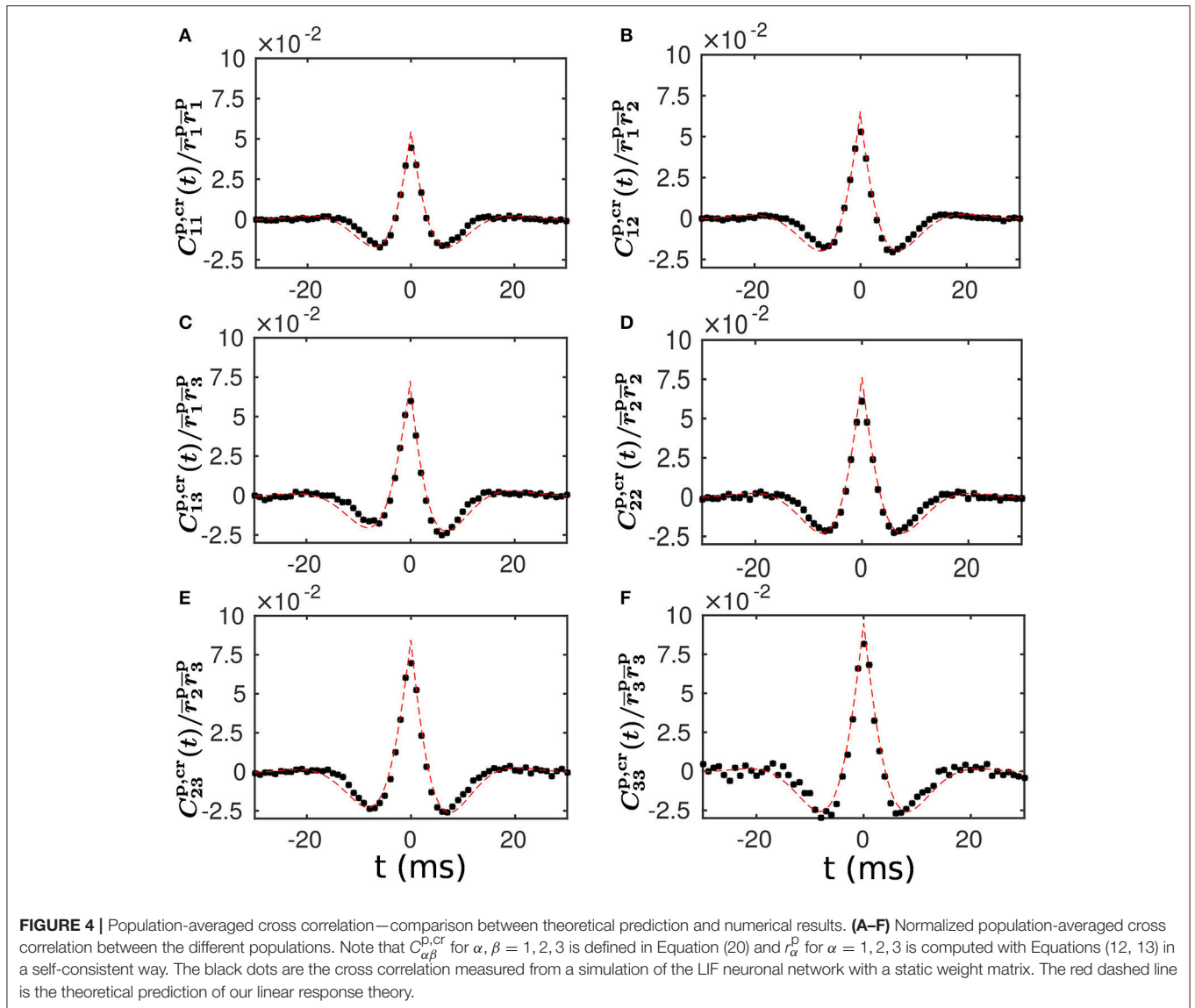
We next discuss the terms associated with the indirect connections. For the fourth term on the RHS of Equation (33), the corresponding component with respect to the correlation between neuron  $i$  and neuron  $j$  is  $\sum_k w_{ik}\bar{r}_k w_{jk}[h_i * h_{-j}](\tau)$ . This involves the firing rate  $\bar{r}_k$  of neuron  $k$ , the connection strengths  $w_{\alpha k}$  and the response kernels  $h_\alpha(\cdot)$  for  $\alpha = i, j$ . Therefore, this term corresponds to the correlation induced by the common inputs from neuron  $k$  to both neuron  $i$  and neuron  $j$ . As we have demonstrated above,  $h_\alpha(\cdot)$ ,  $\alpha = i, j$ , exhibits an exponential decay profile, which will be expressed as  $B_\alpha \exp(-t/\tau_\alpha)\Theta(t)$ . Then, this component becomes

$$\sum_k B_i B_j \frac{\tau_i \tau_j}{\tau_i + \tau_j} w_{ik} \bar{r}_k w_{jk} \begin{cases} \exp(-t/\tau_i) & \text{for } t \geq 0, \\ \exp(t/\tau_j) & \text{for } t < 0. \end{cases} \quad (34)$$

