ilmedia

Technische Universität Ilmenau

Gast, Richard; Knösche, Thomas R.; Schmidt, Helmut

Mean-field approximations of networks of spiking neurons with short-term synaptic plasticity

Mean-field approximations of networks of spiking neurons with short-term synaptic plasticity

Richard Gast [*], Thomas R. Knösche [†], and Helmut Schmidt

Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Low-dimensional descriptions of spiking neural network dynamics are an effective tool for bridging different scales of organization of brain structure and function. Recent advances in deriving mean-field descriptions for networks of coupled oscillators have sparked the development of a new generation of neural mass models. Of notable interest are mean-field descriptions of all-to-all coupled quadratic integrate-and-fire (QIF) neurons, which have already seen numerous extensions and applications. These extensions include different forms of short-term adaptation considered to play an important role in generating and sustaining dynamic regimes of interest in the brain. It is an open question, however, whether the incorporation of presynaptic forms of synaptic plasticity driven by single neuron activity would still permit the derivation of mean-field equations using the same method. Here we discuss this problem using an established model of short-term synaptic plasticity at the single neuron level, for which we present two different approaches for the derivation of the mean-field equations. We compare these models with a recently proposed mean-field approximation that assumes stochastic spike timings. In general, the latter fails to accurately reproduce the macroscopic activity in networks of deterministic QIF neurons with distributed parameters. We show that the mean-field models we propose provide a more accurate description of the network dynamics, although they are mathematically more involved. Using bifurcation analysis, we find that QIF networks with presynaptic short-term plasticity can express regimes of periodic bursting activity as well as bistable regimes. Together, we provide novel insight into the macroscopic effects of short-term synaptic plasticity in spiking neural networks, as well as two different mean-field descriptions for future investigations of such networks.

I. LOW-DIMENSIONAL MANIFOLDS OF SPIKING NEURAL NETWORK ACTIVITY

The brain can generate a variety of highly complex and chaotic patterns of neural activity [1]. However, given the vast number of neurons in the brain, these patterns appear to be less complex than they could be theoretically, indicating a high level of neuronal redundancy [2,3]. Electrophysiological recordings of macroscopic neural activity have revealed highly stereotyped responses to sensory stimulation as well as strongly synchronized regimes of neural activity [4–7]. More recently, multiunit recordings have demonstrated that strong redundancies are present at the level of spiking neurons as well [8,9]. These findings indicate the existence of low-dimensional manifolds in the state space of the brain that typically govern its neural dynamics and its response to extrinsic stimulation. The identification and description of such low-dimensional manifolds has been a central topic of neuroscientific research for many years [10–15]. Different approaches for the derivation of mathematical descriptions of the temporal evolution of low-dimensional neural activity have been proposed [16]. Among those are classic neural mass models that use direct, phenomenological descriptions of macroscopic measures of neural dynamics [17–21]. For these neural mass models, equivalent spiking neural networks do not exist in general. Other approaches make use of probabilistic descriptions of the evolution of the collective behavior inside a neural population [22–24], which make it possible to capture the statistics inside the spiking neural network up to a certain order. However, some of these approaches are restricted to asynchronous regimes of neural activity [22,23], whereas others use approximations of random fluctuations in the spiking neural network [24]. Hence, neither of these approaches provide a mathematically exact set of mean-field equations that can describe the macroscopic dynamics of a spiking neural network in general.

The Ott-Antonsen ansatz has provided a new tool to derive mean-field models of spiking neural networks [25]. While originally devised for networks of all-to-all coupled Kuramoto oscillators [26], it has since been applied to networks of theta neurons [27,28], and, most relevant to this study, to networks of all-to-all coupled quadratic integrate-and-fire (QIF) neurons [29]. For future applications of this method, it is of interest to know how well the derivation of the mean-field equations generalizes to other descriptions of neural dynamics

─────────
[*]rgast@cbs.mpg.de
[†]Also at Institute for Biomedical Engineering and Informatics, TU Ilmenau, Germany.

