&

Robust Rhythmogenesis in the Gamma Band via Spike Timing Dependent Plasticity

Gabi Socolovsky[1,2*] and Maoz Shamir[1,2,3]
[1] *Department of Physics, Faculty of Natural Sciences,*
[2] *Zlotowski Center for Neuroscience, and*
[3] *Department of Physiology and Cell Biology, Faculty of Health Sciences,*
Ben-Gurion University of the Negev, Be'er-Sheva, Israel
(Dated: July 22, 2020)

Rhythmic activity in the gamma band (30-100Hz) has been observed in numerous animal species ranging from insects to humans, and in relation to a wide range of cognitive tasks. Various experimental and theoretical studies have investigated this rhythmic activity. The theoretical efforts have mainly been focused on the neuronal dynamics, under the assumption that network connectivity satisfies certain fine-tuning conditions required to generate gamma oscillations. However, it remains unclear how this fine tuning is achieved.

Here we investigated the hypothesis that spike timing dependent plasticity (STDP) can provide the underlying mechanism for tuning synaptic connectivity to generate rhythmic activity in the gamma band. We addressed this question in a modeling study. We examined STDP dynamics in the framework of a network of excitatory and inhibitory neuronal populations that has been suggested to underlie the generation of gamma. Mean field Fokker Planck equations for the synaptic weights dynamics are derived in the limit of slow learning. We drew on this approximation to determine which types of STDP rules drive the system to exhibit gamma oscillations, and demonstrate how the parameters that characterize the plasticity rule govern the rhythmic activity. Finally, we propose a novel mechanism that can ensure the robustness of self-developing processes, in general and for rhythmogenesis in particular.

I. INTRODUCTION

Rhythmic activity in the brain has been observed for more than a century [1, 2]. Oscillations in different frequency bands have been associated with different cognitive tasks and mental states [2–6]. Specifically, rhythmic activity in the Gamma band has been described in association with sensory stimulation [7], attentional selection [8, 9], working memory [10] and other measures [11]. Deviation from normal rhythmic activity has been associated with pathology [12–15].

Considerable theoretical efforts have been devoted to unraveling the neural mechanism responsible for generating rhythmic activity in the Gamma band [16–21]. One possible mechanism is based on delayed inhibitory feedback [21–24]. The basic architecture of this mechanism is composed of one excitatory and one inhibitory neuronal populations, with reciprocal connections (Fig. 1a). A target rhythm is obtained by tuning the strengths of the excitatory and inhibitory interactions (Fig. 1b-1c). However, it is unclear which mechanism results in the required fine-tuning [23, 25].

We hypothesized that activity dependent synaptic plasticity can provide the mechanism for tuning the interaction strengths in order to stabilize a specific rhythmic activity in the gamma band.

Here we focused on spike timing dependent plasticity (STDP) as the rhythmogenic process [23, 25]. Below, we briefly describe STDP and derive the dynamics of the synaptic weights in the limit of slow learning. Since the cross-correlation of neural activity is central to STDP dynamics, we next define the network dynamics and analyze its phase diagram and the dependence of the correlations on the synaptic weights. Using the separation of timescales in the limit of slow learning we analyze STDP dynamics and investigate under what conditions STDP can stabilize a specific rhythmic activity and how the characteristics of the STDP rule govern the resultant rhythmic activity. Finally, we summarize our results, discuss possible extensions and limitations and propose a general principle for robust rhythmogenesis.

II. STDP DYNAMICS

The basic coin of information transfer in the central nervous system is the spike: a short electrical pulse that propagates along the axon (output branch) of the transmitting neuron to synaptic terminals that relay the information to the dendrites (input branch) of the receiving neurons downstream. While spikes are stereotypical, the relayed signal depends on the synaptic weight, which can be thought of as interaction strength. Learning is the process that modifies synaptic weights (the dynamics of the interaction strengths themselves), and typically occurs on a slower timescale than the timescale of the neuronal responses.

STDP is an empirically observed microscopic learning rule in which the modification of the synaptic weight depends on the temporal relation between the spike times

* E-mail me at: gabisoco@post.bgu.ac.il

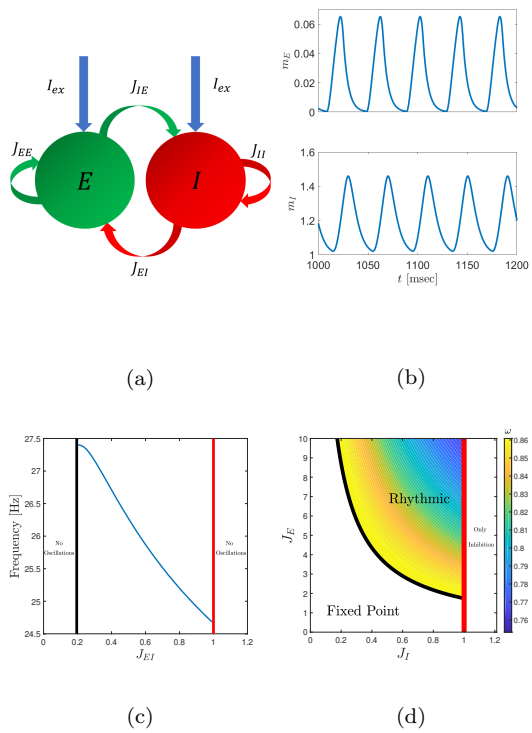

FIG. 1: The excitatory-inhibitory network. (**a**) Model architecture. The neuronal network here is composed of an excitatory (E) and an inhibitory (I) populations with inter- (J_{IE} and J_{EI}) and intra- (J_{II} and J_{EE}, that will be taken to zero hereafter) connections. The interaction is not symmetric and is delayed. (**b**) The oscillatory dynamics of the mean excitatory end inhibitory population firing rates m_E and m_I, respectively, at a frequency of 24.9Hz. Here we used $J_{EE} = J_{II} = 0$, $J_{IE} = 8.91$, $J_{EI} = 0.9$, $\tau_m = d = 5$ms. (**c**) Rhythmic activity. The oscillation frequency is depicted as a function of the strength of the inhibitory to excitatory connection, J_{EI}. The following parameters were implemented here: $J_{EE} = J_{II} = 0$, $\tau_m = d = 5$ms, $J_{IE} = 8.91$. (**d**) Phase diagram of the delayed rate model. Strong inhibition, $J_I > 1$, leads the system to a purely inhibitory fixed point, in which m_E is fully suppressed, and $m_I = 1$. For weak to moderate inhibition, $J_I \in (0, 1)$, and below the black line, the system converges to a fixed point, where both populations are active. For $\bar{J} > \bar{J}_d$ (above the black line) the model exhibits rhythmic activity. For a given time delay, the frequency is governed by $\bar{J}$. When $\bar{J}$ is increased, higher frequencies are observed. Here we used $\tau_m = d = 1$.