For ease of discussion below, we denote this term as  $C_{ij}^{\text{co}}(t)$ , the remaining higher-order terms in Equation (33) as  $C_{ij}^{\text{hi}}(t)$ , and the summation of  $C_{ij}^{\text{co}}(t)$  and  $C_{ij}^{\text{hi}}(t)$  as the indirect term  $C_{ij}^{\text{in}}(t)$ . We can observe from Equation (34) that, if there is a difference between response time constants  $\tau_i$  and  $\tau_j$ ,  $C_{ij}^{\text{co}}(t)$  will show an asymmetric property with respect to the axis  $t = 0$ . More precisely, if we assume  $\tau_i > \tau_j$ ,  $C_{ij}^{\text{co}}(t) > C_{ij}^{\text{co}}(-t)$  for any  $t > 0$ . This asymmetry can be understood intuitively as follows. When there is a common input from neuron  $k$  into both neurons  $i$  and  $j$ , the neuron with the smaller response time constant (neuron  $j$ ) will respond more rapidly than the neuron with the larger response time constant (neuron  $i$ ) on average. Therefore, there is a higher likelihood that neuron  $j$  will fire before neuron  $i$ . In terms of the spike correlation, this implies that  $C_{ij}^{\text{co}}(t) > C_{ji}^{\text{co}}(t) = C_{ij}^{\text{co}}(-t)$  for any  $t > 0$ . We note that, as shown in Equation (34), the common inputs from both excitatory and inhibitory neurons into neuron  $i$  and neuron  $j$  induce a positive spike correlation between neuron  $i$  and neuron  $j$ . Therefore, if there is a significant difference between  $\tau_i$  and  $\tau_j$ , there will be a substantial asymmetry between the positive- $t$  component of  $C_{ij}^{\text{co}}(t)$  and the corresponding negative- $t$  component. This will have a profound impact on the structure of  $C_{ij}(t)$  as discussed below.

Now we study how different connectivity patterns affect the cross correlation structure between different neuronal populations for Cases I and II in **Figures 1, 2**. As an example, we pick neuron  $i$  from Population 3 and neuron  $j$  from Population 1. Note that, for Case I in **Figure 1**, neuron  $i$  ( $j$ ) has a lower (higher) firing rate and a higher (lower) firing variability. Whereas, for Case II in **Figure 2**, neuron  $i$  ( $j$ ) has a lower (higher) firing rate and a lower (higher) firing variability.

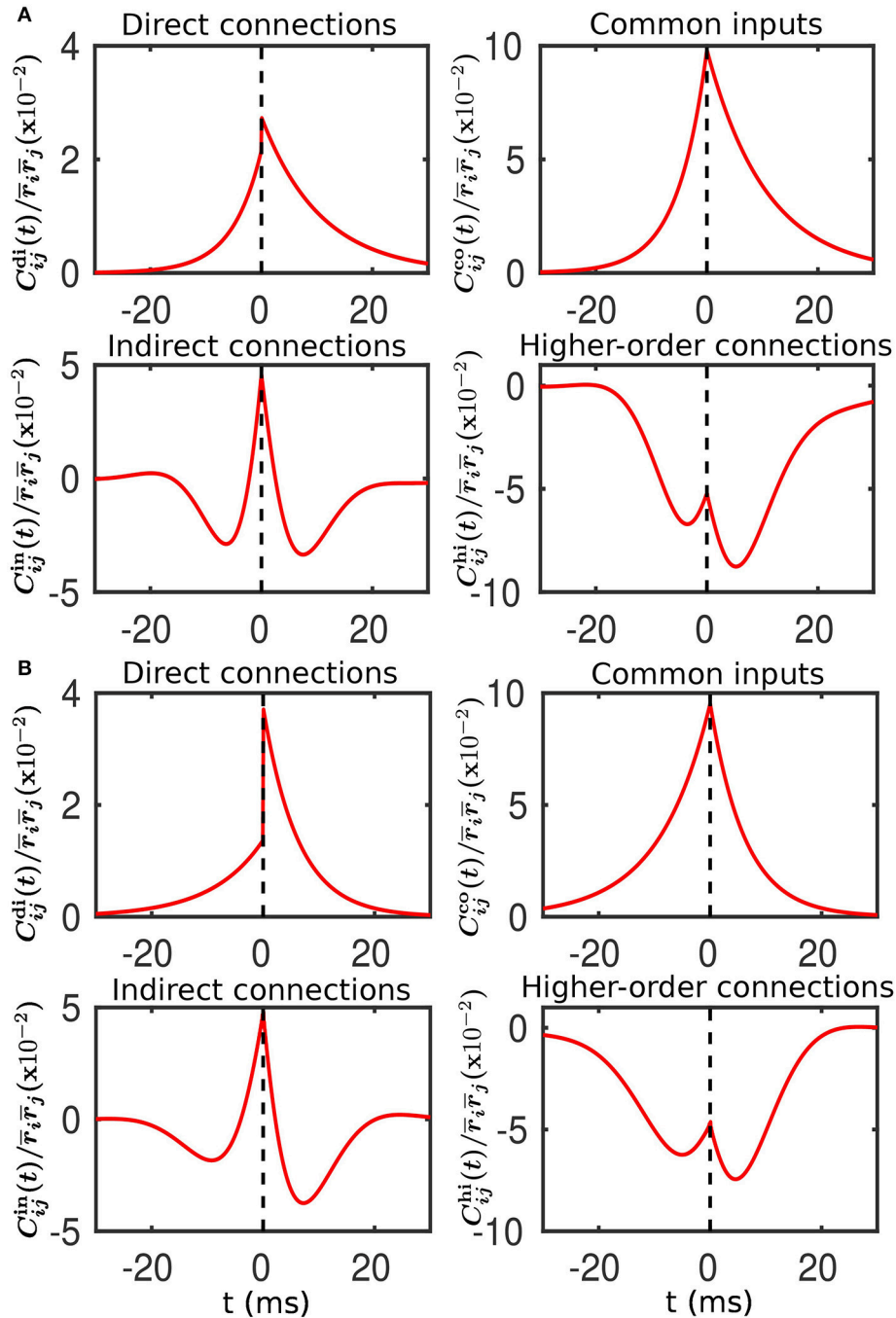
For Case I in **Figure 1**, we find that the net mean inputs to neuron  $j$  in Population 1 and neuron  $i$  in Population 3 are 31.45 and 11.55 mV, respectively. Note that the external mean inputs  $\mu_{\text{ext},j}^{\text{p}}$  and  $\mu_{\text{ext},i}^{\text{p}}$  are 40 and 20 mV, respectively. Since the net mean input is the summation of external mean input and recurrent input, the net recurrent inputs to neuron  $j$  and neuron  $i$  are  $-8.55$  mV and  $-8.45$  mV, respectively. This means the overall recurrent inputs to both neurons are inhibitory. In Case I, the standard deviation values for both neurons are 15.81 mV. With these net



