Functional stability and recurrent STDP in rhythmogenesis

Gabi Socolovsky[1, 2, *] and Maoz Shamir[1, 2, 3]

[1] *Department of Physics, Faculty of Natural Sciences,*
[2] *The School of Brain Sciences and Cognition,*
[3] *Department of Physiology and Cell Biology, Faculty of Health Sciences,*
Ben-Gurion University of the Negev, Be'er-Sheva, Israel

(Dated: March 4, 2025)

Synapses in the nervous system show considerable volatility, raising the question of how the brain maintains functional stability despite continuous synaptic motility. Previous studies have suggested that functionality may be maintained by an ongoing process of activity-dependent plasticity. Here, we address this question in the context of rhythmogenesis. Specifically, we investigated the hypothesis that rhythmic activity in the brain can develop and be stabilized via activity-dependent plasticity in the form of spike-timing-dependent-plasticity (STDP), extending our previous work that demonstrated rhythmogenesis in a toy model of two effective neurons. We examined STDP dynamics in large recurrent networks in two stages. We first derived the effective dynamics of the order parameters of the synaptic connectivity. Then, a perturbative approach was applied to investigate stability. We show that for a wide range of parameters STDP can induce rhythmogenesis. Moreover, STDP can suppress synaptic fluctuations that disrupt functionality. Interestingly, STDP can channel fluctuations in the synaptic weights into a manifold on which the network activity is not affected, thus, maintaining functionality while allowing a subspace in which synaptic weights can be widely distributed.

I. INTRODUCTION

It is widely accepted that our sensory perception as well as planned motor commands are represented in the brain by the dynamical state of the neural activity. Synaptic connections shape the space of this activity and hence functionality of the brain. However, empirical findings report that the strength of synaptic connections; i.e., synaptic weights, is highly volatile [1–7]. Synaptic weights have been shown to fluctuate by about 50 percent over the course of several days [6, 7]. Moreover, activity-independent fluctuations in the synaptic weights with a similar order of magnitude have also been observed [4, 5, 7]. This raises the question as to how functionality can be maintained in the face of constant synaptic motility.

Here, we address the question of functional stability within the context of rhythmogenesis. Rhythmic activity has been widely observed in the central nervous systems of a diverse range of animals and has been linked to various cognitive tasks [8–16]. Numerous computational studies have investigated models that can generate this type of rhythmic activity [8, 17–25]. In these models, for rhythmic activity to emerge, the synaptic weights must be within a certain range of values. Thus, in the context of rhythmogenesis, the question of functionality is translated into the mechanism that ensures that the synaptic weights develop to the target range and stabilizes them.

Activity-dependent synaptic plasticity, and in particular Spike-Timing-Dependent Plasticity (STDP), has been suggested as a potential mechanism to maintain functionality [26–30]. In STDP, the synaptic weight is modified according to the time difference between the pre- and post-synaptic firings [31–35]. Various STDP learning rules have been observed across multiple brain regions and in numerous animal species [32, 33, 36–41].

We previously showed, as a proof of concept, that STDP can lead the development of the synaptic weights into the required range for rhythmic activity [42]. However, this was done in the framework of a minimal model, which is equivalent to a network of one excitatory and one inhibitory neuron with only reciprocal connections. Thus, it remains unclear whether STDP can provide a mechanism for rhythmogenesis in large networks with other types of connectivity and motility of individual synapses.

To address these questions, here we investigated rhythmogenesis in a model with large excitatory and inhibitory populations that incorporates various types of synaptic connections. STDP was applied to the synapses that are crucial for rhythmogenesis (fig. 1).

This paper is structured as follows. In section II we define the dynamical model of the neural responses. First, in II A, we derive the phase diagram of the network in the case where all synapses of the same type are equal, and identify the mean (across the population) synaptic weights as order parameters of the system. Then, in II B, we extend the analysis by incorporating small fluctuations in the synaptic weights around the order parameters, and compute the cross-correlations in the neural responses. In Section III we introduce the STDP dynamics. In III A we define the STDP learning rules used in this study and develop the dynamical equations for synaptic weights in the limit of slow learning. In our approach, STDP dynamics is studied in two stages. First, we study the induced STDP dynamics of the order parameters,

* gabisoco@post.bgu.ac.il

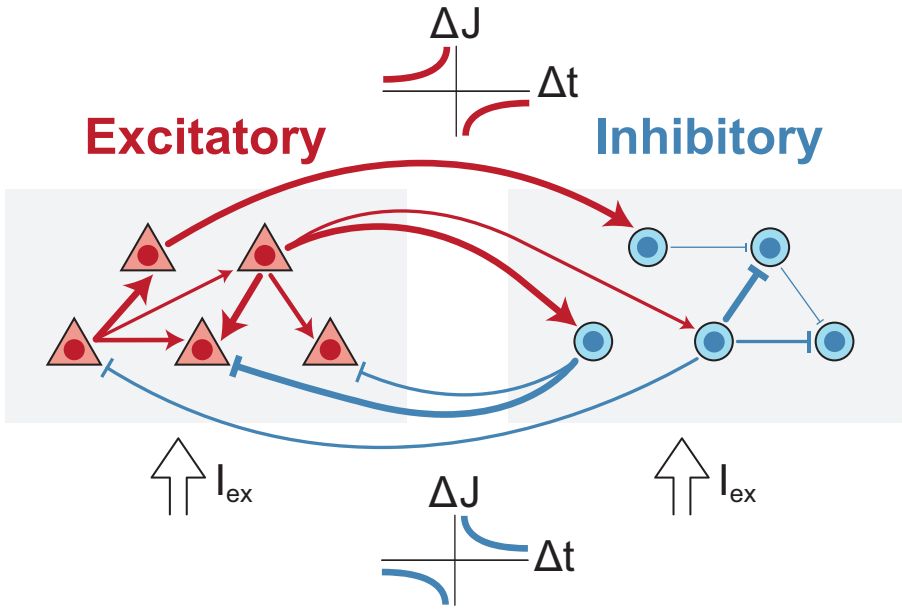


FIG. 1: Network architecture.

III B. Then, the STDP dynamics of fluctuations around the order parameters are analyzed using a perturbative approach, III C. Finally, we summarize our results and discuss our work in the context of maintaining functional stability in the face of constant synaptic motility, and theoretical approaches for studying recurrent STDP.

II. FIRING RATES DYNAMICS

We model a system of N_E excitatory (E) neurons and N_I inhibitory (I) neurons (fig. 1). The spiking activity of each neuron follows inhomogeneous Poisson statistics, where $\rho_{X,i}(t) = \sum_k \delta\left(t - t^k_{X,i}\right)$ is the spike train of neuron i in population $X \in \{E, I\}$, expressed as the sum of Dirac's delta functions at times $t^k_{X,i}$. Following Roxin et al. [22], the instantaneous firing rate of neuron i in population X, $m_{X,i}(t) = \langle \rho_{X,i}(t) \rangle$, where $\langle ... \rangle$ denotes ensemble averaging, is governed by:

$$\tau_m \dot{m}_{X,i}(t) = -m_{X,i}(t) + \Bigg[I_{ex} + \frac{1}{N_E} \sum_{n=1}^{N_E} J_{XE,in} m_{E,n}(t-d) - \frac{1}{N_I} \sum_{k=1}^{N_I} J_{XI,ik} m_{I,k}(t-d) \Bigg]_+ \tag{1}$$

where $J_{XY,ij}$ is the synaptic connection from neuron j in population $Y \in \{E, I\}$ to neuron i in population X. The characteristic time of the neuronal response is τ_m, and I_{ex} is an external input. The parameter d is the delay in the neural response to its input (see [25] for discussion of various physiological sources for the delay). The function $[x]_+$ denotes the threshold linear function; i.e., $[x]_+ = H(x)\,x$, where $H(x)$ is the Heaviside function. Unless stated otherwise, we measure time in units of the characteristic time of the neural response; i.e., $\tau_m = 1$, and the neural responses are measured relative to the external input, i.e., $I_{ex} = 1$.

Note that although this model contains autapses, $J_{XX,ii}$, their contribution to the total synaptic input to each neuron is negligible for large networks, $N_E, N_I \gg 1$.

In the following subsections we analyze the neural dynamics in the case of uniform synapses and extend these results to small fluctuations around the uniform synapses solution.

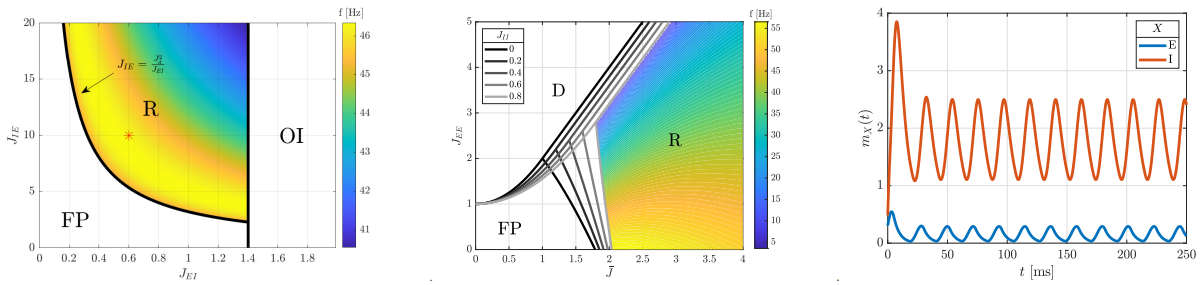
A. Homogeneous network and order parameters

In the case of uniform synaptic coefficients, $J_{XY,ij} = J_{XY}\ \forall i,j$, a uniform solution, in which $m_{X,i}(t) = m_X(t)\ \forall i$, exists and is stable. In this case eq. (1), is reduced to

$$\dot{m}_X(t) = -m_X(t) + \left[1 + J_{XE} m_E(t-d) - J_{XI} m_I(t-d)\right]_+, \quad X \in \{E, I\}. \tag{2}$$

The range of the synaptic weights we chose to investigate revealed four global states of activity (See detailed derivation in Appendix A 1). Specifically, since we were interested in rhythmic activity that results from E-I interaction, we assumed moderate intra-connection strengths; i.e., $J_{EE}, J_{II} < 1$.

Depending on the values of the synaptic weights, eq. (2) exhibits different types of behaviors, figs. 2a and 2b. For sufficiently strong I-to-E inhibition, $J_{EI} > 1 + J_{II}$, the activity of the excitatory population can be fully suppressed by inhibition and the system converges



(a) **(b)** **(c)**

FIG. 2: The phase diagram. (a) Phase diagram in the plane of $[J_{EI}, J_{IE}]$. Three regions are delineated: Fixed Point (FP), Only Inhibition (OI), and Rhythmic (R). The bifurcation line separating the FP and R regions is given by $J_{IE} = \bar{J}_d^2/J_{EI}$. The frequency of the rhythmic activity in the R region is shown by color. Here, we used $J_{EE} = 0.6$ and $J_{II} = 0.4$. (b) Phase diagram in the plane of $[\bar{J}, J_{EE}]$. Three regions are delineated: Fixed Point (FP), Rhythmic (R), and Divergence (D) for $J_{II} = 0, 0.2, ..., 0.8$. The frequency of the rhythmic activity in the R region is shown by color. (c) Example of the mean E and I firing rates in the rhythmic region, depicting $m_E(t)$ and $m_I(t)$ as a function of time, for $J_{EE} = J_{EI} = 0.6$, $J_{IE} = 10$ and $J_{II} = 0.4$ (See red asterisk in fig. 2a). For all figures, $\tau_m = 2.5d = 5$ ms was used.

to a fixed-point in which Only Inhibition (OI) is active, $\mathbf{m}^* \equiv \begin{pmatrix} m_E^* \\ m_I^* \end{pmatrix} = \frac{1}{1+J_{II}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

When the E-to-E connectivity exceeds a certain threshold, $J_{EE} > \begin{cases} 1 + \bar{J}^2/(1+J_{II}) & \bar{J} < \bar{J}_{tc} \\ 2\bar{J} - J_{II} & \bar{J} > \bar{J}_{tc} \end{cases}$ where $\bar{J} \equiv \sqrt{J_{EI} J_{IE}}$, and $\bar{J}_{tc} = 1 + J_{II}$, the positive feedback results in a Divergence (D) of the neuronal activities. This scenario represents pathological activity where diverging neuronal responses are suppressed due to a saturating non-linearity that is not described in our model.