of the pre (transmitting) and post (receiving) synaptic neurons [26–30]. Following [31] the change, ΔJ_{ij}, in the synaptic weight J_{ij} from the pre-synaptic neuron j to the post-synaptic neuron i is expressed as the sum of two processes: potentiation (i.e., increasing the synaptic weight) and depression (i.e., decreasing the synaptic weight):

$$\Delta J_{ij} = \lambda \left[K_+(\Delta t) - \alpha K_-(\Delta t) \right], \tag{1}$$

where $\Delta t = t_i - t_j$ is the pre and post spike time difference. Functions $K_\pm(t) \geq 0$ describe the temporal structure of the potentiation (+) and depression (−) of the STDP rule. Parameter α denotes the relative strength of the depression and λ is the learning rate. We assume that learning occurs on a slower timescale than the characteristic timescales that describe neuronal activity (for given fixed synaptic weights).

A wide range of temporal structures of STDP rules has been reported [32–39]. Here we focus on two families of rules. One is composed of temporally symmetric rules and the other is made up of temporally asymmetric rules.

For the temporally symmetric family we apply a difference of Gaussian STDP rule; namely:

$$K_\pm(t) = \frac{1}{\sqrt{2\pi}\tau_\pm} e^{-(t/\tau_\pm)^2/2}, \tag{2}$$

where $\tau_\pm$ denotes the characteristic time scales of the potentiation (+) and depression (−). Consistent with the popular description of the famous Hebb rule that 'neurons that fire together wire together' [40], we refer to the case of $\tau_+ < \tau_-$ as Hebbian, and $\tau_+ < \tau_-$ anti-Hebbian.

For the temporally a-symmetric STDP rules we take:

$$K_\pm(t) = \frac{1}{\tau_\pm} e^{\mp H t/\tau_\pm} \Theta(\pm H t), \tag{3}$$

where $\Theta(x)$ is the Heaviside step function, $\tau_\pm$ are the characteristic time scales of the potentiation (+) and depression (−), and $H = \pm 1$ dictates the Hebbianity of the STDP rule. The rule will be termed Hebbian for $H = 1$, when potentiation occurs in the causal branch, $t_{\text{post}} > t_{\text{pre}}$, and anti-Hebbian for $H = -1$.

Different types of synapses have been reported to exhibit different types of STDP rules. Consequently, there is no a-priori reason to assume that excitatory and inhibitory synapses share the exact same learning rule. In particular, the characteristic time constants $\tau_{E,\pm}$ for excitatory synapses and $\tau_{I,\pm}$ for inhibitory synapses may differ.

Changes to synaptic weights due to the plasticity rule of Eq. (1) at short time intervals occur as a result of either a pre or post-synaptic spike during this interval. Thus,

$$\begin{aligned} \dot{J}_{i,j}(t) = \lambda \rho_i(t) \int_0^\infty \rho_j(t - t') \left[K_+(t') - \alpha K_-(t') \right] dt' \\ + \lambda \rho_j(t) \int_0^\infty \rho_i(t - t') \left[K_+(-t') - \alpha K_-(-t') \right] dt', \end{aligned} \tag{4}$$

where $\rho_{\text{post/pre}}(t) = \sum_l \delta(t - t_l^{\text{post/pre}})$ is the spike train of the post/pre neuron written as the sum of the delta function at the neuron's spike times $\{t_l^{\text{post/pre}}\}_l$. In the limit of slow learning, $\lambda \to 0$, the right hand side of Eq. (4) can be replaced by its temporal mean (see [31] for complete derivations). This approximation, has been termed the

mean field Fokker-Planck approximation [27]. As fluctuations vanish in this limit and deterministic dynamics are retained for the mean synaptic weights, we get

$$\dot{J}_{ij}(t) = \lambda \int_{-\infty}^{\infty} \Gamma_{ij}(-t') \left[K_+(t') - \alpha K_-(t')\right] dt', \quad (5)$$

where $\Gamma_{ij}(t)$ is the cross correlation of neurons i and j:

$$\Gamma_{ij}(t) = \langle \rho_i(t')\rho_j(t'+t)\rangle. \quad (6)$$

The angular brackets $\langle \cdots \rangle$ denote ensemble averaging over the neurnal noise and temporal averaging over one period in the case of rhythmic activity (see Soloduchin [23] & Shamir and Luz & Shamir [31] for more details). Note that the dependence of the r.h.s. of Eq. (5) on time, t, occurs through the dependence of the cross-correlations on the synaptic weights at time t. Thus, the key to analyzing STDP dynamics is the ability to compute the cross-correlations of the neural activities and grasp their dependence on the synaptic weights. To this end we examined rhythmogenesis in the gamma band using the framework of a reduced rate model with delay, proposed by Roxin and colleagues (Fig. 1a). A complete analysis of the model appears in [21, 22, 24]. Below we briefly describe the phase diagram of the system and derive the cross-correlations.

III. THE DELAYED EXCITATORY INHIBITORY NETWORK

The firing of different neurons is assumed to follow independent inhomogeneous Poisson process statistics with instantaneous firing rates that adhere to the reduced model in Roxin et al. [21]. In their work, Roxin and colleagues considered a full model that included inter-population as well as intra-population interactions (i.e., excitatory-excitatory and inhibitory-inhibitory). Here, for simplicity we restrict the analysis to the minimal model that can reproduce oscillations in the gamma band. To do so, we model the rate dynamics of the gamma generating network by:

$$\tau_m \dot{m}_E(t) = -m_E(t) + [I - J_I m_I(t-d)]_+ \quad (7)$$
$$\tau_m \dot{m}_I(t) = -m_I(t) + [I + J_E m_E(t-d)]_+, \quad (8)$$

where $m_{E/I}(t)$ is the mean firing rate of the excitatory (E) and inhibitory (I) population at time t. τ_m is the neuronal time constant. Unless stated otherwise, we take $\tau_m = 1$, which is equivalent to measuring time in units of the neuronal time constant. Parameter d denotes the delay, and I is the external input to the system. In our analysis we took $I = 1$. J_E and J_I are the effective interaction strengths between the two populations. J_E (J_I) can be thought of as a global order parameter reflecting the mean synaptic weight from the excitatory (inhibitory) pre-synaptic population to the inhibitory (excitatory) post-synaptic population.