mean input values and standard deviation values, we compute the response kernel by using the reverse correlation method (see section Methods for more details). We find that  $\int h_j(t)dt > \int h_i(t)dt$  and  $\tau_j < \tau_i$ , which is evident in **Figures 3C,D**. From the left panels of **Figure 5A**, we can observe that (i) the positive- $t$  component of  $C_{ij}^{di}(t)$  is larger than the corresponding negative- $t$  component; and (ii) the positive- $t$  component of  $C_{ij}^{in}(t)$  is slightly smaller than the corresponding negative- $t$  component. To further understand how different connectivity patterns shape  $C_{ij}^{in}(t)$ , we split  $C_{ij}^{in}(t)$  into the common-input correlation  $C_{ij}^{co}(t)$  and the higher-order-connection correlation  $C_{ij}^{hi}(t)$ . It can be seen clearly from the right panels of **Figure 5A** that the positive- $t$  component of  $C_{ij}^{hi}(t)$  is smaller than the corresponding negative- $t$  component whereas the positive- $t$  component of  $C_{ij}^{co}(t)$  is larger than the corresponding negative- $t$  component. A combination of these two facts yields a small difference between the

positive- $t$  component of  $C_{ij}^{in}(t)$  and the corresponding negative- $t$  component. We note that the asymmetry of  $C_{ij}^{co}(t)$  with respect to the axis  $t = 0$  exhibited in the top right panel of **Figure 5A** is a consequence of  $\tau_j < \tau_i$ .

For Case II in **Figure 2**, we find that the net mean inputs to neuron  $j$  in Population 1 and neuron  $i$  in Population 3 are 15.58 and 20.50 mV, respectively. Note that the external mean inputs  $\mu_{ext,j}^p$  and  $\mu_{ext,i}^p$  are 27.5 and 32.5 mV, respectively. Since the net mean input is the summation of external mean input and recurrent input, the net recurrent inputs to neuron  $j$  and neuron  $i$  are  $-11.92$  and  $-12.00$  mV, respectively. This means the overall recurrent inputs to both neurons also are inhibitory. In Case II, the standard deviation values for neuron  $j$  and neuron  $i$  are 31.6 and 11.1 mV, respectively. With these net mean input values and standard deviation values, we compute the response kernel by using the reverse correlation method (see section Methods for



**FIGURE 5** | Decomposition of the spike cross correlation between neurons from the different populations before learning. Here, we choose neuron  $i$  from Population 3 and neuron  $j$  from Population 1. The connection strengths  $w_{ij}$  and  $w_{ji}$  are fixed at 0.5 mV. The top left, top right, bottom left and bottom right panels of both **(A,B)** correspond to the correlations associated with direct connections, common inputs, indirect connections, and higher-order connections, respectively. **(A)** Compared to neuron  $i$ , neuron  $j$  receives a larger mean and an identical variance of external inputs. This corresponds to Case I in **Figure 1**. **(B)** Compared to neuron  $i$ , neuron  $j$  receives a smaller mean but a larger variance of external inputs. This corresponds to Case II in **Figure 2**.