For moderate to low I-to-E and E-to-E levels; i.e., outside of the OI and D regions, the system has a Fixed Point (FP) that represents an asynchronous state in which both populations are active,

$$\mathbf{m}^* = \frac{1}{(1 - J_{EE})(1 + J_{II}) + \bar{J}^2} \begin{pmatrix} 1 + J_{II} - J_{EI} \\ J_{IE} + 1 - J_{EE} \end{pmatrix}. \quad (3)$$

However, when $\bar{J}$ crosses a certain threshold, $\bar{J}_d$, the asynchronous fixed point loses its stability and a limit cycle emerges so that the network relaxes to a rhythmic (R) activity (See figs. 2a and 2c). As the E-to-E connectivity, J_{EE}, is strengthened, the oscillation frequency in the R region decreases monotonically and approaches zero on the transition to the D region, $J_{EE} = 2\bar{J} - J_{II}$ (See fig. 2b).

Solving the delayed differential equations for the neuronal dynamics in the R domain is not an easy task. However, on the bifurcation line the dynamics is linear, yielding:

$$\mathbf{m}(t) = \begin{pmatrix} \bar{m}_E + \tilde{m}_E \cos(\omega_d t + \psi) \\ \bar{m}_I + \tilde{m}_I \cos(\omega_d t) \end{pmatrix}, \quad (4)$$

where $\bar{m}_X$ is the temporal mean of the firing rate, $\tilde{m}_X$ is the amplitude modulation of the firing rate, ω_d is the angular frequency and ψ is the relative temporal phase between the excitatory and inhibitory populations (see the derivations for ω_d in Appendix A 1 and for $\bar{m}_X$, $\tilde{m}_X$ and ψ in Appendix A 2).

B. Beyond order parameters

In the following section we investigate the STDP dynamics, in which the synaptic weights change over time. The driving force of the STDP dynamics is the product of the STDP rule and the neuronal cross-correlations. This makes it imperative important to understand how the cross-correlations depend on the synaptic weights. To this end, we describe the weights by their means, which serve as the order parameters, and the fluctuations around the means

$$J_{XY} \equiv \frac{1}{N_X N_Y} \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} J_{XY,ij}, \quad (5)$$

$$\delta J_{XY,ij} = J_{XY,ij} - J_{XY}. \quad (6)$$

In our model, the cross correlations between different neurons is given by the product of their mean firing rates. We define the temporal averaged cross correlations $\Gamma_{XY,ij}(\Delta) = \langle m_{X,i}(t)\, m_{Y,j}(t + \Delta) \rangle_T$, where $\langle \cdot \rangle$ denotes averaging over one period of the limit cycle (in the R region). Expanding to first order in $\delta\mathbf{J}$, we write

$$\Gamma_{IE,ji}(\Delta) = \Gamma_{IE}(\Delta) + \delta\Gamma_{IE,ji}(\Delta) + \mathcal{O}(\|\delta\mathbf{J}\|^2) \quad (7)$$

In the FP region, in which the firing rates converge to a fixed point, to a leading order in the fluctuations, the cross correlations are given by (see Appendix A 3 and A 4)

$$\Gamma_{IE}(\Delta) = m_E^* m_I^*, \quad (8)$$

$$\delta\Gamma_{IE,ji}(\Delta) = m_E^{*2} \delta s_{I,j} - m_I^{*2} \delta s_{E,i}, \quad (9)$$

where $\delta s_{X,i} = \frac{1}{N_Y}\sum_{j=1}^{N_Y}\delta J_{XY,ij}$ $(X \neq Y)$ is the mean fluctuation in the inter-synapses to neuron i in population X.

In general, in the rhythmic region the dynamics is non-linear and the computation of the limit cycle is not trivial. However, near the bifurcation line we can approximate the neural responses by their dynamics on the bifurcation line, which is linear, yielding (see Appendix A 3 and A 4)

$$\Gamma_{IE}(\Delta) = \bar{m}_E\bar{m}_I + \frac{\tilde{m}_E\tilde{m}_I}{2}\cos(\omega_d\Delta + \psi) \tag{10}$$

$$\begin{aligned}\delta\Gamma_{IE,ji}(\Delta) = &\left(\bar{m}_E^2 + \frac{\tilde{m}_E^2}{2g}\cos(\omega_d\Delta + \varphi)\right)\delta s_{I,j} \\ &- \left(\bar{m}_I^2 + \frac{\tilde{m}_I^2}{2g}\cos(\omega_d\Delta - \varphi)\right)\delta s_{E,i},\end{aligned} \tag{11}$$

where $\bar{m}_X$, $\tilde{m}_X$, g, ψ and φ are given in Appendix A 1 and A 2. Note that $\Gamma_{EI,ij}(\Delta) = \Gamma_{IE,ji}(-\Delta)$. For convenience, we further define the following quantities dependent on the order parameters

$$\begin{aligned}\bar{\Gamma} &\equiv \frac{1}{T_d}\int_0^{T_d}\Gamma_{XY}(\Delta)\,d\Delta \\ \tilde{\Gamma} &\equiv \frac{2}{T_d}\int_0^{T_d}\Gamma_{XY}(\Delta)\cos(\omega_d\Delta_X + \psi)d\Delta\end{aligned} \tag{12}$$

where $T_d = 2\pi/\omega_d$, $\Delta_E = -\Delta$, and $\Delta_I = \Delta$. In the FP region $\bar{\Gamma} = m_E^* m_I^*$ and $\tilde{\Gamma} = 0$. In the R region near the bifurcation line $\bar{\Gamma} = \bar{m}_E\bar{m}_I$ and $\tilde{\Gamma} = \tilde{m}_E\tilde{m}_I/2$.

III. STDP DYNAMICS

In section III we define the STDP learning rule and introduce the STDP dynamics in the limit of slow learning. Next, in section III B we study the induced STDP dynamics of the order parameters. Then, in section III C we study the dynamics of the fluctuations around the order parameters.

A. The STDP rule

We now define the synaptic plasticity rule. As the inter-connections, J_{EI} and J_{IE}, are the essential synaptic connections needed to generate rhythmic activity, we focus on their plasticity, and assume that the intra-connections, J_{EE} and J_{II}, are not plastic (illustrated in fig. 1). In our STDP learning rule, the modification of the weight, $dJ_{XY,ij}$, of a synapse from pre-synaptic neuron j in population Y to post-synaptic neuron i in population X, resulting from a pre-synaptic spike at time t_j and a post-synaptic spike at time $t_i = t_j + \Delta$ is given by (see also [35, 42, 43])

$$dJ_{XY,ij} = \lambda_Y\left[f_Y(J_{XY,ij})K_{Y,+}(\Delta) - \alpha K_{Y,-}(\Delta)\right]. \tag{13}$$

The change in the synaptic weight $dJ_{XY,ij}$ is given by the contribution of two processes: potentiation (+) and depression (−). Each process may depend on both the spike time difference, Δ, and on the synaptic weight, $J_{XY,ij}$. Here for mathematical convenience we assumed a separation of variables, and wrote each process as the product of the synaptic weight dependent function, $f_Y(J_{XY,ij})$, and the temporal kernel $K_{Y,\pm}(\Delta)$. The parameters λ_Y and α correspond to the learning rate and the relative strength of depression, respectively. Note that since only two types of synapses exhibit plasticity, it suffices to denote λ_Y, f_Y and $K_{Y,\pm}$ solely by the type of their pre-synaptic neuron, Y.

The weight dependent function, f_Y, restrains the potentiation of Y-to-X synapses from diverging. Following Gütig et al. [35] (see also [42]) for the excitatory synapses we chose

$$f_E(z) = (1 - z/J_{max})^\mu \tag{14}$$

where μ characterizes the non-linearity of the weight dependence and J_{max} is the maximal value of the excitatory synaptic weights. For inhibitory synapses we took $f_I = 1$.

For the temporal kernels of the STDP rule we focused on temporally asymmetric Hebbian [33, 37, 40, 44, 45] and anti-Hebbian rules [36, 43, 46]. Specifically, we used exponential kernels:

$$K_{Y,\pm}(\Delta) = \frac{1}{\tau_{Y,\pm}}e^{\mp H\Delta/\tau_{Y,\pm}}\Theta(\pm H\Delta) \tag{15}$$

The change in the synaptic weight, eq. (13), during short time intervals, yields (see also [35, 43, 47])

$$\begin{aligned}&\dot{J}_{XY,ij}(t) = \\ &\lambda_Y\rho_{X,i}(t)\int_{-\infty}^{t}\rho_{Y,j}(t')K_{XY,ij}(t-t')\,dt' \\ &+ \lambda_Y\rho_{Y,j}(t)\int_{-\infty}^{t}\rho_{X,i}(t')K_{XY,ij}(t'-t)\,dt'\end{aligned} \tag{16}$$

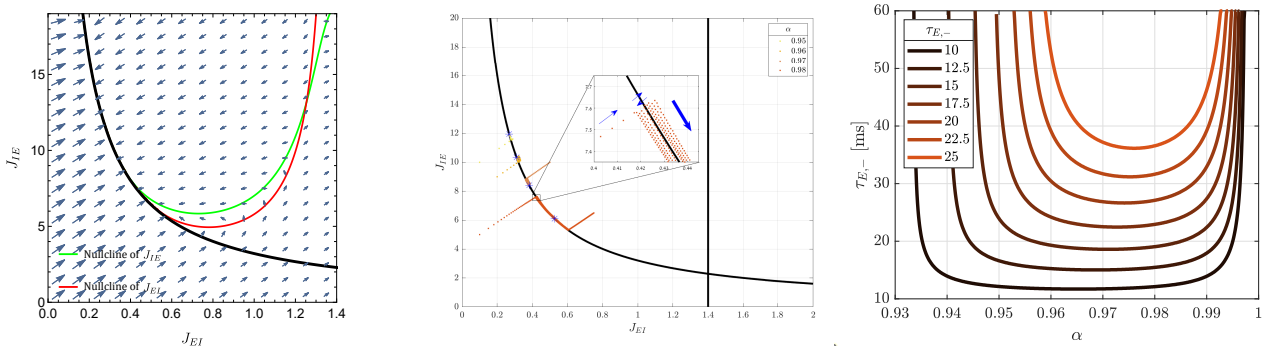
where $K_{XY,ij}(z) = f_Y(J_{XY,ij})K_{Y,+}(z) - \alpha K_{Y,-}(z)$. In the limit of slow learning, $\lambda_Y \ll \omega_d$, the STDP dynamics, eq. (16), averages the neural responses over time, yielding (see [43] for a more detailed derivation)

$$\dot{J}_{XY,ij} = \lambda_Y\int_{-\infty}^{\infty}K_{XY,ij}(\Delta)\Gamma_{XY,ij}(-\Delta)\,d\Delta. \tag{17}$$

B. Induced dynamics of the order parameters

Assuming a uniform solution in which $J_{XY,ij} = J_{XY}$, reduces the complex high-dimensional dynamics in eq. (17) to the following synaptic dynamics

$$\dot{J}_{XY} = \lambda_Y\int_{-\infty}^{\infty}K_Y(\Delta)\Gamma_{XY}(-\Delta)\,d\Delta \tag{18}$$



(a) **(b)** **(c)**

FIG. 3: Critical Rhythmogenesis. (a) The phase diagram of neural activity (fig. 2a) is the phase space of the STDP dynamics. Flow in the phase space $[J_{EI}, J_{IE}]$. The green and red lines denote the nullclines of the excitatory (J_{IE}) and inhibitory (J_{EI}) synaptic weights order parameters, respectively. We chose $\alpha = 0.96$, $\tau_{E,-} = 35$ ms and $\lambda_E = 10\lambda_I = 1$. (b) STDP dynamics of the order parameters, governed by eq. (18), were simulated for different values of $\alpha = 0.95, 0.96, 0.97, 0.98$. Critical points $\mathbf{J} = \mathbf{J}^*$ for each α are marked with blue asterisks. For $\alpha = 0.98$, the dynamics were computed for two different initial conditions. (Inset) The thin blue arrows highlight the arrival to the bifurcation line and the alternate transitions across it. The thick blue arrow highlights the drift towards $\mathbf{J}^*$. Here we used $\tau_{E,-} = 25$ ms and $\lambda_E = 10\lambda_I = 4$. (c) The region in parameter space–specifically here the plane of $[\alpha, \tau_{E,-}]$–in which critical rhythmogenesis can occur, is shown for various values of $\tau_{I,-} = 10, 12.5, ..., 25$ ms. Critical rhythmogeneis is achievable above the corresponding lines. In all the figures we used $\mu = 0.01$, $J_{IE,max} = 20$, $\tau_{E,+} = \tau_{I,+} = 10$ ms, $\tau_{I,-} = 15$ ms, $J_{EE} = 0.6$, $J_{II} = 0.4$ and $\tau_m = 2.5d = 5$ ms.