For strong inhibition, $J_I > 1$, the system converges to a fixed point in which the excitatory population is fully suppressed by the inhibitory population, $\vec{m}^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. For weak to moderate levels of inhibition, $J_I \in (0,1)$, the system has a fixed point in which both populations are active, $\vec{m}^* \equiv \begin{pmatrix} m_E^* \\ m_I^* \end{pmatrix} = \frac{1}{1+\bar{J}^2} \begin{pmatrix} 1 - J_I \\ 1 + J_E \end{pmatrix}$, with $\bar{J} \equiv \sqrt{J_E J_I}$. However, this fixed point is not stable for $\bar{J} > \bar{J}_d$, where $\bar{J}_d^2 = 1 + \omega_d^2$, $\omega_d = \cot(\omega_d d)$ and $\omega_d \in [0, \pi/2d]$ (see Roxin et al. [21]). In this region ($\bar{J} > \bar{J}_d$ and $J_I < 1$) the system converges to a limit cycle solution, Fig. 1d.

By rescaling the firing rates, the two dimensional first order delayed dynamics, Eq. (7)-(8), can be reduced to a one dimensional delayed dynamic equation:

$$\ddot{x}(t) + 2\dot{x}(t) + x(t) - [1 - \bar{J}^2 x(t-2d)]_+ = 0, \quad (9)$$

with

$$m_I(t) = 1 + \left(J_E - \bar{J}^2\right) x(t) \quad (10)$$

$$m_E(t-d) = \frac{J_E - \bar{J}^2}{J_E} \left(x(t) + \dot{x}(t)\right). \quad (11)$$

Equation (9) highlights the fact that the temporal structure of the limit cycle solution depends on the synaptic weights, J_E and J_I, only via $\bar{J}$. In particular, the period of oscillations is solely a function of $\bar{J}$ and d. As shown in Fig. 2a the period is a monotonically increasing function of both $\bar{J}$ and d.

In our model the cross-correlations are given by temporal averaging of the mean firing rates, $\Gamma_{IE}(\Delta) \equiv \langle m_I(t)\, m_E(t+\Delta)\rangle$. In the fixed point region of the phase diagram, the cross-correlation is simply given by the product of the mean rates, $\Gamma_{IE}(\Delta) = (1 - J_I)(1 + J_E) / \left(1 + \bar{J}^2\right)^2$.

In the region of the phase diagram where the system converges to a limit cycle solution, Eqs. (10) & (11) provide the scaling of the cross-correlations; namely,

$$\Gamma_{IE}(\Delta) = a\bar{x} + b\left(\Gamma_x(\Delta + d) + \frac{d}{d\Delta}\Gamma_x(\Delta + d)\right) \quad (12)$$

where $a \equiv \left(J_E - \bar{J}^2\right)/J_E$, $b \equiv \left(J_E - \bar{J}^2\right)^2/J_E$, $\bar{x} \equiv \langle x(t)\rangle$, and $\Gamma_x(s) \equiv \langle x(t)\, x(t+s)\rangle$. Note that $\bar{x}$ and Γ_x depend solely on the delay, d, and $\bar{J}$. Numerical investigation reveals that the auto-correlation of $x(t)$ is well approximated by a cosine function,

$$\Gamma_x(s) \approx \bar{\Gamma}_x + \tilde{\Gamma}_x \cos(\omega s) \quad (13)$$

where we used

$$\bar{\Gamma}_x = \int_0^T \Gamma_x(s)\, ds/T \quad (14)$$

$$\tilde{\Gamma}_x = 2\int_0^T \Gamma_x(s) \cos(2\pi s/T) ds/T, \quad (15)$$

with T denoting the period of the limit cycle, as can be seen from the value of R^2, Fig. 2b. The goodness of fit