more details). We find that  $\int h_j(t)dt < \int h_i(t)dt$  and  $\tau_j > \tau_i$ , which is also evident in **Figures 3C,D**. From the left panels of **Figure 5B**, we can observe that (i) the positive- $t$  component of  $C_{ij}^{di}(t)$  is again larger than the corresponding negative- $t$

component; and (ii) the positive- $t$  component of  $C_{ij}^{in}(t)$  is much smaller than the corresponding negative- $t$  component. From the right panels of **Figure 5B**, we find that the positive- $t$  components of both  $C_{ij}^{co}(t)$  and  $C_{ij}^{hi}(t)$  are smaller than the corresponding

negative- $t$  components. A combination of these two effects yields a large difference between the positive- $t$  component of  $C_{ij}^{\text{in}}(t)$  and the corresponding negative- $t$  component. Note that the fact that  $C_{ij}^{\text{co}}(t) < C_{ij}^{\text{co}}(-t)$  for  $t > 0$  in **Figure 5B** is a consequence of the inequality  $\tau_j > \tau_i$ . In the next section, we will demonstrate the extent of asymmetry between the positive- $t$  component of  $C_{ij}^{\text{in}}(t)$  and the corresponding negative- $t$  component underlies the difference of network structures learned through STDP presented in **Figures 1, 2**.

## Learning Dynamics with Spike Correlation Decomposition

Combining the STDP learning dynamics (Equations 29, 30) with the direct- and indirect-connection decomposition of the correlation structure in the previous section, we arrive at the following learning dynamics:

$$\frac{dw_{ij}}{dt} = A_i w_{ij} \bar{r}_j - A_j w_{ji} \bar{r}_i + B, \quad (35)$$

$$\frac{dw_{ji}}{dt} = A_j w_{ji} \bar{r}_i - A_i w_{ij} \bar{r}_j - B, \quad (36)$$

where  $A_\alpha = \int_0^\infty A_+ \exp\{-\tau/\tau_+\} h_\alpha(\tau) d\tau$  for  $\alpha = i, j$  and  $B = \int L(\tau) C_{ij}^{\text{in}}(\tau) d\tau$ . In Equations (35, 36), the first and second terms on the RHS correspond to the effect of direct connections between neuron  $i$  and neuron  $j$ , whereas the third term corresponds to the effect of indirect connections between neuron  $i$  and neuron  $j$ . Note that here we have used the balanced STDP assumption, i.e.,  $A_+ = A_-$  and  $\tau_+ = \tau_-$ .

As argued in Babadi and Abbott (2013), while a linear system of differential equations such as Equations (35, 36) cannot have more than two attractors, the hard lower and upper bounds enforced by the STDP learning rule, i.e., the condition  $0 \leq w_{ij} \leq w_{EE}^{\text{max}}$ , can lead to the following two attractors:  $(w_{ij}, w_{ji}) = (w_{EE}^{\text{max}}, 0)$  and  $(w_{ij}, w_{ji}) = (0, w_{EE}^{\text{max}})$ . For the sake of discussion, we consider  $A_i \bar{r}_j > A_j \bar{r}_i$  below. Then, the line  $w_{ji} = \frac{1}{A_j \bar{r}_i} (A_i \bar{r}_j w_{ij} + B)$  has the slope greater than 1, as illustrated by the black dash line in **Figure 6**. If neglecting the indirect connection effect, that is, setting  $B = 0$ , we can observe from Equations (35, 36) that, for any  $(w_{ij}, w_{ji})$  pair satisfying  $A_i w_{ij} \bar{r}_j - A_j w_{ji} \bar{r}_i > 0$ , the pair will converge to  $(w_{EE}^{\text{max}}, 0)$ . Whereas, for any  $(w_{ij}, w_{ji})$  pair satisfying  $A_i w_{ij} \bar{r}_j - A_j w_{ji} \bar{r}_i < 0$ , the pair will converge to  $(0, w_{EE}^{\text{max}})$ . This is illustrated in **Figure 6A** with the black dash line separating the two basins. In this case, the size of basin of the attractor  $(w_{EE}^{\text{max}}, 0)$  is larger than that of the attractor  $(0, w_{EE}^{\text{max}})$ . Therefore,  $w_{ij}$  would be more likely to be potentiated, whereas  $w_{ji}$  would be more likely to be depressed.