where $K_Y(\Delta) = f_Y(J_{XY}) K_{Y,+}(\Delta) - \alpha K_{Y,-}(\Delta)$. The separation of time scales between neural and STDP dynamics enables examination of the neural activity while maintaining constant synaptic weights. In this framework, the synaptic weights J_{XY} are treated as parameters of the firing rates $m_X(t)$ phase diagram (See figs. 2a and 2b). According to eq. (18), on a large time scale, λ^{-1}, STDP induces a dynamic flow on the phase diagram, simultaneously defined as the phase space of synapses $[J_{EI}, J_{IE}]$. The STDP parameters shape this flow and indirectly dictate the collective behavior of the neural network $m_X(t)$ in accordance with the current position of the synaptic weights on the phase diagram.
Substituting Γ_{XY} from eqs. (8) and (10) into eq. (18) yields

$$\dot{J}_{XY} = \lambda_Y \left(\bar{\Gamma} \bar{K}_Y + \tilde{\Gamma} \tilde{K}_Y(\psi) \right) \tag{19}$$

where

$$\bar{K}_Y \equiv \int_{-\infty}^{\infty} K_Y(\Delta)\, d\Delta$$
$$\tilde{K}_Y(z) \equiv \int_{-\infty}^{\infty} K_Y(\Delta) \cos(\omega_d \Delta_Y + z) d\Delta. \tag{20}$$

The flow on the phase diagram, eq. (19), is controlled by the first two temporal Fourier transform components of the mean cross-correlations between populations and the learning rule. In the FP region only the first term exists, $\tilde{\Gamma} = 0$; therefore, due to our choice of normalization $\bar{K}_\pm = 1$, we have $\dot{J}_{XY} \propto f_Y - \alpha$. Consequently, for a weak relative strength of depression, characterized by $\alpha < f_E < 1$, STDP induces a flow that attracts the system from the majority of the FP region toward the R region, as illustrated in fig. 3a. When the order parameters traverse the bifurcation line into the R region, $\tilde{\Gamma}$ undergoes a discontinuous change from 0 to $\tilde{m}_E \tilde{m}_I / 2$. Therefore, as the second term in eq. (19) begins to influence the flow on the phase diagram, it may lead to a change in the sign of $\dot{J}_{XY}$. The effect of the second term in eq. (19) on the dynamics will be also determined by the second Fourier transform component of the temporal kernel, $\tilde{K}_Y(\psi)$. In the case of an asymmetric Hebbian rule for J_{EI} and an asymmetric anti-Hebbian rule for J_{IE} we get

$$\tilde{K}_Y(\psi) = f_Y(J_{XY}) \tilde{K}_{Y,+}(\psi) - \alpha \tilde{K}_{Y,-}(\psi) \tag{21}$$

where $\tilde{K}_{Y,\pm}(z) = \cos(\theta_{Y,\pm}) \cos(\theta_{Y,\pm} \pm z)$, $\cos(\theta_{Y,\pm}) = \left[1 + (\omega_d \tau_{Y,\pm})^2\right]^{-1/2}$. Due to $\theta_{Y,\pm}$ and ψ, the term $\tilde{\Gamma} \tilde{K}_Y(\psi)$ is negative, resulting in a sign change of $\dot{J}_{XY}$ across a significant part of the bifurcation line, as depicted in fig. 3a. Usually a change of sign in the time derivative of all the dynamic variables may indicate an attractor, where the dynamics slowly converge towards it. However, in this case, due to the discontinuity in $\tilde{\Gamma}$, $\dot{J}_{XY}$ change their sign discontinuously, resulting in repetitive transitions between the FP and R regions (inset in fig. 3b). These alternating transitions generate a continual toggling between fixed firing rates and rhythmic

activity. The transitions between the FP and R regions exhibit a drift along the bifurcation line, as shown in the inset of fig. 3b. This drift arises because the direction of $\dot{\mathbf{J}} \equiv \begin{pmatrix} \dot{J}_{EI} \\ \dot{J}_{IE} \end{pmatrix}$ on one side of the bifurcation line does not necessarily oppose the direction of $\dot{\mathbf{J}}$ on the subsequent step in the other region. To identify the critical point on the bifurcation line $\mathbf{J}^* = \begin{pmatrix} \bar{J}_d^2 / J_{IE}^* \\ J_{IE}^* \end{pmatrix}$, at which the drift vanishes and the dynamics stabilize in a state of repetitive transitions across $\mathbf{J}^*$, the discussed vectors must be oriented in opposing directions. This condition can be mathematically expressed as $\dot{\mathbf{J}}^{(\mathrm{FP})} \times \dot{\mathbf{J}}^{(\mathrm{R})}\Big|_{\mathbf{J}=\mathbf{J}^*} = \mathbf{0}$, which results in

$$\frac{\tilde{K}_I(\psi)}{\bar{K}_I} = \frac{\tilde{K}_E(\psi)}{\bar{K}_E} \tag{22}$$

and reveals the critical point

$$J_{IE}^* = J_{IE,max}\Bigg(1- \left(\frac{\left(1-\bar{K}_I\right)\left(\bar{K}_I\tilde{K}_{E,-}(\psi) - \tilde{K}_I(\psi)\right)}{\bar{K}_I\tilde{K}_{E,+}(\psi) - \tilde{K}_I(\psi)}\right)^{1/\mu}\Bigg). \tag{23}$$

which is determined by the STDP parameters. In fig. 3b, the dynamics of the order parameters are illustrated for varying values of the relative strength of depression α. For each α the system achieves critical rhythmogenesis at different values of $\mathbf{J}^*$ (indicated by the blue asterisks).

Two essential conditions must be met for critical rhythmogensis to occur. First, the left branches of the J_{XY} nullclines must exhibit stability. Second, the excitatory nullcline must diverge from the bifurcation line at a point higher than the divergence of the inhibitory nullcline. As demonstrated in fig. 3a, when these conditions are met, the flow between the nullclines facilitates a return to the bifurcation line, thereby ensuring its stability and enabling critical rhythmogenesis (see Socolovsky and Shamir [42] for details in the case of only E-to-I and I-to-E synapses). Note that at the points where the J_{XY} nullclines diverge from the bifurcation line, $\dot{J}_{XY} = 0$. Furthermore, to find the critical values of the STDP parameters that result in critical rhythmogenesis, the nullclines must depart from the bifurcation line at the same point. This means that $\dot{J}_{EI}\Big|_{\mathbf{J}=\mathbf{J}^*} = \dot{J}_{IE}\Big|_{\mathbf{J}=\mathbf{J}^*} = 0$. This imposes an additional condition on eq. (22)

$$\bar{K}_I\bar{\Gamma} + \tilde{K}_I(\psi)\tilde{\Gamma} = 0. \tag{24}$$

Figure 3c illustrates the phase space for critical rhythmogensis, by delineating its domain within the $[\alpha, \tau_{E,-}]$ phase space. Each line represents the boundary for different $\tau_{I,-}$ values, above which critical rhythmogenesis emerges.

C. STDP dynamics of the fluctuations around the order parameters

We next examined the STDP dynamics with synaptic fluctuations around the uniform solution, while the order parameters are relaxed to critical rhythmogensis. Assuming small fluctuations, we expand eq. (17) to the leading order in the fluctuations, $\delta J_{XY,ij}$, around $\mathbf{J}^*$, to obtain

$$\delta\dot{J}_{XY,ij} = \sum_{\substack{U=E,I \\ V\neq U}} \sum_{k=1}^{N_U}\sum_{l=1}^{N_V} M_{XYUV,ijkl}\,\delta J_{UV,kl} \tag{25}$$

where $M_{XYUV,ijkl}$ is a tensor relating a fluctuation in synaptic weight $\delta J_{UV,kl}$ to the dynamics of a fluctuation in the synaptic weight $\delta J_{XY,ij}$. The tensor depends on the order parameters $\mathbf{J}^*$, and on the collective state of the system; i.e., FP or R states. The details for $M_{XYUV,ijkl}$ appear in Appendix A 5.

The eigenvectors of the tensor $M_{XYUV,ijkl}$ can be classified into four families which represent four qualitatively different structures where synaptic fluctuations may evolve in time. These types of synaptic fluctuations are orthogonal to the uniform configuration of the synapses. Moreover, these families emerge regardless of the network size (although not for $N_E = 1$ and/or $N_I = 1$), so that we simply visualized them for $N_E = 3$ and $N_I = 2$ in fig. 4.

Eigenvectors in family I (fig. 3a) are fluctuations in inhibitory synapses to an excitatory cell, k ($k \in \{1, \dots N_E\}$), such that $\sum_l \Delta J_{EI,kl} = 0$ and thus conserve the total inhibitory input to neuron k and consequently do not affect the response dynamics. The constancy in neural dynamics does not alter the STDP dynamics, resulting in marginal stability for any order in these synaptic fluctuations. The multiplicity of these eigenvectors is $N_E(N_I - 1)$.

Eigenvectors in family II (fig. 3b) are fluctuations in the excitatory synapses to an inhibitory cell, k ($k \in \{1, \dots N_I\}$), such that $\sum_l \Delta J_{IE,kl} = 0$ and thus conserve the total excitatory input to neuron k and consequently do not affect the response dynamics. Although constancy in neural dynamics does not alter the STDP dynamics, built-in depression in excitatory dynamics, due to f_E, suppresses these fluctuations, resulting in a negative (stable) eigenvalue, (see Appendix A 5). The multiplicity of these eigenvectors is $N_I(N_E - 1)$.

In family III (fig. 3c) firing rates of individual excitatory neurons change while keeping the mean excitatory activity, $m_E(t) = \sum_i m_{E,i}(t)$, as well as activity of the inhibitory neurons fixed. In the example in fig. 3c, the white arrows indicate the neurons that change their activity. In the eigenvectors in family III, all the inhibitory connections to excitatory neuron k are modified by the same amount, $\Delta J_{EI,kj_1} = \Delta J_{EI,kj_2}\ \forall j_1, j_2$, and all of the connections of excitatory neuron k to the inhibitory