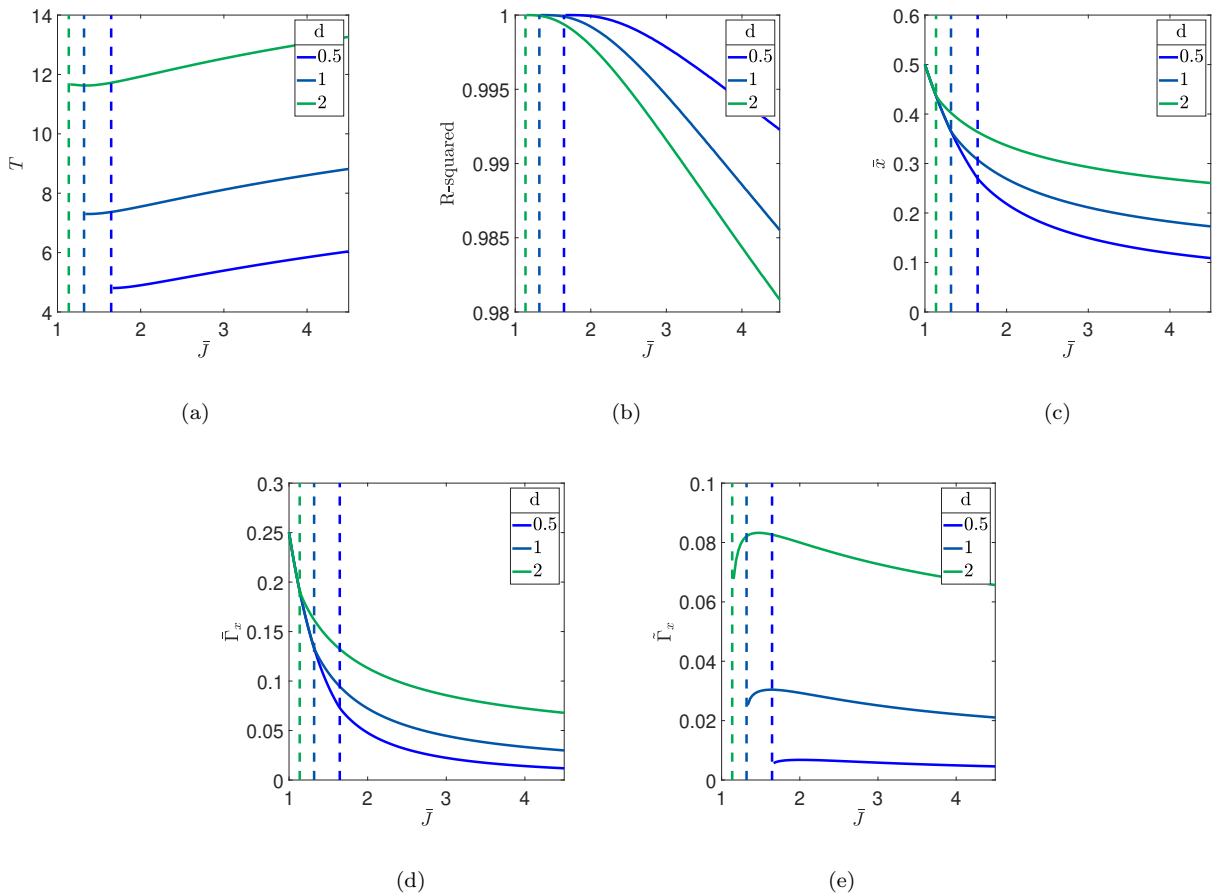

FIG. 2: Dynamics of the one dimensional variable x, Eq. (9). Different features characterizng the dynamics of x are shown as a function of $\bar{J}$, for different values of the delay, d, differentiated by color. The dashed lines indicate the values of $\bar{J}_d$, for different delays, shown by color. The parameters plotted are: (**a**) The period of the oscillations. (**b**) The goodness of fit of the cosine approximation, Eq. (13), of Γ_x. R^2. (**c**) The temporal average of x, $\bar{x}$. (**d**) The zeroth order Fourier component of the auto-correlation of x, $\bar{\Gamma}_x$. (**e**) The first order Fourier component of the auto-correlation of x, $\tilde{\Gamma}_x$.

.

of the cosine approximation decreases when $\bar{J}$ or d are increased. Nevertheless, for a wide range of parameters relevant to the generation of gamma oscillations R^2 is extremely high ($R^2 > 0.98$ throughout Fig. 2b). Both $\bar{x}$ and $\bar{\Gamma}_x$ monotonically decrease as $\bar{J}$ increases, Fig. 2c-2d. In addition, they transition the bifurcation line continuously. On the other hand, $\tilde{\Gamma}_x$ does not transition in a continuous manner: it is zero in the fixed point region and jumps to a positive value in the rhythmic region, Fig. 2e.

Using the cosine approximation for the correlations, Eq. (13) and the scaling Eqs. (10) & (11) yields the semi-empirical excitatory-inhibitory cross-correlations function:

$$\Gamma_{IE}(\Delta) \approx \bar{\Gamma} + \tilde{\Gamma}\cos(\omega\Delta + \tilde{\varphi}) \tag{16}$$

with $\bar{\Gamma} = a\bar{x} + b\bar{\Gamma}_x$, $\tilde{\Gamma} = b\tilde{\Gamma}_x\left(1+\omega^2\right)^{1/2}$, $\tilde{\varphi} = \omega\left(d + \varphi_\omega\right)$ and $\omega\varphi_\omega = \arcsin\left(\omega/\sqrt{1+\omega^2}\right)$. Figure 3 shows the values of $\tilde{\varphi}$ on the phase diagram. The phase, $\tilde{\varphi}(\bar{J}, d)$ is $\pi/2$ on the bifurcation line and weakly decreases as $\bar{J}$ is further increased. Note that $\Gamma_{EI}(\Delta) = \Gamma_{IE}(-\Delta)$.

IV. STDP INDUCED FLOW ON THE PHASE DIAGRAM

Utilizing the semi-empirical cross-correlations, Eq. (16), yields the following dynamics for the synaptic weights:

$$\dot{J}_\sigma = \lambda\left(\bar{\Gamma}\bar{K} + \tilde{\Gamma}\tilde{K}_\sigma\right) \tag{17}$$

where

$$\bar{K} = \int_{-\infty}^{\infty} K(\Delta)\, d\Delta, \tag{18}$$

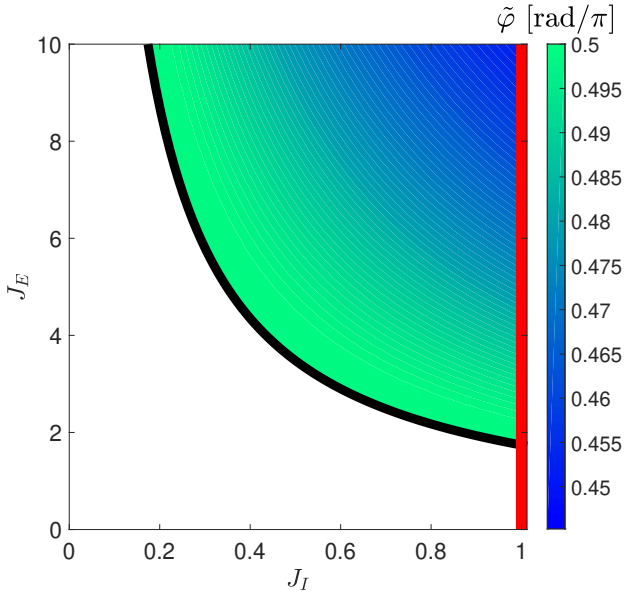

FIG. 3: The values of $\tilde{\varphi}$. The parameter, $\tilde{\varphi}/\pi$ is shown by color as a function of J_E and J_I on the phase diagram. Here $d = 1$ was used.

$$\tilde{K}_\sigma \equiv \int_{-\infty}^{\infty} K(\Delta) \cos(\omega\Delta_\sigma + \tilde{\varphi}) d\Delta, \tag{19}$$

where $\sigma = E, I$, we used the notation $K(\Delta) = K_+(\Delta) - \alpha K_-(\Delta)$, and $\Delta_E = -\Delta$ whereas $\Delta_I = \Delta$.