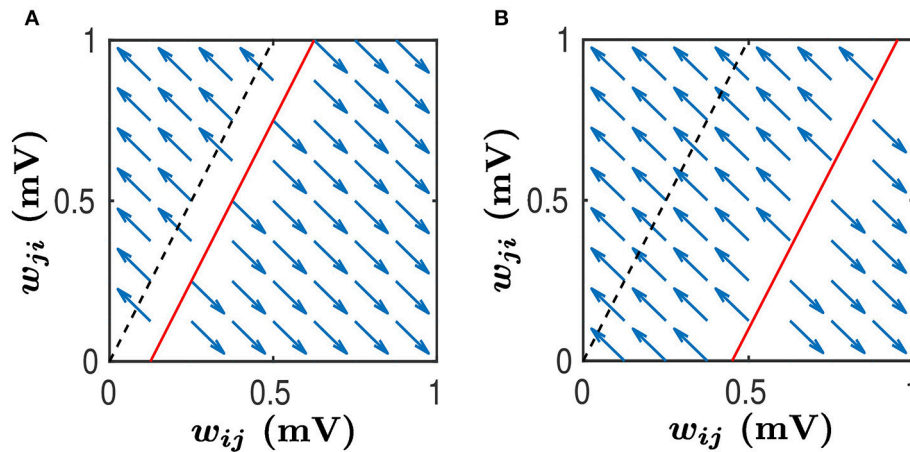
When the indirect connection effect is included, there exist two distinct regimes depending on the value of  $B$ . For the sake of discussion, we denote the size of basin of  $(w_{EE}^{\text{max}}, 0)$  (i.e., the area below the line  $A_i w_{ij} \bar{r}_j - A_j w_{ji} \bar{r}_i + B = 0$ ) as  $S_{ij}$  and the size of basin of  $(0, w_{EE}^{\text{max}})$  (i.e., the area above the line  $A_i w_{ij} \bar{r}_j - A_j w_{ji} \bar{r}_i + B = 0$ ) as  $S_{ji}$ . One is a regime in which  $S_{ij} > S_{ji}$ . This occurs when  $B \geq 0$  or  $B$  is negative but small. As illustrated in **Figure 6A**, the red line separates the two basins. In this regime,  $w_{ij}$  would be more likely to be potentiated whereas  $w_{ji}$  would be more likely

to be depressed. For neuronal populations, the synapses from the neuron  $j$ 's population to neuron  $i$ 's population will have a higher likelihood to be strengthened. In contrast, the synapses with the opposite direction will be more likely to be weakened. The other is a regime in which  $S_{ij} < S_{ji}$ . This occurs when  $B$  is sufficiently negative, as indicated by the red line in **Figure 6B**. In this regime,  $w_{ij}$  would be more likely to be depressed whereas  $w_{ji}$  would be more likely to be potentiated. For neuronal populations, the synapses from neuron  $i$ 's population to neuron  $j$ 's population will have a higher likelihood to be strengthened, whereas the synapses with the opposite direction will be weakened with a larger probability.

We now examine the cases in **Figures 1, 2**. For Case I in **Figure 1** and **Table 1** displays the corresponding values of  $A_i \bar{r}_j$ ,  $A_j \bar{r}_i$ ,  $B$ ,  $S_{ij}$ ,  $S_{ji}$  as well as the average connection strengths  $\bar{w}_{ij}$  and  $\bar{w}_{ji}$  after  $10^6$  s of learning. It can be seen clearly that for all the three  $(i, j)$  pairs, namely  $(i \in P_2, j \in P_1)$ ,  $(i \in P_3, j \in P_1)$  and  $(i \in P_3, j \in P_2)$ ,  $S_{ij} > 0.5 \text{mV}^2 > S_{ji}$ , indicating that  $\bar{w}_{ij}$  for all these three pairs would be strengthened. Indeed, this strengthening is confirmed by the values of  $\bar{w}_{ij}$  presented in **Table 1** (i.e.,  $\bar{w}_{ij} > 0.5 \text{mV} > \bar{w}_{ji}$  for all three pairs). To evaluate the contribution of common inputs and higher-order connections separately, we also include  $B^{\text{co}}$  (defined as  $\int L(\tau) C_{ij}^{\text{co}}(\tau) d\tau$ ) and  $B^{\text{hi}}$  (defined as  $\int L(\tau) C_{ij}^{\text{hi}}(\tau) d\tau$ ) in **Table 1**. We can observe that for all three pairs  $B^{\text{co}} > 0$  whereas  $B^{\text{hi}} < 0$ . Since  $B = B^{\text{co}} + B^{\text{hi}}$ , this gives rise to a negative but small value of  $B$ , yielding a regime in which  $S_{ij} > S_{ji}$ . Since the connections from the populations of higher firing rate to those of lower firing rate are strengthened whereas the connections with the opposite direction are weakened, the rate of the populations of lower firing rate will increase whereas that of the populations of higher firing rate will decrease. The strengthening of connections from the populations of higher firing rate to those of lower firing rate leads to the increase of the mean of total inputs into the populations of lower firing rate. As illustrated in **Figure 3B**, when the variance of total inputs is fixed, the larger the mean total inputs, the smaller the CV. Therefore, this explains why in **Figure 1** the CV of the populations of lower firing rate, which is larger before learning, decreases after learning. Therefore, STDP here acts as a weak equalizer for both firing rate and firing variability, as demonstrated in **Figure 1**.