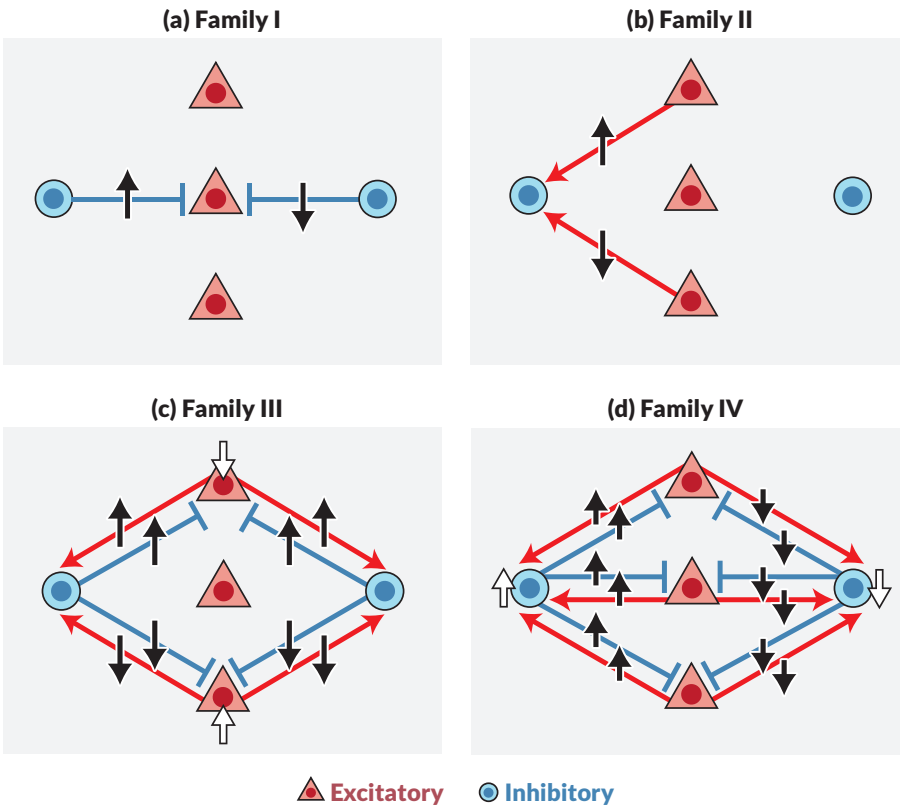


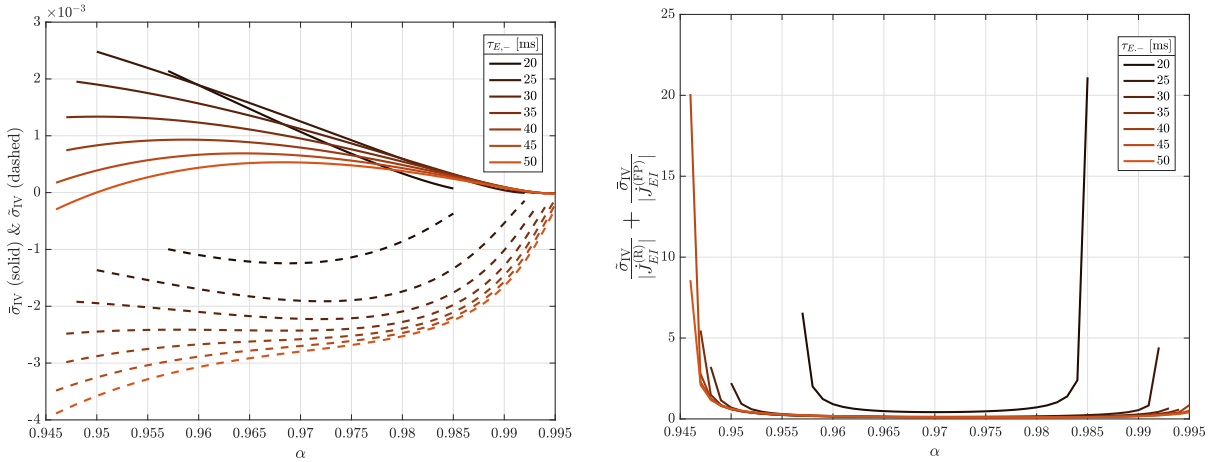
FIG. 4: Illustrations of the four families of synaptic fluctuations, for $N_E = 3$ and $N_I = 2$. (a) Family I - fluctuations in the I-to-E synapses, such that the inhibitory input to each excitatory neuron remains fixed. (b) Family II - fluctuations in the E-to-I synapses, such that the excitatory input to each inhibitory neuron remains fixed. (c) Family III - fluctuations in which excitatory neurons change their activity, while keeping the mean excitatory activity fixed. (d) Family IV - fluctuations in which inhibitory neurons change their activity, while keeping the mean inhibitory activity fixed. Upright (inverted) black arrows indicate strengthening (weakening) of the synapse. The white arrows illustrate the change in neural activity.

neurons obey, $\Delta J_{IE,i_1k} = \Delta J_{IE,i_2k} = \eta_{\rm III}\Delta J_{EI,ki_1}\ \forall i_1, i_2$ (see Appendix A 5). Figure 3c depicts an example eigenvector of this family. In this example, the fluctuations in the excitatory firing rates are in the form of $\delta m_E = (-1, 0, 1)$ (from top to bottom). This is achieved by increasing all the inhibitory connections to the top excitatory neuron and decreasing all the inhibitory connections to the bottom neuron. Fluctuations in this family decay (negative eigenvalue) when the network is around the bifurcation line. The degeneracy of this eigenvector results from the constraint $\sum_i m_{E,i} = const$; hence, this family spans an $N_E - 1$ dimensional space.

In family IV (fig. 3d), firing rates of individual inhibitory neurons change while keeping the mean inhibition, $m_I(t) = \sum_j m_{I,j}(t)$, fixed. In these eigenvectors, all the excitatory connections to inhibitory neuron k are modified by the same amount, $\Delta J_{IE,kj_1} = \Delta J_{IE,kj_2}$ $\forall j_1, j_2$, and all of the connections of inhibitory neuron k to excitatory neurons obey, $\Delta J_{EI,i_1k} = \Delta J_{EI,i_2k} = \eta_{\rm IV}^{-1}\Delta J_{IE,ki_1}\ \forall i_1, i_2$ (see Appendix A 5). Family IV eigenvectors span an $N_I - 1$ dimensional space from the condition $\sum_j m_{I,j} = const$.

The stability analysis of fluctuations in this family is more intricate. We found that the eigenvalue that corresponds to this family is typically positive (unstable) in the FP region and negative (stable) in the R region, fig. 5a. Thus, as the system switches alternately from FP to R states around the critical point, $\mathbf{J}^*$, these fluctuations grow (in the FP) and then decay (R) repeatedly. What can be said about the net effect on these fluctuations?

Heuristically, consider the following scenario. During the time, $T^{(\rm FP)}$, the system spends in the FP state, the fluctuations grow by a factor $e^{\tilde{\sigma}_{\rm IV}T^{(\rm FP)}}$. During the time, $T^{(\rm R)}$, the system spends in the R state, the fluctuations decay by a factor $e^{\tilde{\sigma}_{\rm IV}T^{(\rm R)}}$. Thus, stability is determined by the sum of the eigenvalues multiplied by the duration spent in each state, $\tilde{\sigma}_{\rm IV}T^{(\rm R)} + \bar{\sigma}_{\rm IV}T^{(\rm FP)}$ (see Appendix A 6). The duration spent in each region can be estimated by the inverse of the absolute value of the time derivative of the order parameters, $|\dot{\mathbf{J}}^{(\rm Z)}|^{-1}$, $Z = \{\rm FP,\ R\}$. At the critical point, $\hat{\dot{\mathbf{J}}}^{(\rm FP)} = -\hat{\dot{\mathbf{J}}}^{(\rm R)}$; therefore, the magnitude of the ratio of $\dot{\mathbf{J}}$ is equal to the ratio of the magnitude of each of its components, $T^{(\rm R)}/T^{(\rm FP)} = |\dot{J}_{EI}^{(\rm FP)}/\dot{J}_{EI}^{(\rm R)}| = |\dot{J}_{IE}^{(\rm FP)}/\dot{J}_{IE}^{(\rm R)}|$. Figure 5b, depicts the sum $\tilde{\sigma}_{\rm IV}/|\dot{J}_{EI}^{(\rm R)}| + \bar{\sigma}_{\rm IV}/|\dot{J}_{EI}^{(\rm FP)}|$ for different values of α and $\tau_{E,-}$. As shown,



(a) **(b)**

FIG. 5: The eigenvalues of family IV. (a) The eigenvalues of family IV in the FP regime (solid) and R regime (dashed) near the bifurcation line. (b) The value of $\tilde{\sigma}_{\rm IV}/|\dot{J}_{EI}^{(\rm R)}| + \bar{\sigma}_{\rm IV}/|\dot{J}_{EI}^{(\rm FP)}|$, which governs the stability of family IV fluctuations in critical rhythmogenesis. Each figure is drawn as a function of α for different values of $\tau_{E,-} = 20, 25, ..., 50$ ms differentiated by color. We used $\mu = 0.01$, $J_{IE,max} = 20$, $\lambda_E = \lambda_I = 1$ $\tau_{E,+} = \tau_{I,+} = 10$ ms, $\tau_{I,-} = 15$ ms, $J_{EE} = 0.6$, $J_{II} = 0.4$ and $\tau_m = 2.5d = 5$ ms.

the sum is negative and fluctuations in this family of eigenvectors are expected to decay over time.

To study the stability of critical rhythmogenesis with respect to fluctuations, we simulated STDP dynamics with additive Gaussian white noise (see details on white noise statistics in Appendix A 7). The magnitude of the fluctuations in each family as a function of time is depicted in fig. 6a. As can be seen in the figure, fluctuations in families II, III and IV converge to a steady state. By contrast, fluctuations in family I, with marginal stability accumulate over time. Nevertheless these fluctuations do not affect the neuronal response dynamics. This can be seen in figs. 6b and 6c, which show that the variability of the neural responses across the population converges to a finite size.

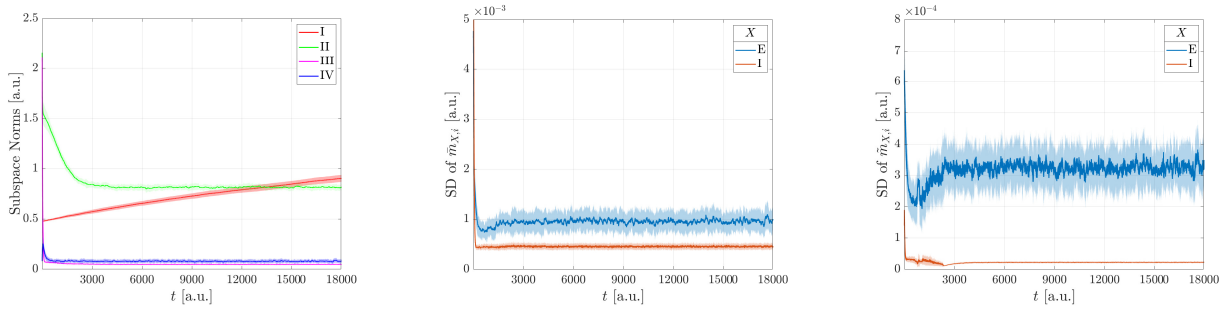
IV. DISCUSSION

Here we extended previous research on critical rhythmogenesis [42] (see also [46]). Our approach to investigating recurrent STDP dynamics has two stages. First, the dimensionality of the problem is reduced by studying the induced STDP dynamics of the order parameters. Then, a perturbative analysis is carried out in the subspace that is orthogonal to the order parameters.

The interconnections; i.e., E-to-I and I-to-E, are essential for generating the rhythmic activity in the specific model studied here [22, 42]. The intra-connections (E-to-E and I-to-I) mainly modulate this activity. However, beyond certain critical values, these intra-connections can drive the network toward either a pathological state (fig. 2b) or non-physiological rapid oscillations (see Appendix A 1). We found that homeostatic plasticity of the intra-connections can prevent the drift into these pathological states and maintain the intra-connections in the desired working range (results not shown). However, here, since we were interested in rhythmogenesis, we focused on the plasticity of the inter-connections and assumed that the intra-connections were non-plastic.

The order parameters span the phase diagram of the neural response (figs. 2a and 2b). Due to the separation of timescales between the neural response and the learning process, the STDP dynamics of the order parameters can be viewed as a flow on the phase diagram. We found that for a wide range of parameters, STDP induced a flow toward the bifurcation line (figs. 3a and 3b). Note that in this case, the system does not converge to a fixed point, and the temporal derivative of the order parameters, $\dot{J}_{XY}$, does not vanish on the bifurcation line. Rather, $\dot{J}_{XY}$ can be positive to the left of the bifurcation line and negative to the right of it, for example. Thus, the attraction to the bifurcation line results from two key features. The first is that, STDP is *activity*-dependent plasticity rule. The second is that, activity changes sharply on the bifurcation line. Hence, this mechanism can be generalized to other scenarios.

Reducing the dimensionality by focusing on the effective dynamics of the order parameters allowed us to study the effects of the non-linearity of the neuronal responses; i.e., the transition to rhythmic activity. To study the stability of the order parameters to arbitrary fluctuations, a stability analysis was carried out in orthogonal direction. Our analysis revealed four families of eigenmodes of fluc-



(a) **(b)** **(c)**

FIG. 6: Noisy STDP dynamics. STDP dynamics were simulated with added Gaussian white noise (see details in Appendix A 7). Results were averaged over 100 realizations of the noise. The initial conditions for the synaptic weights were drawn independently from Gaussian distribution $\sim \mathcal{N}(0.5, 6.25 \times 10^{-4})$. After a short transient, the system converges to critical rhythmogenesis. (a) In each time step the fluctuations were written as the sum of their four constituent vectors, one of each family. Their Euclidean norms, $\|\mathbf{x}\| := \sqrt{\mathbf{x} \cdot \mathbf{x}}$, are shown for each time step. (b) The standard deviation in $\tilde{m}_{E,i}$ (blue) and $\tilde{m}_{I,i}$ (orange) across their respective populations as a function of time. (c) The standard deviation in $\tilde{m}_{E,i}$ (blue) and $\tilde{m}_{I,i}$ (orange) across their population as a function of time. The neural and STDP dynamics (eqs. (1) and (17)) were simulated for a network consisting of $N_E = 40$ excitatory and $N_I = 10$ inhibitory neurons. Shaded areas represent the standard deviation across 100 trials. Here we used $\alpha = 0.98$, $\mu = 0.01$, $J_{IE,max} = 20$, $\lambda_E = 10\lambda_I = 10$, $\tau_{E,+} = \tau_{I,+} = 10$ ms, $\tau_{E,-} = 25$ ms, $\tau_{I,-} = 15$ ms, $J_{EE} = 0.6$, $J_{II} = 0.4$ and $\tau_m = 2.5d = 5$ ms.

tuations. Two families (II and III) are characterized by negative eigenvalues, indicating decay. One family (IV) is characterized with positive eigenvalue in the FP region and a negative eigenvalue in the R region but the net effect was of a decay in this subspace as well.