The STDP dynamics, Eq. (17), induces a flow in the phase plane of the synaptic weights $[J_I, J_E]$, which is also the phase diagram of the neural responses. The right hand side of Eq. (17) depends on the synaptic weights through the Fourier transforms of the cross-correlations $\bar{\Gamma}$ and $\tilde{\Gamma}$. Using the separation of timescales between the fast neuronal responses and the slow learning rate, in the limit of slow learning $\lambda \to 0$, one can compute $\bar{\Gamma}$ and $\tilde{\Gamma}$ from the neuronal dynamics for fixed synaptic weights.

Thus, STDP induces a flow on the phase diagram of the system. Rhythmogenesis is obtained when this flow guides the system and stabilizes it at a fixed point on the phase diagram that is characterized by the desired rhythm.

In the region of the phase diagram in which the mean neuronal firings relax to a fixed point, $\tilde{\Gamma} = 0$. Due to our choice of normalization, $\bar{K}_\pm = 1$, in this region $\text{sign}(\dot{J}_\sigma) = \text{sign}(1 - \alpha_\sigma)$, $\sigma \in \{E, I\}$. Consequently, the STDP dynamics will induce a flow from the fixed point region towards the rhythmic region if and only if the potentiation is strong relative to the depression for both types of synapses, $\alpha_E, \alpha_I < 1$ (except for a small region of the phase diagram with high inhibition and low excitation, which also depends on the learning rates, λ_E and λ_I, of the different synapses). This result holds true for any STDP rule. In contrast, the STDP dynamics in the rhythmic region of the phase diagram depend on the temporal structure of the learning rule.

The difference of Gaussians learning rule, Eq. (2), yields

$$\tilde{K}_\sigma = \cos\tilde{\varphi} \left(e^{-\frac{(\omega\tau_{\sigma,+})^2}{2}} - \alpha e^{-\frac{(\omega\tau_{\sigma,-})^2}{2}} \right), \tag{20}$$

with $\sigma \in \{E, I\}$. Consequently, the dynamical equations of J_E and J_I will be identical if the characteristic time scales of potentiation and depression are the same; namely, if $\tau_{E,+} = \tau_{I,+}$ and $\tau_{E,-} = \tau_{I,-}$. Note that on the r.h.s. of Eq. (20), the term $\cos\tilde{\varphi}$ ensures that $\tilde{K}_\sigma$ is zero on the bifurcation (see Fig. 3). Figures 4a and 4b depict the nullclines of J_E and J_I , respectively, for different values of the relative strength of depression, α (in 4a), and the characterstic time of depression, τ_- (in 4b), differentiated by color. We show that for $\alpha < 1$ ($\alpha > 1$) and $\tau_+ > \tau_-$ ($\tau_+ < \tau_-$) the nullcline of J_E and the left branch of the nullcline of J_I are stable (unstable). A fixed point of the STDP dynamics is obtained by the intersection of J_E and J_I nullclines. For the difference of Gaussians rule, a stable fixed point that exhibits rhythmic activity in the gamma band can thus be obtained. However, this requires a delicate adjustment of the parameters characterizing the STDP learning rules.

For the temporally asymmetric Hebbian exponential rule, Eq. (3) with $H = 1$, we obtain

$$\begin{aligned}\tilde{K}_\sigma = &\cos(\theta_{\sigma,+}) \cos(\theta_{\sigma,+} - \tilde{\varphi}_\sigma) \\ &- \alpha \cos(\theta_{\sigma,-}) \cos(\theta_{\sigma,-} + \tilde{\varphi}_\sigma)\end{aligned} \tag{21}$$

where $\cos(\theta_\pm) = (1+(\omega\tau_\pm)^2)^{-1/2}$, $\tilde{\varphi}_E = \tilde{\varphi}$ and $\tilde{\varphi}_I = -\tilde{\varphi}$. Now, due to $\theta_{\sigma,\pm}$, $\tilde{K}_\sigma$ transitions discontinuously across the bifurcation line, thus inducing discontinuity in $\dot{J}_E$ and $\dot{J}_I$ along the transition from fixed point to the rhythmic region. Figures 4c and 4d depict the nullclines of J_I, for the temporally asymmetric Hebbian learning rule for different values of the relative strength of depression, α (in 4c), and the characteristic time of depression, τ_- (in 4d), differentiated by color. The left branch of the nullclines of J_I is stable (unstable) for $\alpha < 1$ ($\alpha > 1$) and $\tau_+ < \tau_-$ ($\tau_+ > \tau_-$). Interestingly, a considerable part of the J_I nullcline is on the bifurcation line.

For the temporally asymmetric Hebbian learning rule, the dynamics of J_E do not have a nullcline. As a result, a fixed point does not exist, and the temporally asymmetric Hebbian STDP rule cannot stabilize rhythmic activity in the gamma range.

However, as $\Gamma_{EI}(\Delta) = \Gamma_{IE}(-\Delta)$ the temporally asymmetric *anti*-Hebbian exponential rule (Eq. (3) with $H = -1$) for excitatory synapses, J_E, in our model, defines the exact same dynamics as that of an inhibitory synapses, J_I, with a Hebbian rule (Eq. (3) with $H = +1$). Therefore, an asymmetric Hebbian learning rule for J_I and an asymmetric anti-Hebbian learning rule for J_E yield the same nullclines (see Fig. 4c-4d). Moreover, because a considerable part of the nullcline is on the bifurcation line, no fine tuning of the parameters is required to obtain a fixed point of the STDP dynamics that will generate rhythmic activity at ω_d. Thus, the STDP dy-