For Case II in **Figure 2** and **Table 2** shows that for all three  $(i, j)$  pairs, namely  $(i \in P_2, j \in P_1)$ ,  $(i \in P_3, j \in P_1)$  and  $(i \in P_3, j \in P_2)$ ,  $S_{ij} < 0.5 \text{mV}^2 < S_{ji}$  and  $\bar{w}_{ij} < 0.5 \text{mV} < \bar{w}_{ji}$ . Since the connections from the populations of higher firing rate to the populations of lower firing rate are weakened whereas the connections with the opposite direction are strengthened, the rate of the populations of higher firing rate will increase further whereas that of the populations of lower firing rate will decrease further. Meanwhile, the weakening of connections from the populations of higher firing rate to those of lower firing rate leads to the decrease of the mean of total inputs into the populations of lower firing rate. As illustrated in **Figure 3B**, the smaller the mean total inputs, the larger the CV. Therefore, the CV of these populations of lower firing rate, which is smaller before learning, increases after learning. Note that, in Case II, those populations of lower firing rate also have the lower firing





**FIGURE 6** | Phase plane analysis of learning dynamics. This learning dynamics has two attractors —  $(w_{ij}, w_{ji}) = (1, 0)$  and  $(w_{ij}, w_{ji}) = (0, 1)$ . The black dashed line satisfies  $A_i w_{ji} r_j - A_j w_{ij} r_i = 0$  while the red line satisfies  $A_i w_{ji} r_j - A_j w_{ij} r_i - B = 0$ . We assume that  $A_i r_j > A_j r_i$ . Therefore, the slope of both the black dashed line and the red line is greater than 1. Any  $(w_{ij}, w_{ji})$  pair in the area above the red line will converge to  $(0, 1)$  whereas any pair in the area below the red line will converge to  $(1, 0)$ , as indicated by the blue arrows. Therefore, the area above the red line is the basin of the attractor  $(0, 1)$  whereas the area below the red line is the basin of the attractor  $(1, 0)$ . Note that  $w_{ij}$  and  $w_{ji}$  evolve along the line  $w_{ij} + w_{ji} = \text{Const}$  as indicated by the arrows. **(A)** The size of basin of the attractor  $(0, 1)$  is smaller than that of the attractor  $(1, 0)$ . In this case, synapses from the neuronal population to which neuron  $j$  belongs, to the neuronal population to which neuron  $i$  belongs, would be more likely to be strengthened, whereas the synapses with the opposite direction will be weakened. **(B)** The size of basin of the attractor  $(0, 1)$  is larger than that of the attractor  $(1, 0)$ . In this case, synapses from the neuronal population to which neuron  $j$  belongs, to the neuronal population to which neuron  $i$  belongs, would be more likely to be weakened, whereas the synapses with the opposite direction will be strengthened.

**TABLE 1** | Parameters of learning dynamics (Equations 35, 36) and related quantities for **Figure 1**.

	$A_i r_j \times 10^6$ ( $\text{ms}^{-1}$ )	$A_j r_i \times 10^6$ ( $\text{ms}^{-1}$ )	$B \times 10^6$ ( $\text{ms}^{-1} \cdot \text{mV}$ )	$B^{\text{co}} \times 10^6$ ( $\text{ms}^{-1} \cdot \text{mV}$ )	$B^{\text{hi}} \times 10^6$ ( $\text{ms}^{-1} \cdot \text{mV}$ )	$S_{ij}$ ( $\text{mV}^2$ )	$\bar{w}_{ij}$ ( $\text{mV}$ )	$\bar{w}_{ji}$ ( $\text{mV}$ )
$(i \in P_2, j \in P_1)$	1.75	1.22	-0.19	0.47	-0.66	0.54	0.68	0.32
$(i \in P_3, j \in P_1)$	1.12	0.54	-0.18	0.80	-0.98	0.60	0.81	0.19
$(i \in P_3, j \in P_2)$	0.66	0.45	-0.04	0.44	-0.48	0.59	0.70	0.30

**TABLE 2** | Parameters of learning dynamics (Equations 35, 36) and related quantities for **Figure 2**.

	$A_i r_j \times 10^6$ ( $\text{ms}^{-1}$ )	$A_j r_i \times 10^6$ ( $\text{ms}^{-1}$ )	$B \times 10^6$ ( $\text{ms}^{-1} \cdot \text{mV}$ )	$B^{\text{co}} \times 10^6$ ( $\text{ms}^{-1} \cdot \text{mV}$ )	$B^{\text{hi}} \times 10^6$ ( $\text{ms}^{-1} \cdot \text{mV}$ )	$S_{ij}$ ( $\text{mV}^2$ )	$\bar{w}_{ij}$ ( $\text{mV}$ )	$\bar{w}_{ji}$ ( $\text{mV}$ )
$(i \in P_2, j \in P_1)$	0.92	0.71	-0.15	-0.11	-0.04	0.45	0.45	0.55
$(i \in P_3, j \in P_1)$	1.28	0.63	-0.55	-0.53	-0.02	0.32	0.30	0.70
$(i \in P_3, j \in P_2)$	1.15	0.73	-0.40	-0.45	0.05	0.34	0.32	0.68

variability. This explains why STDP acts as a weak equalizer of firing variability, instead of firing rate, in **Figure 2**.