Finally, one family (I) is characterized by zero eigenvalues. Because it does not affect neural activity, the STDP dynamics of its fluctuations neither grow or decay, $\delta\dot{J}_{XY,ij} = 0$. Consequently, this family defines a manifold spanned by its eigenvectors, permitting the system to move freely within it (see also Burkitt et al. [48]). Thus, in our model, STDP maintains functionality by stabilizing the order parameters that govern the function while allowing considerable fluctuations in an orthogonal subspace.

The question of maintaining functional stability in the face of continuous synaptic motility has been addressed in the past, mainly in the framework of stabilizing a certain spatial structure of the neural responses. Three main approaches have been suggested. One approach assumes that only a subgroup of synapses is essential for functionality, and that these synapses are not (or less) volatile, either due to their type or strength [2, 49–52]. Another approach builds on the observation that neurons can have multiple synaptic connections. In this case, the noise in the effective synaptic weight can be decreased [53–55].

An alternative idea posits that functionality is stabilized through activity-dependent plasticity. Thus, fluctuations in the synaptic connectivity that disrupt functionality will cause a change in the neural activity, which, in turn, will induce negative feedback by the activity dependent plasticity. Thus, in this approach, activity dependent plasticity functions as a counterbalancing force that mitigates intrinsic synaptic fluctuations [28–30]. In Manz et al. [27], for example, the target functionality was the formation of a spatial structure in the form of neuronal clusters. Two types of fluctuations were found. The first were fluctuations that could disrupt functionality. These fluctuations were suppressed by STDP. The second were fluctuations that did not disrupt functionality, such as transferring a neuron from one cluster to another. These fluctuations were not suppressed. Thus, STDP allowed the system to drift in the space of possible solutions for network functionality.

Interestingly, Susman and colleagues [56] investigated homeostatic mechanisms that can mitigate the accumulation of random fluctuations in synaptic weights. They found that these mechanisms also degrade the stored memory in the system. However, they reported that memories that were encoded as a dynamical patterns were more robust than memories that are encoded by fix points.

In our work, the target function is the formation of a temporal structure. Here too we found two types of fluctuations. Fluctuations of the order parameters, as well as families III and IV may disrupt functionality and are suppressed by STDP. On the other hand, fluctuations in family I do not affect functionality. Hence, STDP allows the system to drift on the manifold of solutions, spanned by family I.

The analysis of STDP in recurrent neuronal networks is challenging due to the associated mathematical difficulties of solving high dimensional coupled non-linear dynamics. The coupling enters the dynamical equations via the cross-correlations of the neuronal responses, which together with the STDP rule provide the driving force for

the dynamics. Typically, in order to compute the cross-correlations, theoretical studies assume that the neural response is linear [27, 48, 57–63]. This assumption does not make the STDP dynamics low dimensional, uncoupled or linear. It simply makes it possible to derive the dynamical equations. To proceed with the analysis, one of two approaches is usually applied.

One approach is to study certain structures, known as motifs [57, 64, 65] analytically and numerically . While this approach is prevalent and addresses specific structures of interest, it fails to provide a comprehensive understanding of the dynamics associated with all synaptic weights.

The other approach reduces the dimensionality by studying the dynamics of the order parameters [27, 48, 57, 59–63]. Then, in some studies, a complementary step is taken that involves performing a linear stability analysis around these order parameters; see e.g. [48].

Here, we applied the latter approach. However, unlike other studies, we did not assume that the neural response was linear. Instead, to facilitate the analysis, we reduced the dimensionality when deriving the cross-correlations of the neural responses. This enabled us to capture the non-linear effects of the neural responses, such as the transition between a fixed-point and a rhythmic activity.

We applied our approach to investigate the emergence of a specific function within a particular model architecture. However, its applicability extends beyond this specific task or structure. We believe this framework can be utilized to explore the development and maintenance of a wide range of dynamic activities in the central nervous system, including the generation of rhythmic patterns by Central Pattern Generators, the sharpening of stimulus selectivity, predictive coding, and memory formation. By capturing the interplay between synaptic plasticity and network dynamics, our approach offers a foundational tool for addressing these complex phenomena.

Appendix A: Description of the analytical derivations

1. Stability analysis of the homogeneous fixed point solutions

The firing rates dynamics for a homogeneous neural activity is given by

$$\begin{aligned}\dot{m}_X(t) = &- m_X(t) + \big[1 + J_{XE} m_E(t-d) \\ &- J_{XI} m_I(t-d)\big]_+, \quad X \in \{E, I\}.\end{aligned} \tag{A1}$$

These dynamics have two kinds of fixed point solutions, $\mathbf{m}^* \equiv \begin{pmatrix} m_E^* \\ m_I^* \end{pmatrix}$. In one (OI) the inhibitory population fully suppresses the excitatory activity; i.e., $m_E^* = 0$, with the requirement that the total input to the excitatory population be negative, $1 - J_{EI} m_I^* < 0$, leading to

$$\mathbf{m}^* = \frac{1}{1 + J_{II}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{A2}$$

and the existence condition for this solution; i.e., complete suppression of the excitatory population, is $J_{EI} > 1 + J_{II}$. In order to study the stability of the OI state we introduce a perturbation $m_I = m_I^* + \delta m_I(t)$, with $\delta m_I(t) = \delta m_I e^{\lambda t}$. The OI fixed point loses its stability in a Hopf bifurcation, $\lambda = i\omega$ ($\omega \in \mathbb{R}$) for

$$J_{II} = -\frac{1}{\cos(\omega d)} \tag{A3}$$

$$\omega = -\tan(\omega d) \tag{A4}$$

where the solution to these equations is found in $\omega \in [\pi/2d, \pi/d]$.

In the case where the excitatory population is not fully suppressed, $J_{EI} < 1 + J_{II}$, the fixed point (FP) is given by

$$\mathbf{m}^* = \frac{1}{(1 - J_{EE})(1 + J_{II}) + \bar{J}^2} \begin{pmatrix} 1 + J_{II} - J_{EI} \\ J_{IE} + 1 - J_{EE} \end{pmatrix} \tag{A5}$$

where $\bar{J} \equiv \sqrt{J_{EI} J_{IE}}$.

Introducing a perturbation $\mathbf{m} = \mathbf{m}^* + \delta\mathbf{m}(t)$, with $\delta\mathbf{m}(t) = \delta\mathbf{m} e^{\lambda t}$, to this fixed point we obtain from eq. (A1)

$$(\lambda \mathbf{I} - \mathbf{A})\, \delta\mathbf{m} = 0 \tag{A6}$$

where

$$\mathbf{A} = \begin{pmatrix} -1 + J_{EE} e^{-\lambda d} & -J_{EI} e^{-\lambda d} \\ J_{IE} e^{-\lambda d} & -1 - J_{II} e^{-\lambda d} \end{pmatrix}. \tag{A7}$$

Requiring $\det[\lambda \mathbf{I} - \mathbf{A}] = 0$ leads to

$$\begin{aligned}(\lambda + 1)^2 &+ e^{-\lambda d}(-J_{EE} + J_{II})(\lambda + 1) \\ &+ \left(\bar{J}^2 - \tilde{J}^2\right) e^{-2\lambda d} = 0\end{aligned} \tag{A8}$$

where $\tilde{J} = \sqrt{J_{EE} J_{II}}$. Solving for $\lambda + 1$ results in two equations for the eigenvalues

$$\lambda_\pm + 1 = g_\pm e^{-\lambda_\pm d} \tag{A9}$$

where

$$g_\pm = \frac{J_{EE} - J_{II} \pm \sqrt{(J_{EE} + J_{II})^2 - 4\bar{J}^2}}{2}. \tag{A10}$$

For $J_{EE} > 2\bar{J} - J_{II}$, $g_\pm \in \mathbb{R}$. If, in addition, $\lambda_\pm \in \mathbb{R}$, instability toward the D state occurs when $g_+ > 1$; i.e., $J_{EE} > 1 + \bar{J}^2/(1 + J_{II})$. For $J_{EE} = 2\bar{J} - J_{II}$, then $g_+ = g_- \in \mathbb{R}$, and if $\lambda_+ = \lambda_- \in \mathbb{R}$, then instability towards the D regime occurs when $g_\pm > 1$; i.e., $J_{EE} > 2 + J_{II}$ (or equivalently $\bar{J} > 1 + J_{II}$). Therefore, the D domain is obtained by the unification of the two conditions

$$J_{EE} > \begin{cases} 1 + \bar{J}^2/(1 + J_{II}) & \bar{J} < \bar{J}_{tc} \\ 2\bar{J} - J_{II} & \bar{J} > \bar{J}_{tc} \end{cases} \tag{A11}$$

where $\bar{J}_{tc} = 1 + J_{II}$.

In the case where $g_\pm = ge^{\pm i\varphi} \in \mathbb{C}$, that is $J_{EE} < 2\bar{J} - J_{II}$, a Hopf bifurcation; i.e., $\lambda = i\omega$ with $\omega \in \mathbb{R}$, emerges from eq. (A9) when

$$i\omega + 1 = ge^{-i(\omega d \mp \varphi)} \tag{A12}$$

which yields two eigenvalue equations

$$g = \frac{1}{\cos(\omega d \mp \varphi)} \tag{A13}$$

$$\omega = -\tan(\omega d \mp \varphi) \tag{A14}$$

where $g = |g_\pm| = \sqrt{\bar{J}_d^2 - \tilde{J}^2}$ and $\varphi = \arg(g_+)$ ($\cos\varphi = \frac{J_{EE} - J_{II}}{2g}$). The set of equations with $-\varphi$ has a solution at $\omega_d \in [0, \pi/2d]$ (note that $-\omega_d$ is the solution to the set of equations with $+\varphi$). The Hopf bifurcation is terminated, $\omega \to 0$, at $(\bar{J}, J_{EE}) = (\bar{J}_{tc}, \bar{J}_{tc} + 1)$.

A different Hopf bifurcation occurs for $g_\pm \in \mathbb{R}$, yielding

$$g_\pm = \frac{1}{\cos(\omega d)} \tag{A15}$$

$$\omega = -\tan(\omega d). \tag{A16}$$

Thus, we obtain oscillations with a similar frequency as in the case of the oscillations that emerge from the OI state; compare eq. (A4) and eq. (A16). These oscillations are characterized by a fast frequency ($\omega \in [\pi/2d, \pi/d]$). For example, in our numerical analysis $f = \omega/2\pi \approx 142$Hz for typical values $\tau_m = 2.5d = 5$ms. Since our focus is on rhythmic activity emerging from the interplay of the E and I activities, we restricted our study to moderate values of I-to-I connectivity, $J_{II} < 1$, in which these fast oscillations do not occur. In addition, for convenience we also took $J_{EE} > J_{II}$.

2. The homogeneous oscillatory solution on the bifurcation line

In the case of positive input to the inhibitory population, the global firing rates can be rescaled in the following manner

$$\begin{aligned} m_E(t) &= \left(1 - \frac{J_{EI}}{1 + J_{II}}\right)\left(r(t+d) + \dot{r}(t+d) + J_{II} r(t)\right) \\ m_I(t) &= \frac{1}{1 + J_{II}} + \left(J_{IE} - \frac{\bar{J}^2}{1 + J_{II}}\right) r(t) \end{aligned} \tag{A17}$$

Thus expressing the dynamics, eq. (2), via a single variable r, obeying:

$$\begin{aligned} \ddot{r}(t) + 2\dot{r}(t) + r(t) = \\ - J_{II}\left(r(t-d) + \dot{r}(t-d)\right) + [h(t)]_+ \end{aligned} \tag{A18}$$

where

$$\begin{aligned} h(t) = 1 + J_{EE}\big(r(t-d) + \dot{r}(t-d) \\ + J_{II} r(t-2d)\big) - \bar{J}^2 r(t-2d). \end{aligned} \tag{A19}$$

The transition from a fixed point state to oscillations occurs when the fixed point loses its stability. On the bifurcation line the solution can be written in the form of $r(t) = \bar{r} + \tilde{r}\cos(\omega_d t)$, where $\bar{r}$ is given by the fixed point solution of eq. (A18):

$$\bar{r} = \frac{1}{(1 - J_{EE})(1 + J_{II}) + \bar{J}_d^2}. \tag{A20}$$

On the bifurcation line, the oscillation amplitude, $\tilde{r}$, is given by the condition that the minimum of $h(t)$ along the cycle is zero, yielding

$$\tilde{r} = \frac{(1 + J_{II})\bar{r}}{g\bar{J}_d}. \tag{A21}$$

From eq. (A17), on the bifurcation line, the homogeneous firing rates are in the form of $m_I(t) = \bar{m}_I + \tilde{m}_I \cos(\omega_d t)$ and $m_E(t) = \bar{m}_E + \tilde{m}_E \cos(\omega_d t + \psi)$, with

$$\bar{\mathbf{m}} = \frac{1}{(1 - J_{EE})(1 + J_{II}) + \bar{J}_d^2} \begin{pmatrix} 1 + J_{II} - \bar{J}_d^2/J_{IE} \\ J_{IE} + 1 - J_{EE} \end{pmatrix} \tag{A22}$$

$$\tilde{\mathbf{m}} = \bar{m}_E/g \begin{pmatrix} 1 \\ J_{IE}/\bar{J}_d \end{pmatrix} \tag{A23}$$

$$\cos\psi = \frac{\hat{J}}{\bar{J}_d} \tag{A24}$$

where $\hat{J} \equiv (J_{EE} + J_{II})/2$. Note that $\bar{\mathbf{m}}$ equals the fixed point $\mathbf{m}^*$ on the bifurcation line, i.e., $\bar{J} = \bar{J}_d$.