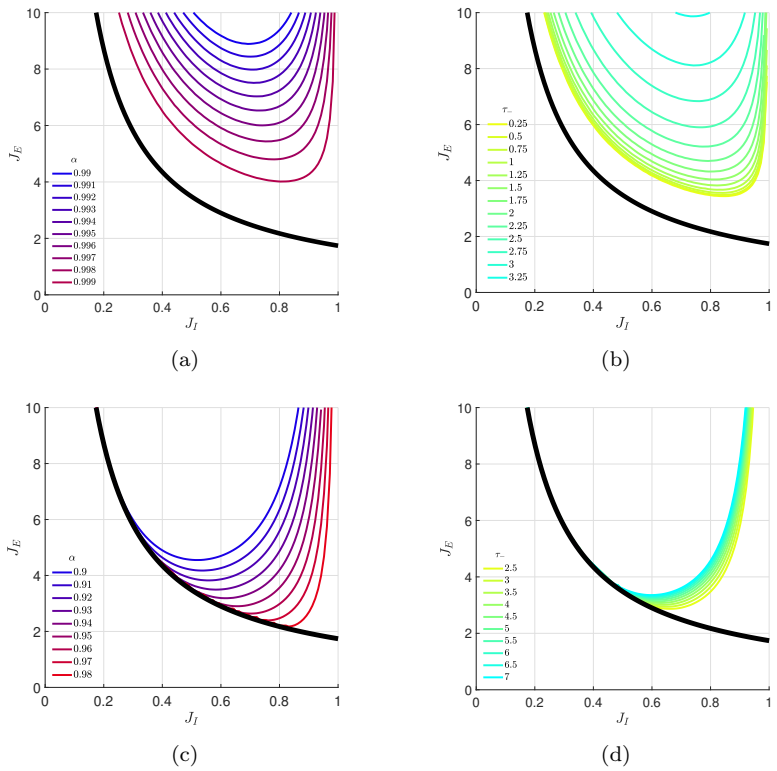

FIG. 4: Nullclines of the STDP dynamics. (**a**) and (**b**) The nullclines of J_E and J_I for the difference of Gaussians learning rule (in this case the nullclines are identical for the same choice of parameters α, $\tau_\pm$) The nullclines are shown for different values of of $\alpha = 0.99, 0.991, ...0.999$, differentiated by color, with $\tau_+ = 2$ and $\tau_- = 1$, in (a). The nullclines are shown for different values of $\tau_- = 0.25, 0.5, ...3.25$, by colors, with $\tau_+ = 5$ and $\alpha = 0.999$, in (b). (**c**) and (**d**) The nullclines of J_I (J_E) for the temporally asymmetric Hebbian (anti-Hebbian) STDP rule. The nullclines are shown for different values of $\alpha = 0.9, 0.91, ...0.99$, with $\tau_+ = 2$ and $\tau_- = 5$, in (c). The nullclines are shown for different values of $\tau_- = 2.5, 3, ...7$, with $\tau_+ = 2$ and $\alpha = 0.94$, in (d). The nullclines were computed using the cosine approximation for the neuronal cross-correlations.

.

namics have a line attractor on (part of) the bifurcation line.

V. DISCUSSION

Previous studies have investigated the effects of rhythmic activity on STDP [20, 41–50]. However, in these studies, rhythmic activity was hard-wired in the system and the issue of rhythmogenesis was not addressed.

Rhythmogenesis can be thought of as self organizing temporal activity; i.e., the ability of a non-rhythmic system to spontaneously develop rhythmic activity. In our approach, the process of rhythmogenesis was mapped to a flow on the phase diagram. This mapping relies on the separation of timescales.

Previously, Soloduchin & Shamir investigated rhythmogenesis using the framework of two neuronal populations with reciprocal inhibition and short term adaptation in the form of firing rate adaptation [23, 25]. The network motif of reciprocal inhibition has been widely reported in the central nervous system [51–53]. However, it is mainly associated with winner-take-all like competition [54–60] rather than generating rhythmic activity (but see [61, 62] in the spinal cord). Here, rhythmogenesis was studied in the framework of a network that is considered a valid hypothesis for generating gamma rhythm in the brain [21, 22, 24].

In Soloduchin & Shamir [23], rhythmogenesis was obtained as a specific stable fixed point on the phase diagram of the system, in which due to the temporal characteristics of the STDP rule, the dynamics of the synaptic weights vanish at a specific frequency. This scenario is similar to the case of temporally symmetric STDP (Fig. 4a-4b). However, scientifically, this scenario is somewhat disappointing, since we have traded the problem of fine-tuning of the synaptic weights for the problem of fine-tuning of characteristics of the STDP rule [25].

The temporally asymmetric STDP rule provides a possible solution to the fine-tuning problem of rhythmogenesis, which we term *critical rhythmogenesis* (Fig. 4c-4d). Rhythmogenesis in the temporally asymmetric STDP rule is not obtained as a fixed point of the STDP dynamics. Rather, in this case, rhythmogenesis utilizes the

discontinuity of $\tilde{\Gamma}$ across the bifurcation line. For a wide range of parameters the flow induced by the STDP is directed from the fixed point region towards the rhythmic region and from the rhythmic region to the fixed point region, and the system will settle on the bifurcation line itself. Consequently, the resultant rhythmic activity will be dictated by bifurcation (e.g., the firing rates will oscillate at ω_d), which is independent of the synaptic plasticity thus accounting for the robustness to the parameters that characterize the STDP.

Recently, Pernelle and colleagues studied the possible contribution of gap junction plasticity to rhythmic activity [63]. They postulated a plasticity rule for gap junctions that tuned the system to operate on the boundary of asynchronous regular firing activity, on one hand, and rhythmic activity of synchronous bursts, on the other. In the context of our work this can be viewed as an example of critical rhythmogenesis, which explains the robustness of their putative mechanism to variations in the plasticity rule.

We suggest that the scenario of critical rhythmogenesis may provide a general principle for robustness in biological systems. Assume a certain biological system, which is characterized by set parameters: $\{x_1, x_2 \dots x_n\}$ (i.e., $\vec{x}$ is a point in the phase diagram of the system), is required to maintain a certain living condition, $f(\{x_1, x_2 \dots x_n\}) = 0$. This condition is met by a homeostatic process. The homeostatic process defines the dynamics on $\{x_1, x_2 \dots x_n\}$, which are characterized by another set of parameters, $\{\alpha_1, \alpha_2 \dots \alpha_m\}$, namely: $\dot{\vec{x}} = \mathcal{F}(\vec{x}, \vec{\alpha})$. A viable homeostatic process is a choice of parameters, $\vec{\alpha}^*$, such that the homeostatic dynamics will lead the system to a set of parameters, x_∞, that satisfy the living condition, $f(\vec{x}_\infty) = 0$.