Furthermore, for Case II in **Figure 2**, by examining the contribution of common inputs and higher-order connections, we can find from **Table 2** that, compared to  $B^{\text{co}}$ ,  $B^{\text{hi}}$  is negligible for the pairs  $(i \in P_2, j \in P_1)$  and  $(i \in P_3, j \in P_1)$ . Therefore, for these two pairs, it is the contribution of common-input correlations that yields a large negative value of  $B$  and the corresponding network structures in **Figure 2**. Note that the asymmetric property of common-input correlations arises from the difference of response time constants between different neurons. As demonstrated in **Figure 3**, the response time constant encodes the information of the firing variability of

spiking activity. In this sense, we have established a direct link between firing variability and learning dynamics by analyzing the spike correlation structures, and have demonstrated explicitly how firing variability shapes the network structures learned through STDP.

## DISCUSSION

We have investigated how the high-order statistics of spiking activity affect the learning dynamics in a plastic LIF neuronal network. We have developed a self-consistent linear response theory, which incorporates the information of both firing rate and firing variability. By decomposing the spike correlation

structure into the component associated with local direct connections and that associated with indirect connections, we have identified two distinct regimes regarding the network structures learned through STDP. In one regime, the contribution of the correlation associated with local direct connections dominates in STDP learning dynamics, leading to network structures consistent with the previous work (Morrison et al., 2007; Kozloski and Cecchi, 2010; Babadi and Abbott, 2013). In the other regime, the contribution of the correlation associated with indirect connections, instead of local direct connections, dominates, yielding network structures at variance with the prediction by the simple firing rate theory. The dominance of the indirect-connection correlation mainly arises from the asymmetric common-input correlations, which is induced by the heterogeneity of firing variability across different neuronal populations. Therefore, our study highlights the importance of firing variability of spiking activity for the evolution of plastic neural networks.

In a series of work (Gilson et al., 2009a,b,c,d, 2010), the STDP learning dynamics in recurrent networks has been investigated with Hawkes processes. However, the response kernel of the Hawkes process used there is identical across neuronal populations, thus the effect of firing variability was not addressed. Here, we study the STDP learning dynamics with the LIF neuron, in which the effect of firing variability is addressed by varying the mean and variance of external inputs across different neuronal populations. In Babadi and Abbott (2013), the authors developed a theory to explain the formation of in- and out-hubs. However, the effect of firing variability has not been discussed. Our theory can be viewed as an extension of the work (Babadi and Abbott, 2013). More recently, motif dynamics have been studied with linear response theory for one homogeneous population (Ocker et al., 2015). Instead, our work focuses on the evolution of connection strengths across different neuronal populations.

We have focused here on the balanced STDP rule. However, it is straightforward to extend our analysis to the unbalanced case by including additional terms in learning dynamics (Equations 35, 36). An unbalanced STDP rule can be expected to act as a loop-generating or loop-eliminating mechanism, depending on whether the rule is potentiation-dominated or depression-dominated (Babadi and Abbott, 2013). The neuron model we used here is the current-based LIF neuron. As we have demonstrated, the key property of a single neuron model to have an indirect-connection-dominated regime is that the firing variability of spiking activity can be encoded in the time constant of linear response kernel. Therefore, it can be expected that the indirect-connection-dominated regime can

exist with other neuron models, such as conductance-based models, exponential integrate-and-fire models or more realistic Hodgkin-Huxley neurons. The connectivity here is of the all-to-all type. We can readily extend our self-consistent linear response theory to the sparsely connected network case. For larger network size, such as the network of 500 excitatory neurons and 500 inhibitory neurons, we have also found that there is an indirect-connection-dominated regime, provided that there is a significant difference of firing variability between different neuronal populations. In this work, the input statistics (i.e., input means and input variances) are kept constant, indicating that the circuit we studied receives the same sensory stimuli during the whole simulation. Under this assumption, the neural population activity can be simply quantified by mean firing rate and CV of ISI. It would be very interesting to study the learning dynamics when the stimulation periods are interleaved by spontaneous states. Another interesting issue to be addressed in the future work is how the high-order firing statistics of spiking activity affect the connectivity structures in a circuit with more complex learning rules, such as triplet STDP learning rule (Gjorgjieva et al., 2011).

In conclusion, we have demonstrated that the introduction of heterogeneity of firing variability across different neuronal populations can give rise to counterintuitive network structures in a plastic neural circuit. Our work suggests that the high-order statistics of spiking activity can play an important role in shaping the connectivity patterns in our brain.

## AUTHOR CONTRIBUTIONS

Conceived and designed the research, performed experiments, analyzed data, and wrote the paper: BM, DZ, and DC.