3. Perturbative approach in the Fixed Point domain

It is convenient to represent the firing rates of a network by an $N = N_E + N_I$ dimensional vector $\mathbf{m} = \begin{pmatrix} \mathbf{m_E} \\ \mathbf{m_I} \end{pmatrix} \in \mathbb{R}^N$ where $\mathbf{m_E} \in \mathbb{R}^{N_E}$ and $\mathbf{m_I} \in \mathbb{R}^{N_I}$ represent the activity of the excitatory and inhibitory populations, respectively. The neural dynamics, eq. (1), is now written as

$$\dot{\mathbf{m}}(t) = -\mathbf{m}(t) + \left[\sqrt{N}\mathbf{u} + \mathbf{J}\mathbf{m}(t-d)\right]_+ \tag{A25}$$

where the connectivity matrix, $\mathbf{J} \in \mathbb{R}^{N \times N}$

$$\mathbf{J} = \begin{pmatrix} \mathbf{J_{EE}}/N_E & -\mathbf{J_{EI}}/N_I \\ \mathbf{J_{IE}}/N_E & -\mathbf{J_{II}}/N_I \end{pmatrix} \tag{A26}$$

consists of the intra connection matrices $\mathbf{J_{EE}} \in \mathbb{R}^{N_E \times N_E}$, $\mathbf{J_{II}} \in \mathbb{R}^{N_I \times N_I}$ and the inter connection matrices $\mathbf{J_{EI}} \in \mathbb{R}^{N_E \times N_I}$ and $\mathbf{J_{IE}} \in \mathbb{R}^{N_I \times N_E}$. The external input to the entire population is represented by $\sqrt{N}\mathbf{u}$ where $\mathbf{u} \in \mathbb{R}^N$ is a uniform normalized vector; i.e. $u_i = 1/\sqrt{N}$.

The fixed point solution to eq. (A25) is

$$\mathbf{m}^* = \sqrt{N}(\mathbf{I} - \mathbf{J})^{-1}\mathbf{u}, \tag{A27}$$

where $\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix. The connectivity matrix can be written as $\mathbf{J} = \langle \mathbf{J} \rangle + \delta \mathbf{J}$ where $\langle \mathbf{J} \rangle$ is the homogeneous matrix $\langle J_{XY} \rangle_{ij} = J_{XY}$ $(X, Y = E, I)$, and $\delta \mathbf{J}$ is the fluctuations. Using this notation, the firing rates, $\mathbf{m}^*$, are given by

$$\begin{aligned} \mathbf{m}^* &= \sqrt{N} \left(\mathbf{I} - \langle \mathbf{J} \rangle - \delta \mathbf{J} \right)^{-1} \mathbf{u} \\ &= \sqrt{N} \left[\left(\mathbf{I} - \langle \mathbf{J} \rangle \right) \left(\mathbf{I} - \left(\mathbf{I} - \langle \mathbf{J} \rangle \right)^{-1} \delta \mathbf{J} \right) \right]^{-1} \mathbf{u} \\ &= \sqrt{N} \left(\mathbf{I} - \left(\mathbf{I} - \langle \mathbf{J} \rangle \right)^{-1} \delta \mathbf{J} \right)^{-1} \left(\mathbf{I} - \langle \mathbf{J} \rangle \right)^{-1} \mathbf{u}. \end{aligned} \tag{A28}$$

For small fluctuations, we can expand $\mathbf{m}^*$ to a leading order in $\delta \mathbf{J}$

$$\mathbf{m}^* \approx \mathbf{m}^{*(0)} + \mathbf{m}^{*(1)} \tag{A29}$$

where

$$\mathbf{m}^{*(0)} = \sqrt{N} \left(\mathbf{I} - \langle \mathbf{J} \rangle \right)^{-1} \mathbf{u} \tag{A30}$$

is the homogeneous fixed point solution, and

$$\mathbf{m}^{*(1)} = \left(\mathbf{I} - \langle \mathbf{J} \rangle \right)^{-1} \delta \mathbf{J} \mathbf{m}^{*(0)}. \tag{A31}$$

Note that $\delta \mathbf{J} \mathbf{m}^{*(0)}$ is an eigenvector of $\langle \mathbf{J} \rangle$ with zero eigenvalue. Consequently $(\mathbf{I} - \langle \mathbf{J} \rangle)^{-1} \delta \mathbf{J} \mathbf{m}^{*(0)} = \mathbf{I} \delta \mathbf{J} \mathbf{m}^{*(0)}$, yielding

$$\mathbf{m}^{*(1)} = \delta \mathbf{J} \mathbf{m}^{(0)}. \tag{A32}$$

Here we only considered variability in the interconnections; i.e., $\delta J_{XX,ij} = 0$; hence, the first order term of $m^*_{X,i}$ is

$$m^{*(1)}_{E,i} = -m^*_I \delta s_{E,i} \tag{A33}$$

$$m^{*(1)}_{I,i} = m^*_E \delta s_{I,i} \tag{A34}$$

where $\delta s_{X,i} = \frac{1}{N_Y} \sum_{j=1}^{N_Y} \delta J_{XY,ij}$, and $X \neq Y$. To a leading order in the fluctuations we obtain eqs. (8) and (9).

4. Perturbative approach in the R domain

In critical rhythmogenesis, once the system reaches the R domain it will remain near the bifurcation line. Consequently, to analyze the STDP dynamics in the R domain, we can approximate the neuronal responses by their linear dynamics on the bifurcation line. On the bifurcation line, the firing rate of neuron n in population Z is given by $m_{Z,n}(t) = \bar{m}_{Z,n} + \tilde{m}_{Z,n} e^{i(\omega t + \psi_{Z,n})}$, where for mathematical convenience we used complex notation for $m_{Z,n}(t)$ with the understanding that the firing rates are given by the real part. For small fluctuations in the synaptic weights, the firing rates can be expanded around the spatially homogeneous firing rates (i.e. for $\mathbf{J} = \langle \mathbf{J} \rangle$) to first order in $\delta \mathbf{J}$, writing $m_{Z,n}(t) \approx m_Z(t) + \delta m_{Z,n}(t)$ where

$$m_Z(t) = \bar{m}_Z(t) + \tilde{m}_Z e^{i(\omega t + \psi_Z)} \tag{A35}$$

$$\begin{aligned} \delta m_{Z,n}(t) =& \delta \bar{m}_{Z,n} + \delta \tilde{m}_{Z,n} e^{i(\omega t + \psi_Z)} \\ &+ i \delta \psi_{Z,n} \tilde{m}_Z e^{i(\omega t + \psi_Z)} \end{aligned} \tag{A36}$$

where $\psi_I = 0$, $\psi_E = \psi$. The fluctuations in the temporal means, $\delta \bar{m}_{Z,n}$, are given by eqs. (A33) and (A34).

The time dependent fluctuations, $\delta \tilde{m}_{Z,n}$ and $\delta \psi$, are given by

$$\begin{aligned} & i\omega \left(\delta \tilde{m}_{X,i} + i \delta \psi_{X,i} \tilde{m}_X \right) = \\ & - \delta \tilde{m}_{X,i} - i \delta \psi_{X,i} \tilde{m}_X + \tilde{m}_Y e^{i(\psi_Y - \psi_X - \omega d)} Q_Y \delta s_{X,i} \end{aligned} \tag{A37}$$

where $\delta s_{X,i} = \frac{1}{N_Y} \sum_{j=1}^{N_Y} \delta J_{XY,ij}$, $Q_E = 1$, $Q_I = -1$, and $X \neq Y$. Solving the two equations for the real and imaginary parts of eq. (A37) we obtain

$$\delta \tilde{m}_{E,i} = -\frac{\tilde{m}_I \delta s_{E,i}}{g} \cos(\varphi + \psi) \tag{A38}$$

$$\delta \tilde{m}_{I,i} = \frac{\tilde{m}_E \delta s_{I,i}}{g} \cos(\varphi - \psi) \tag{A39}$$

$$\delta \psi_{E,i} = \frac{\tilde{m}_I \delta s_{E,i}}{\tilde{m}_E g} \sin(\varphi + \psi) \tag{A40}$$

$$\delta \psi_{I,i} = -\frac{\tilde{m}_E \delta s_{I,i}}{\tilde{m}_I g} \sin(\varphi - \psi) \tag{A41}$$

where ψ is given by eq. (A24). Using eqs. (A38) to (A41), to a leading order in the fluctuations $\delta J_{UV,kl}$, we obtain eqs. (10) and (11).

5. STDP dynamics of the fluctuations

The STDP dynamics for a synaptic weight $J_{XY,ij}$, eq. (17), is given by

$$\dot{J}_{XY,ij} = \lambda_{XY} \int_{-\infty}^{\infty} K_{XY,ij}(\Delta) \Gamma_{XY,ij}(-\Delta) \, d\Delta \tag{A42}$$

where $K_{XY,ij}(\Delta) = f_{XY}(J_{XY,ij}) K_{XY,+}(\Delta) - \alpha K_{XY,-}(\Delta)$ and $X \neq Y$.

Considering small deviations from the homogeneous solution, $\delta J_{XY,ij} = J_{XY,ij} - J_{XY}$, we can expand eq. (A42) to the first order in these deviations. In the FP domain we obtain

$$\delta \dot{J}_{XY,ij} = \bar{M}_{YY} \delta s_{X,i} + \bar{M}_{YX} \delta s_{Y,j} + \bar{F}_Y \delta J_{XY,ij} \tag{A43}$$

and in the R domain near the bifurcation line

$$\delta \dot{J}_{XY,ij} = \tilde{M}_{YY}\delta s_{X,i} + \tilde{M}_{YX}\delta s_{Y,j} + \tilde{F}_Y \delta J_{XY,ij} \quad \text{(A44)}$$

where: $\bar{M}_{UV} = \lambda_U \bar{K}_U Q_V \bar{m}_V^2$, $\bar{F}_U = \lambda_U f'_U \bar{K}_U \bar{\Gamma}$, $\tilde{M}_{UV} = \bar{M}_{UV} + \lambda_U \tilde{K}_U (Q_V \varphi) Q_V \tilde{m}_V^2 / (2g)$, $\tilde{F}_U = \bar{F}_U + \lambda_U f'_U \tilde{K}_{U,+}(\psi)\tilde{\Gamma}$, $f_U / dJ_{IE}\big|_{J^*_{IE}}$, $Q_E = 1$ and $Q_I = -1$.

Eqs. (A43) and (A44) describe the linear dynamics of the fluctuations, which are characterized by four families of eigenvectors. Eigenvectors of family I represent fluctuations in the inhibitory synapses to an excitatory cell k ($k \in 1, \ldots, N_E$), such that $\sum_l \Delta J_{EI,kl} = 0$. These eigenvectors are characterized by zero eigenvalues, $\bar{\sigma}_{\rm I} = \tilde{\sigma}_{\rm I} = 0$, where we denoted by $\bar{\sigma}_{\rm A}$ ($\tilde{\sigma}_{\rm A}$) the eigenvalues of family A$\in$ {I, II, III, IV} in FP (R) region. The multiplicity of this family is $N_E(N_I - 1)$.