How can this be achieved? One possibility is that x_∞ is a stable fixed point of the homeostatic dynamics, in which $\mathcal{F}(\vec{x}_\infty, \vec{\alpha}^*) = 0$. This solution requires fine-tuning of the parameters that define the homeostatic process, $\vec{\alpha}$. In this case, fluctuations in $\vec{\alpha}$ will generate fluctuations in $\vec{x}$ away from x_∞.

An alternative scenario is that the homeostatic dynamics utilize some discontinuity in the phase diagram. In this scenario the dynamics do not necessarily vanish on x_∞, $\mathcal{F}(\vec{x}_\infty, \vec{\alpha}^*) \neq 0$. Rather, due to the discontinuity there exists a wide range of parameters, $\{\alpha_1, \alpha_2 \dots \alpha_m\}$, such that the dynamics draw the system towards the discontinuity from both sides, as is the case for critical rhythmogenesis. Consequently, this scenario can stabilize the system in a critical condition on the boundary of two phases.

The idea that the central nervous system may operate in (or near) a critical condition has been suggested in the past, and may have computational advantages [64–66]. However, this latter scenario also has shortcomings. The most obvious is that it can only be used to ensure and stabilize critical behavior. In addition, it cannot be used when one of the phases near the critical line is lethal.

An advantage of critical rhythmogenesis is that it allows for rapid switching between rhythmic and non-rhythmic phases, for example by neuromodulators, since the system is on the boundaries of these phases. This raises the question of the likelihood of critical rhythmogenesis: what is the probability that a biological system will 'choose' the exact rhythmic activity that also characterizes the bifurcation? One possible explanation is that the opposite took place. In other words, biological systems have evolved to operate at the critical conditions chosen by the critical rhythmogenesis mechanism. Thus, in our example, the characteristic delays, d, do not miraculously fit the desired rhythm. Rather, due to the specific values of d, the critical rhythmogenesis tunes the system to oscillate at ω_d which is why the biological system 'uses' this specific frequency band.

In our work we made several simplifying assumption to facilitate the analysis. We studied the dynamics of the effective couplings between excitatory and inhibitory populations and did not incorporate the STDP dynamics of individual synapses. We estimated neuronal correlations using a simplified rate model, and did not study the effects of spiking neurons. These issues are beyond the scope of the current study and will be addressed elsewhere. Nevertheless, this work lays the foundation for studying a novel mechanism for robust homeostatic plasticity, in general, and rhythmogenesis in particular.

ACKNOWLEDGMENTS

This work was supported by the Israel Science Foundation, grant number 300/16.

[1] R. Jung and W. Berger, Archiv fur Psychiatrie und Nervenkrankheiten **227**, 279 (1979).

[2] G. Buzsaki, *Rhythms of the Brain* (Oxford University Press, 2006).

[3] G. Buzsáki and W. Freeman, Current opinion in neurobiology **31**, v (2015).

[4] M. Bocchio, S. Nabavi, and M. Capogna, Neuron **94**, 731 (2017).

[5] M. Shamir, O. Ghitza, S. Epstein, and N. Kopell, PLoS Comput Biol **5**, e1000370 (2009).

[6] A. H. Taub, R. Perets, E. Kahana, and R. Paz, Neuron **97**, 291 (2018).

[7] C. M. Gray, P. König, A. K. Engel, and W. Singer, Nature **338**, 334 (1989).

[8] P. Fries, J. H. Reynolds, A. E. Rorie, and R. Desimone, Science **291**, 1560 (2001).

[9] N. P. Bichot, A. F. Rossi, and R. Desimone, Science **308**, 529 (2005).

[10] B. Pesaran, J. S. Pezaris, M. Sahani, P. P. Mitra, and R. A. Andersen, Nature neuroscience **5**, 805 (2002).