## ACKNOWLEDGMENTS

This work was supported by NYU Abu Dhabi Institute G1301 (BM, DZ, and DC), NSFC-11671259, NSFC-91630208, NSFC-11722107, Shanghai Rising-Star Program-15QA1402600 (DZ), NSFC 31571071, NSF DMS-1009575 (DC), Shanghai 14JC1403800, Shanghai 15JC1400104, SJTU-UM Collaborative Research Program (DZ and DC).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2018.00001/full#supplementary-material>

## REFERENCES

- 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
- Babadi, B., and Abbott, L. F. (2016). Stability and competition in multi-spike models of spike-timing dependent plasticity. *PLoS Comput. Biol.* 12:e1004750. doi: 10.1371/journal.pcbi.1004750
- 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.
- Bi, Z., and Zhou, C. (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

- 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
- Clopath, C., Büsing, L., Vasilaki, E., and Gerstner, W. (2010). Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nat. Neurosci.* 13, 344–352. doi: 10.1038/nn.2479
- Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience*, vol. 806. Cambridge, MA: MIT Press.
- Fourcaud, N., and Brunel, N. (2002). Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural Comput.* 14, 2057–2110. doi: 10.1162/089976602320264015
- Gilson, M. (2010). STDP in recurrent neuronal networks. *Front. Comput. Neurosci.* 4:23. doi: 10.3389/fncom.2010.00023
- 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 III: partially connected neurons driven by spontaneous activity. *Biol. Cybern.* 101, 411–426. doi: 10.1007/s00422-009-0343-4
- Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2009d). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks IV. *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
- Gjorgjieva, J., Clopath, C., Audet, J., and Pfister, J.-P. (2011). A triplet spike-timing-dependent plasticity model generalizes the biestock-cooper-munro rule to higher-order spatiotemporal correlations. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19383–19388. doi: 10.1073/pnas.1105933108
- Grytskyy, D., Tetzlaff, T., Diesmann, M., and Helias, M. (2013). A unified view on weakly correlated recurrent networks. *Front. Comput. Neurosci.* 7:131. doi: 10.3389/fncom.2013.00131
- Hawkes, A. G. (1971a). Point spectra of some mutually exciting point processes. *J. R. Stat. Soc. Ser. B (Methodol.)* 33, 438–443.
- Hawkes, A. G. (1971b). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58, 83–90. doi: 10.1093/biomet/58.1.83
- Helias, M., Tetzlaff, T., and Diesmann, M. (2013). Echoes in correlated neural systems. *New J. Phys.* 15:023002. doi: 10.1088/1367-2630/15/2/023002
- Hu, Y., Trousdale, J., Josic, K., and Shea-Brown, E. (2013). Motif statistics and spike correlations in neuronal networks. *J. Stat. Mech. Theory Exp.* 2013:P03012. doi: 10.1088/1742-5468/2013/03/P03012
- Kozloski, J., and Cecchi, G. A. (2010). A theory of loop formation and elimination by spike timing-dependent plasticity. *Front. Neural Circuits* 4:7. doi: 10.3389/fncir.2010.00007
- Litwin-Kumar, A., and Doiron, B. (2014). Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nat. Commun.* 5:5319. doi: 10.1038/ncomms6319
- Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275, 213–215. doi: 10.1126/science.275.5297.213
- Morrison, A., Aertsen, A., and Diesmann, M. (2007). Spike-timing-dependent plasticity in balanced random networks. *Neural Comput.* 19, 1437–1467. doi: 10.1162/neco.2007.19.6.1437
- 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
- 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
- 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
- Pernice, V., Staudé, B., Cardanobile, S., and Rotter, S. (2012). Recurrent interactions in spiking networks with arbitrary topology. *Phys. Rev. E* 85:031916. doi: 10.1103/PhysRevE.85.031916
- Rao, R. P., and Sejnowski, T. J. (2001). Spike-timing-dependent hebbian plasticity as temporal difference learning. *Neural Comput.* 13, 2221–2237. doi: 10.1162/089976601750541787
- Ravid Tannenbaum, N., and Burak, Y. (2016). Shaping neural circuits by high order synaptic interactions. *PLoS Comput. Biol.* 12:e1005056. doi: 10.1371/journal.pcbi.1005056
- Schuecker, J., Diesmann, M., and Helias, M. (2015). Modulated escape from a metastable state driven by colored noise. *Phys. Rev. E* 92:052119. doi: 10.1103/PhysRevE.92.052119
- Song, S., Miller, K. D., and Abbott, L. F. (2000). Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nat. Neurosci.* 3, 919–926. doi: 10.1038/78829
- Trousdale, J., Hu, Y., Shea-Brown, E., and Josic, K. (2012). Impact of network structure and cellular response on spike time correlations. *PLoS Comput. Biol.* 8:e1002408. doi: 10.1371/journal.pcbi.1002408
- Zenke, F., Agnes, E. J., and Gerstner, W. (2015). Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nat. Commun.* 6:6922. doi: 10.1038/ncomms7922

**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 © 2018 Min, Zhou and Cai. 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.