Eigenvectors of family II represent fluctuations in the excitatory synapses to an inhibitory cell, k ($k \in \{1, \ldots N_I\}$), such that $\sum_l \Delta J_{IE,kl} = 0$. The eigenvalues of these eigenvectors in the FP and R regions are $\bar{\sigma}_{\rm II} = \bar{F}_E < 0$ and $\tilde{\sigma}_{\rm II} = \tilde{F}_E < 0$, respectively. The multiplicity of this family are $N_I(N_E - 1)$.

Eigenvectors of family III - represent fluctuations in the firing of excitatory neurons such that a constant mean excitatory activity, $m_E(t) = \sum_i m_{E,i}(t)$, is maintained. In these eigenvectors, all the inhibitory connections to each excitatory neuron $k \in \{1, \ldots M_E\}$ are modified by the same amount, $\Delta J_{EI,kj_1} = \Delta J_{EI,kj_2}$ $\forall j_1, j_2$, and all of the connections of excitatory neuron k to the inhibitory neurons obey, $\Delta J_{IE,i_1k} = \Delta J_{IE,i_2k} = \bar{\eta}_{\rm III}\Delta J_{EI,ki_1}$ $\forall i_1, i_2$ in the FP region, and $\Delta J_{IE,i_1k} = \Delta J_{IE,i_2k} = \tilde{\eta}_{\rm III}\Delta J_{EI,ki_1}$ $\forall i_1, i_2$ in the R region. The ratios in synaptic fluctuation are given by $\bar{\eta}_{\rm III} = \bar{M}_{EI}/\left(\bar{M}_{II} - \bar{F}_E\right)$ and $\tilde{\eta}_{\rm III} = \tilde{M}_{EI}/\left(\tilde{M}_{II} - \tilde{F}_E\right)$. The eigenvalues of these eigenvectors in the FP and R regions are negative, $\bar{\sigma}_{\rm III} = \bar{M}_{II} < 0$ and $\tilde{\sigma}_{\rm III} = \tilde{M}_{II} < 0$. Eigenvectors in this family are determined by the deviations of the excitatory activities from the mean excitation, $\mathbf{m}_E - m_E$. Thus, from the condition of fixed mean activity, $\sum_i m_{E,i} = const$, the multiplicity of this family is $N_E - 1$.

Eigenvectors of family IV - represent fluctuations in the firings of individual inhibitory neurons such that the mean inhibitory activity, $m_I(t) = \sum_i m_{I,i}(t)$, remains constant. In these eigenvectors, all the excitatory connections to each inhibitory neuron k are modified by the same amount, $\Delta J_{IE,kj_1} = \Delta J_{IE,kj_2}$ $\forall j_1, j_2$, and all of the connections of excitatory neuron k to inhibitory neurons obey, $\Delta J_{EI,i_1k} = \Delta J_{EI,i_2k} = \bar{\eta}_{\rm IV}^{-1}\Delta J_{IE,ki_1}$ $\forall i_1, i_2$ in the FP region, and $\Delta J_{EI,i_1k} = \Delta J_{EI,i_2k} = \tilde{\eta}_{\rm IV}^{-1}\Delta J_{IE,ki_1}$ $\forall i_1, i_2$ in the R region. The ratios in synaptic fluctuation are given by $\bar{\eta}_{\rm IV} = \left(\bar{M}_{EE} + \bar{F}_E\right)/\bar{M}_{IE}$ and $\tilde{\eta}_{\rm IV} = \left(\tilde{M}_{EE} + \tilde{F}_E\right)/\tilde{M}_{IE}$. The eigenvalues of these eigenvectors in the FP and R regions are $\bar{\sigma}_{\rm IV} = \bar{M}_{EE} + \bar{F}_E$ and $\tilde{\sigma}_{\rm IV} = \tilde{M}_{EE} + \tilde{F}_E$, respectively. Fig. 5a depicts the eigenvalues $\bar{\sigma}_{\rm IV}$ and $\tilde{\sigma}_{\rm IV}$ as function of α for different values of $\tau_{E,-}$ (differentiated by color). As can be seen, these eigenvalues may have different signs in the FP and R regions (see section A 6 below). The multiplicity of these eigenvectors is $N_I - 1$, from the condition of fixed mean inhibition $\sum_i m_{I,i} = const$.

6. STDP dynamics of synaptic fluctuations in critical rhythmogenesis

Once the system reaches the bifurcation line, the order parameters J_{XY} alternate between the FP and R regions. After a transient, the order parameters will converge to alternate between asynchronous and rhythmic activities around the critical point $\mathbf{J}^*$. The eigenvectors delineated in section A 5 result from the linearization of the STDP dynamics. Thus, in general, there are two sets of eigenvectors: one for the FP region and one for the R region.

Eigenvectors of family I (II) in the FP region are also eigenvectors of family I (II) in the R region and vice-versa. In contrast, although eigenvectors of family III (IV) have a similar structure in the FP and R regions, they are not the same, since $\bar{\eta}_{\rm III} \neq \tilde{\eta}_{\rm III}$ ($\bar{\eta}_{\rm IV} \neq \tilde{\eta}_{\rm IV}$). However, any eigenvector of family III (IV) from the FP region can be expressed as an eigenvector of family III (IV) from the R region plus an eigenvector of family II (I), and vice-versa. Thus, the subspaces spanned by the eigenvectors of family II and III (I and IV) in the FP and the R regions are the same. Secondly, these subspaces are invariant under the STDP dynamics that alternates between the FP and R regions.

The eigenvalues of families II and III are negative. Consequently, any fluctuation in the subspace spanned by the eigenvectors of families II and III will decay. Consider now the dynamics of a small fluctuation in the subspace spanned by the eigenvectors of family I and IV. The eigenvalue of family IV is negative in the R region and positive in the FP region. Below, we present a heuristic argument to determine whether or not fluctuations in this subspace will decay with time.

Let us denote initial conditions by $\mathbf{w}(0)$ and examine its dynamics over a full cycle of FP-R alternation. Assuming the initial state of the system is FP, it is convenient to express the initial condition by eigenvectors of the FP region: $\mathbf{w}(0) = x(0)\mathbf{v}^{(\rm FP)} + y(0)\mathbf{u}$, where $\mathbf{v}^{(\rm FP)}$ is an eigenvector from family IV, $\mathbf{u}$ is an eigenvector of family I, and $x(t), y(t) \in \mathbb{R}$. After a duration of $T^{(\rm FP)}$ in the FP region the fluctuation becomes $\mathbf{w}(T^{(\rm FP)}) = x(T^{(\rm FP)})\mathbf{v}^{(\rm FP)} + y(T^{(\rm FP)})\mathbf{u}$, with

$$\begin{pmatrix} x(T^{(\rm FP)}) \\ y(T^{(\rm FP)}) \end{pmatrix} = \bar{\mathbf{E}} \begin{pmatrix} x(0) \\ y(0) \end{pmatrix}, \quad \text{(A45)}$$

$$\bar{\mathbf{E}} = \begin{bmatrix} e^{\bar{\sigma}_{\rm IV} T^{(\rm FP)}} & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{(A46)}$$

where assuming small fluctuations we can use the linearization of the STDP dynamics.

After transitioning to the R state at time $T^{(\rm FP)}$ it is convenient to express the fluctuation by the eigenvectors of the R region: $\mathbf{w}(T^{(\rm FP)}) = x'(T^{(\rm FP)})\mathbf{v}^{(\rm R)} + y'(T^{(\rm FP)})\mathbf{u}$,

where $\mathbf{v}^{(\mathrm{R})}$ is an eigenvector from family IV, and $x', y' \in \mathbb{R}$ can be obtained by a linear transformation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{T} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{A47}$$

After a duration of $T^{(\mathrm{R})}$ in the R region the fluctuation evolves to $\mathbf{w}(T^{(\mathrm{FP})} + T^{(\mathrm{R})}) = x'(T^{(\mathrm{FP})} + T^{(\mathrm{R})})\mathbf{v}^{(\mathrm{R})} + y'(T^{(\mathrm{FP})} + T^{(\mathrm{R})})\mathbf{u}$, with

$$\begin{pmatrix} x'(T^{(\mathrm{FP})} + T^{(\mathrm{R})}) \\ y'(T^{(\mathrm{FP})} + T^{(\mathrm{R})}) \end{pmatrix} = \tilde{\mathbf{E}} \begin{pmatrix} x'(T^{(\mathrm{FP})}) \\ y'(T^{(\mathrm{FP})}) \end{pmatrix}, \tag{A48}$$

$$\tilde{\mathbf{E}} = \begin{bmatrix} e^{\tilde{\sigma}_{\mathrm{IV}} T^{(\mathrm{R})}} & 0 \\ 0 & 1 \end{bmatrix}. \tag{A49}$$

Transitioning back to the FP region, the fluctuation $\mathbf{w}(T^{(\mathrm{FP})} + T^{(\mathrm{R})})$ can be expressed again by $\mathbf{v}^{(\mathrm{FP})}$ and $\mathbf{u}$, using the inverse transformation matrix $\mathbf{T}^{-1}$. Therefore, an entire cycle affects the fluctuation as follows

$$\mathbf{w}(T^{(\mathrm{FP})} + T^{(\mathrm{R})}) = \mathbf{L}\mathbf{w}(0) \tag{A50}$$

with

$$\mathbf{L} = \mathbf{T}^{-1}\tilde{\mathbf{E}}\mathbf{T}\bar{\mathbf{E}}. \tag{A51}$$

Hence, the determinant of $\mathbf{L}$

$$\det \mathbf{L} = \exp\left(\bar{\sigma}_{\mathrm{IV}} T^{(\mathrm{FP})} + \tilde{\sigma}_{\mathrm{IV}} T^{(\mathrm{R})}\right). \tag{A52}$$

determines whether the fluctuation will grow or vanish in time.

To complete the argument we need to estimate $T^{(\mathrm{FP})}$ and $T^{(\mathrm{R})}$. The time spent in each region is determined by the dynamics of the order parameters around $\mathbf{J}^*$. We estimated $T^{(\mathrm{FP})}$ by the inverse of the rate of exiting the FP region: $T^{(\mathrm{FP})} = 1/|\dot{\mathbf{J}}^{(\mathrm{FP})}|$, and similarly in the R region. At the critical point, $\hat{\dot{\mathbf{J}}}^{(\mathrm{FP})} = -\hat{\dot{\mathbf{J}}}^{(\mathrm{R})}$; hence, the magnitude of the ratio of $\dot{\mathbf{J}}$ is equal to the ratio of the magnitude of each of its components, $T^{(\mathrm{R})}/T^{(\mathrm{FP})} = |\dot{J}_{EI}^{(\mathrm{FP})}/\dot{J}_{EI}^{(\mathrm{R})}| = |\dot{J}_{IE}^{(\mathrm{FP})}/\dot{J}_{IE}^{(\mathrm{R})}|$. Thus,

$$\log(\det \mathbf{L}) = \tilde{\sigma}_{\mathrm{IV}}/|\dot{J}_{EI}^{(\mathrm{R})}| + \bar{\sigma}_{\mathrm{IV}}/|\dot{J}_{EI}^{(\mathrm{FP})}|, \tag{A53}$$

which is presented in fig. 5b.

7. Specifics of the numerical simulations and figures

The numerical simulations presented in this study were run with MATLAB. Within the R region, the neural and STDP dynamics described in eqs. (1), (2), (17) and (18) were solved using Euler's method for differential equations. Specifically, eqs. (1) and (2) were integrated over a time interval of $60\tau_m$ with a time step of $\Delta t = 0.01\tau_m$. The final period of $m_{X,i}(t)$ and $m_{Y,j}(t)$ ($m_X(t)$ and $m_Y(t)$) was used to compute the cross-correlation $\Gamma_{XY,ij}(\Delta)$ ($\Gamma_{XY}(\Delta)$) over a single cycle. By using the periodicity of $\Gamma_{XY,ij}(\Delta)$ ($\Gamma_{XY}(\Delta)$), the integral in eq. (17) (eq. (18)) was computed with integration limits of $[-200\tau_m, 200\tau_m]$ and a time bin of $\Delta = 0.01\tau_m$.

For each iteration of the STDP dynamics in eq. (17) (depicted in figs. 6a to 6c), Gaussian white noise $\xi_{XY,ij}$ was added to its right-hand side. The noise terms were sampled from distributions $\xi_{EI,ij} \sim \mathcal{N}(0, 9 \times 10^{-8})$ and $\xi_{IE,ij} \sim \mathcal{N}(0, 9 \times 10^{-6})$.

ACKNOWLEDGMENTS

This research was supported by the Israel Science Foundation (ISF) grant numbers 300/16 and 624/22.