[11] R. T. Canolty, E. Edwards, S. S. Dalal, M. Soltani, S. S. Nagarajan, H. E. Kirsch, M. S. Berger, N. M. Barbaro, and R. T. Knight, science **313**, 1626 (2006).
[12] A. J. Mably and L. L. Colgin, Current opinion in neurobiology **52**, 182 (2018).
[13] W. Cao, S. Lin, Q.-q. Xia, Y.-l. Du, Q. Yang, M.-y. Zhang, Y.-q. Lu, J. Xu, S.-m. Duan, J. Xia, *et al.*, Neuron **97**, 1253 (2018).
[14] S. Ghosh, T. R. Laxmi, and S. Chattarji, Journal of Neuroscience **33**, 7234 (2013).
[15] P. J. Uhlhaas and W. Singer, neuron **52**, 155 (2006).
[16] P. Fries, D. Nikolić, and W. Singer, Trends in neurosciences **30**, 309 (2007).
[17] A. Amir, D. B. Headley, S.-C. Lee, D. Haufler, and D. Pare, Neuron **97**, 656 (2018).
[18] C. G. Welle and D. Contreras, Neuron **93**, 247 (2017).
[19] T. Womelsdorf, P. Fries, P. P. Mitra, and R. Desimone, Nature **439**, 733 (2006).
[20] Y. Luz and M. Shamir, PLoS computational biology **12**, e1004878 (2016).
[21] A. Roxin, N. Brunel, and D. Hansel, Progress of Theoretical Physics Supplement **161**, 68 (2006).
[22] D. Battaglia, N. Brunel, and D. Hansel, Physical review letters **99**, 238106 (2007).
[23] S. Soloduchin and M. Shamir, Scientific reports **8**, 13050 (2018).
[24] D. Battaglia and D. Hansel, PLoS Comput Biol **7**, e1002176 (2011).
[25] M. Shamir, Current opinion in neurobiology **58**, 70 (2019).
[26] H. Câteau and T. Fukai, Neural Computation **15**, 597 (2003).
[27] R. Gütig, R. Aharonov, S. Rotter, and H. Sompolinsky, Journal of Neuroscience **23**, 3697 (2003).
[28] A. Morrison, M. Diesmann, and W. Gerstner, Biological cybernetics **98**, 459 (2008).
[29] J. Rubin, D. D. Lee, and H. Sompolinsky, Physical review letters **86**, 364 (2001).
[30] S. Song and L. F. Abbott, Neuron **32**, 339 (2001).
[31] Y. Luz and M. Shamir, PloS one **9**, e101109 (2014).
[32] G.-q. Bi and M.-m. Poo, Journal of neuroscience **18**, 10464 (1998).
[33] M. A. Woodin, K. Ganguly, and M.-m. Poo, Neuron **39**, 807 (2003).
[34] A. Suvrathan, H. L. Payne, and J. L. Raymond, Neuron **92**, 959 (2016).
[35] J. S. Haas, T. Nowotny, and H. D. Abarbanel, Journal of neurophysiology (2006).
[36] C. C. Bell, V. Z. Han, Y. Sugawara, and K. Grant, Nature **387**, 278 (1997).
[37] M. Nishiyama, K. Hong, K. Mikoshiba, M.-M. Poo, and K. Kato, Nature **408**, 584 (2000).
[38] R. C. Froemke, M.-m. Poo, and Y. Dan, Nature **434**, 221 (2005).
[39] T. Tzounopoulos, M. E. Rubio, J. E. Keen, and L. O. Trussell, Neuron **54**, 291 (2007).
[40] D. O. Hebb, *The organization of behavior* (na, 1961).
[41] N. Sherf and S. Maoz, PLOS Computational Biology **16**, e1008000 (2020).
[42] H. Câteau, K. Kitano, and T. Fukai, Physical Review E **77**, 051909 (2008).
[43] W. Gerstner, R. Kempter, J. L. Van Hemmen, and H. Wagner, Nature **383**, 76 (1996).
[44] M. Gilson, M. Bürck, A. N. Burkitt, and J. L. van Hemmen, Neural computation **24**, 2251 (2012).
[45] J. Karbowski and G. B. Ermentrout, Physical Review E **65**, 031902 (2002).
[46] R. R. Kerr, A. N. Burkitt, D. A. Thomas, M. Gilson, and D. B. Grayden, PLoS Comput Biol **9**, e1002897 (2013).
[47] S. Lee, K. Sen, and N. Kopell, PLoS Comput Biol **5**, e1000602 (2009).
[48] T. Masquelier, E. Hugues, G. Deco, and S. J. Thorpe, Journal of Neuroscience **29**, 13484 (2009).
[49] L. E. Muller, R. Brette, and B. Gutkin, Frontiers in Computational Neuroscience **5**, 45 (2011).
[50] J.-P. Pfister and P. Tass, Frontiers in computational neuroscience **4**, 22 (2010).
[51] E. Aksay, I. Olasagasti, B. D. Mensh, R. Baker, M. S. Goldman, and D. W. Tank, Nature neuroscience **10**, 494 (2007).
[52] J. Kim, M. Pignatelli, S. Xu, S. Itohara, and S. Tonegawa, Nature neuroscience **19**, 1636 (2016).
[53] M. Koyama and A. Pujala, Current Opinion in Neurobiology **49**, 69 (2018).
[54] D. Z. Jin and H. S. Seung, Physical Review E **65**, 051922 (2002).
[55] T. Fukai and S. Tanaka, Neural computation **9**, 77 (1997).
[56] M. Shamir, Neural computation **18**, 2719 (2006).
[57] J. Hertz, A. Krogh, R. G. Palmer, and H. Horner, PhT **44**, 70 (1991).
[58] M. Bazhenov, R. Huerta, and B. H. Smith, Journal of Neuroscience **33**, 5686 (2013).
[59] C. K. Machens, R. Romo, and C. D. Brody, Science **307**, 1121 (2005).
[60] O. Zohar and M. Shamir, Frontiers in computational neuroscience **10**, 107 (2016).
[61] S. M. Danner, N. A. Shevtsova, A. Frigon, and I. A. Rybak, Elife **6**, e31050 (2017).
[62] J. Ausborn, A. C. Snyder, N. A. Shevtsova, I. A. Rybak, and J. E. Rubin, Journal of neurophysiology **119**, 96 (2018).
[63] G. Pernelle, W. Nicola, and C. Clopath, PLoS computational biology **14**, e1006025 (2018).
[64] O. Kinouchi and M. Copelli, Nature physics **2**, 348 (2006).
[65] W. L. Shew, H. Yang, S. Yu, R. Roy, and D. Plenz, Journal of neuroscience **31**, 55 (2011).
[66] O. Shriki and D. Yellin, PLoS computational biology **12**, e1004698 (2016).
[67] G.-q. Bi and M.-m. Poo, Annual review of neuroscience **24**, 139 (2001).
[68] N. Caporale and Y. Dan, Annu. Rev. Neurosci. **31**, 25 (2008).
[69] T. P. Vogels, R. C. Froemke, N. Doyon, M. Gilson, J. S. Haas, R. Liu, A. Maffei, P. Miller, C. Wierenga, M. A. Woodin, *et al.*, Frontiers in neural circuits **7**, 119 (2013).
[70] M. Gilson, A. N. Burkitt, D. B. Grayden, D. A. Thomas, and J. L. van Hemmen, Biological Cybernetics **101**, 411 (2009).
[71] W. O. Friesen, Neuroscience & Biobehavioral Reviews **18**, 547 (1994).
[72] J. C. Magee and D. Johnston, Science **275**, 209 (1997).
[73] L. I. Zhang, H. W. Tao, C. E. Holt, W. A. Harris, and M.-m. Poo, Nature **395**, 37 (1998).
[74] Y. Dan and M.-M. Poo, Physiological reviews **86**, 1033 (2006).
[75] M. Zilberter, C. Holmgren, I. Shemer, G. Silberberg, S. Grillner, T. Harkany, and Y. Zilberter, Cerebral Cor-

tex **19**, 2308 (2009).
[76] M. Stopfer, S. Bhagavan, B. H. Smith, and G. Laurent, Nature **390**, 70 (1997).
[77] R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse, M. Munk, and H. Reitboeck, Biological cybernetics **60**, 121 (1988).
[78] A. K. Engel, P. König, A. K. Kreiter, and W. Singer, Science , 1177 (1991).
[79] V. N. Murthy and E. E. Fetz, Proceedings of the National Academy of Sciences of the United States of America **89**, 5670 (1992).
[80] C. Tallon-Baudry, O. Bertrand, C. Delpuech, and J. Pernier, Journal of Neuroscience **16**, 4240 (1996).
[81] J.-M. Schoffelen, R. Oostenveld, and P. Fries, Science **308**, 111 (2005).
[82] Y. Luz and M. Shamir, PLoS computational biology **8**, e1002334 (2012).