[1] K Fox. Anatomical pathways and molecular mechanisms for plasticity in the barrel cortex. *Neuroscience*, 111(4):799–814, 2002.

[2] Yonatan Loewenstein, Annerose Kuras, and Simon Rumpel. Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. *Journal of Neuroscience*, 31(26):9481–9488, 2011.

[3] Yonatan Loewenstein, Uri Yanover, and Simon Rumpel. Predicting the dynamics of network connectivity in the neocortex. *Journal of Neuroscience*, 35(36):12535–12544, 2015.

[4] Amir Minerbi, Roni Kahana, Larissa Goldfeld, Maya Kaufman, Shimon Marom, and Noam E Ziv. Long-term relationships between synaptic tenacity, synaptic remodeling, and network activity. *PLoS biology*, 7(6):e1000136, 2009.

[5] Liran Hazan and Noam E Ziv. Activity dependent and independent determinants of synaptic size diversity. *Journal of Neuroscience*, 40(14):2828–2848, 2020.

[6] Noam E Ziv and Naama Brenner. Synaptic tenacity or lack thereof: spontaneous remodeling of synapses. *Trends in neurosciences*, 41(2):89–99, 2018.

[7] Nobuaki Yasumatsu, Masanori Matsuzaki, Takashi Miyazaki, Jun Noguchi, and Haruo Kasai. Principles of long-term dynamics of dendritic spines. *Journal of Neuroscience*, 28(50):13592–13608, 2008.

[8] György Buzsaki. Rhythms of the brain, 2006.

[9] György Buzsáki and Walter Freeman. Editorial overview: brain rhythms and dynamic coordination. *Current opinion in neurobiology*, 31:v–ix, 2015.

[10] Marco Bocchio, Sadegh Nabavi, and Marco Capogna. Synaptic plasticity, engrams, and network oscillations in amygdala circuits for storage and retrieval of emotional memories. *Neuron*, 94(4):731–743, 2017.

[11] Maoz Shamir, Oded Ghitza, Steven Epstein, and Nancy

Kopell. Representation of time-varying stimuli by a network exhibiting oscillations on a faster time scale. *PLoS computational biology*, 5(5):e1000370, 2009.

[12] Aryeh Hai Taub, Rita Perets, Eilat Kahana, and Rony Paz. Oscillations synchronize amygdala-to-prefrontal primate circuits during aversive learning. *Neuron*, 97(2):291–298, 2018.

[13] Charles M Gray, Peter König, Andreas K Engel, and Wolf Singer. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338(6213):334–337, 1989.

[14] Pascal Fries, John H Reynolds, Alan E Rorie, and Robert Desimone. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 291(5508):1560–1563, 2001.

[15] Narcisse P Bichot, Andrew F Rossi, and Robert Desimone. Parallel and serial neural mechanisms for visual search in macaque area v4. *Science*, 308(5721):529–534, 2005.

[16] Bijan Pesaran, John S Pezaris, Maneesh Sahani, Partha P Mitra, and Richard A Andersen. Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nature neuroscience*, 5(8):805–811, 2002.

[17] Wulfram Gerstner. Time structure of the activity in neural network models. *Physical review E*, 51(1):738, 1995.

[18] N Kopell and GB Ermentrout. Gamma rhythms and beta rhythms have different synchronization properties. *Proceedings of the National Academy of Sciences*, 97(4):1867–1872, 2000.

[19] Christoph Börgers and Nancy Kopell. Synchronization in networks of excitatory and inhibitory neurons with sparse, random connectivity. *Neural computation*, 15(3):509–538, 2003.

[20] Nicolas Brunel and David Hansel. How noise affects the synchronization properties of recurrent networks of inhibitory neurons. *Neural computation*, 18(5):1066–1110, 2006.

[21] Pascal Fries, Danko Nikolić, and Wolf Singer. The gamma cycle. *Trends in neurosciences*, 30(7):309–316, 2007.

[22] Alex Roxin, Nicolas Brunel, and David Hansel. Rate models with delays and the dynamics of large networks of spiking neurons. *Progress of Theoretical Physics Supplement*, 161:68–85, 2006.

[23] Demian Battaglia, Nicolas Brunel, and David Hansel. Temporal decorrelation of collective oscillations in neural networks¡? format?¿ with local inhibition and long-range excitation. *Physical review letters*, 99(23):238106, 2007.

[24] Demian Battaglia and David Hansel. Synchronous chaos and broad band gamma rhythm in a minimal multi-layer model of primary visual cortex. *PLoS computational biology*, 7(10):e1002176, 2011.

[25] Alex Roxin and Ernest Montbrió. How effective delays shape oscillatory dynamics in neuronal networks. *Physica D: Nonlinear Phenomena*, 240(3):323–345, 2011.

[26] Genki Shimizu, Kensuke Yoshida, Haruo Kasai, and Taro Toyoizumi. Computational roles of intrinsic synaptic dynamics. *Current opinion in neurobiology*, 70:34–42, 2021.

[27] Paul Manz and Raoul-Martin Memmesheimer. Purely stdp-based assembly dynamics: Stability, learning, overlaps, drift and aging. *PLOS Computational Biology*, 19(4):e1011006, 2023.

[28] Claudia Clopath, Tobias Bonhoeffer, Mark Hübener, and Tobias Rose. Variance and invariance of neuronal long-term representations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1715):20160161, 2017.

[29] Michael Jan Fauth and Mark CW Van Rossum. Self-organized reactivation maintains and reinforces memories despite synaptic turnover. *Elife*, 8:e43717, 2019.

[30] Daniel Acker, Suzanne Paradis, and Paul Miller. Stable memory and computation in randomly rewiring neural networks. *Journal of neurophysiology*, 122(1):66–80, 2019.

[31] Wulfram Gerstner, Richard Kempter, J Leo Van Hemmen, and Hermann Wagner. A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 383(6595):76–78, 1996.

[32] Henry Markram, Joachim Lübke, Michael Frotscher, and Bert Sakmann. Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. *Science*, 275(5297):213–215, 1997.

[33] Guo-qiang Bi and Mu-ming Poo. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of neuroscience*, 18(24):10464–10472, 1998.

[34] Sen Song, Kenneth D Miller, and Larry F Abbott. Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nature neuroscience*, 3(9):919–926, 2000.

[35] Robert Gütig, Ranit Aharonov, Stefan Rotter, and Haim Sompolinsky. Learning input correlations through nonlinear temporally asymmetric hebbian plasticity. *Journal of Neuroscience*, 23(9):3697–3714, 2003.

[36] Curtis C Bell, Victor Z Han, Yoshiko Sugawara, and Kirsty Grant. Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature*, 387(6630):278–281, 1997.

[37] Li I Zhang, Huizhong W Tao, Christine E Holt, William A Harris, and Mu-ming Poo. A critical window for cooperation and competition among developing retinotectal synapses. *Nature*, 395(6697):37–44, 1998.

[38] Per Jesper Sjöström, Gina G Turrigiano, and Sacha B Nelson. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32(6):1149–1164, 2001.

[39] Melanie A Woodin, Karunesh Ganguly, and Mu-ming Poo. Coincident pre-and postsynaptic activity modifies gabaergic synapses by postsynaptic changes in cl- transporter activity. *Neuron*, 39(5):807–820, 2003.

[40] Robert C Froemke, Ishan A Tsay, Mohamad Raad, John D Long, and Yang Dan. Contribution of individual spikes in burst-induced long-term synaptic modification. *Journal of neurophysiology*, 95(3):1620–1629, 2006.

[41] Daniel E Feldman. The spike-timing dependence of plasticity. *Neuron*, 75(4):556–571, 2012.

[42] Gabi Socolovsky and Maoz Shamir. Robust rhythmogenesis via spike-timing-dependent plasticity. *Physical Review E*, 104(2):024413, 2021.

[43] Yotam Luz and Maoz Shamir. The effect of stdp temporal kernel structure on the learning dynamics of single excitatory and inhibitory synapses. *PloS one*, 9(7):e101109, 2014.

[44] Larry F Abbott and Sacha B Nelson. Synaptic plasticity: taming the beast. *Nature neuroscience*, 3(11):1178–1183, 2000.

[45] Yotam Luz and Maoz Shamir. Oscillations via spike-timing dependent plasticity in a feed-forward model. *PLoS computational biology*, 12(4):e1004878, 2016.

[46] Evgueniy V Lubenov and Athanassios G Siapas. Decoupling through synchrony in neuronal circuits with propagation delays. *Neuron*, 58(1):118–131, 2008.

[47] Richard Kempter, Wulfram Gerstner, and J Leo Van Hemmen. Hebbian learning and spiking neurons. *Physical Review E*, 59(4):4498, 1999.

[48] Anthony N Burkitt, Matthieu Gilson, and J Leo van Hemmen. Spike-timing-dependent plasticity for neurons with recurrent connections. *Biological cybernetics*, 96(5):533–546, 2007.

[49] Kalen P Berry and Elly Nedivi. Spine dynamics: are they all the same? *Neuron*, 96(1):43–55, 2017.

[50] Stefano Fusi, Patrick J Drew, and Larry F Abbott. Cascade models of synaptically stored memories. *Neuron*, 45(4):599–611, 2005.

[51] Marcus K Benna and Stefano Fusi. Computational principles of synaptic memory consolidation. *Nature neuroscience*, 19(12):1697–1706, 2016.

[52] Gianluigi Mongillo, Simon Rumpel, and Yonatan Loewenstein. Inhibitory connectivity defines the realm of excitatory plasticity. *Nature neuroscience*, 21(10):1463–1470, 2018.

[53] Moritz Deger, Moritz Helias, Stefan Rotter, and Markus Diesmann. Spike-timing dependence of structural plasticity explains cooperative synapse formation in the neocortex. *PLoS computational biology*, 8(9):e1002689, 2012.

[54] Moritz Deger, Alexander Seeholzer, and Wulfram Gerstner. Multicontact co-operativity in spike-timing–dependent structural plasticity stabilizes networks. *Cerebral Cortex*, 28(4):1396–1415, 2018.

[55] Sven Dorkenwald, Nicholas L Turner, Thomas Macrina, Kisuk Lee, Ran Lu, Jingpeng Wu, Agnes L Bodor, Adam A Bleckert, Derrick Brittain, Nico Kemnitz, et al. Binary and analog variation of synapses between cortical pyramidal neurons. *Elife*, 11:e76120, 2022.

[56] Lee Susman, Naama Brenner, and Omri Barak. Stable memory with unstable synapses. *Nature communications*, 10(1):4441, 2019.

[57] Gabriel Koch Ocker, Ashok Litwin-Kumar, and Brent Doiron. Self-organization of microcircuits in networks of spiking neurons with plastic synapses. *PLoS computational biology*, 11(8):e1004458, 2015.

[58] Yaroslav Felipe Kalle Kossio, Sven Goedeke, Christian Klos, and Raoul-Martin Memmesheimer. Drifting assemblies for persistent memory: Neuron transitions and unsupervised compensation. *Proceedings of the National Academy of Sciences*, 118(46):e2023832118, 2021.

[59] Matthieu Gilson, Anthony N Burkitt, David B Grayden, Doreen A Thomas, and J Leo Van Hemmen. Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks. i. input selectivity–strengthening correlated input pathways. *Biological cybernetics*, 101:81–102, 2009.

[60] Matthieu Gilson, Anthony N Burkitt, David B Grayden, Doreen A Thomas, and J Leo van Hemmen. Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks. ii. input selectivity—symmetry breaking. *Biological Cybernetics*, 101:103–114, 2009.

[61] Matthieu Gilson, Anthony N Burkitt, David B Grayden, Doreen A Thomas, and J Leo van Hemmen. Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks iii: Partially connected neurons driven by spontaneous activity. *Biological Cybernetics*, 101:411–426, 2009.

[62] Matthieu Gilson, Anthony N Burkitt, David B Grayden, Doreen A Thomas, and J Leo van Hemmen. Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks iv: Structuring synaptic pathways among recurrent connections. *Biological cybernetics*, 101:427–444, 2009.

[63] Matthieu Gilson, Anthony N Burkitt, David B Grayden, Doreen A Thomas, and J Leo van Hemmen. Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks v: self-organization schemes and weight dependence. *Biological Cybernetics*, 103:365–386, 2010.

[64] Neta Ravid Tannenbaum and Yoram Burak. Shaping neural circuits by high order synaptic interactions. *PLOS Computational Biology*, 12(8):e1005056, 2016.

[65] Lisandro Montangie, Christoph Miehl, and Julijana Gjorgjieva. Autonomous emergence of connectivity assemblies via spike triplet interactions. *PLOS Computational Biology*, 16(5):e1007835, 2020.