than the particular QIF networks considered in Ref. [29]. Consequently, different extensions of the QIF model have been proposed that added biophysical mechanisms or structural details to the model in order to explain interesting neurodynamic phenomena, such as the onset of synchronized neural activity [30–34]. Particularly interesting are extensions that include dynamic variables which are not driven by the mean-field activity of the network, but by neuron- or synapse-specific processes instead. In such cases, it is unclear whether mean-field equations can still be found. In Ref. [34], the QIF network was extended by a spike-frequency adaptation mechanism, where a neuron-specific adaptation current was elicited by the spiking activity of the same neuron. Thus, the adaptation variable was not simply driven by the mean-field activity of the network. To derive the mean-field equations nonetheless, the authors applied an adiabatic approximation to the adaptation dynamics. This approximation assumes that the adaptation variable evolves slowly in comparison to the membrane potential dynamics and permits one to apply the mean-field derivation on the fast timescale. Based on this mean-field model it will be possible to investigate the effects of neuron-specific currents at meso- and macroscopic scales, such as for example the effects of calcium-dependent spikes on thalamic dynamics [35] or the effects of spike-frequency adaptation on cortical microcircuits [36].

In this work, we address the question of whether exact mean-field equations can be derived for QIF networks with synapse-specific dynamic variables. Synaptic dynamics are especially interesting for the computational modeling of macroscopic neurodynamic phenomena. This is because synaptic currents are thought to trigger the potential changes visible in macroscopic electrophysiological recordings of brain activity, and different synapse types come with different dynamic characteristics that are pivotal for our understanding of brain dynamics. Classic neural mass models, for example, typically use different synaptic timescales to model rhythm generation in the brain [18,20,21]. The QIF mean-field reduction generalizes to any convolution of the synaptic input with a synaptic response kernel [29,30] and, hence, allows one to derive mean-field descriptions of QIF networks with standard descriptions of synaptic dynamics such as the alpha kernel convolution [20,21]. However, given appropriate stimulation, synaptic dynamics also undergo short-term plasticity (STP) that changes properties of the synaptic response. It has been shown that synapses can express short-term depression and facilitation and that timescales and strengths of these two STP forms differ between synapse and neuron types. Moreover, synaptic STP has been linked to various functions and dynamic properties of the brain, such as working memory [37] or operating in a critical regime [38]. A generalization of the above discussed mean-field approaches to neural networks with synaptic STP would thus provide a valuable tool for modeling brain dynamics and function at the meso- and macroscopic level.

Here we discuss the descriptions of synaptic STP that are allowed for in the context of deriving Ott-Antonsen manifolds for heterogeneous QIF networks. Recent work has demonstrated that mean-field equations can be derived for QIF networks with synaptic STP if two conditions are satisfied [34]: First, each time a neuron spikes in the network,

Pre-Synaptic Plasticity   **Post-Synaptic Plasticity**

$$\tau\dot{v}_i = v_i^2 + \eta_i + I(t) + \frac{J}{N}\sum_{j=1}^{N} A_j s_j \qquad \tau\dot{v}_i = v_i^2 + \eta_i + I(t) + \frac{JA_i}{N}\sum_{j=1}^{N} s_j$$
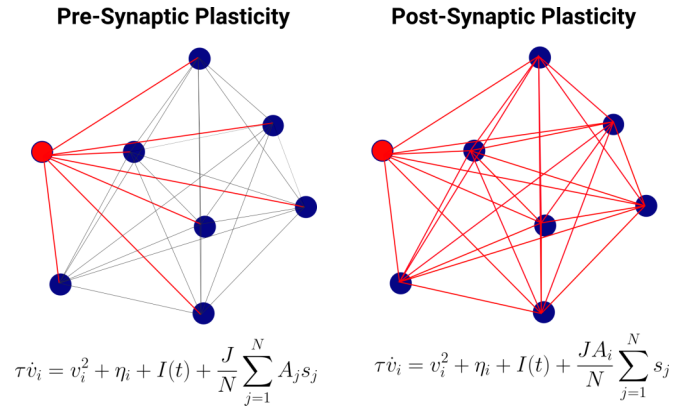
FIG. 1. Pre- vs. postsynaptic forms of short-term plasticity. Nodes represent neurons in an all-to-all coupled network and edges between the nodes represent bidirectional synaptic couplings. Red nodes are active, i.e., did just spike, whereas blue nodes have not spiked for a sufficient period in time. Edges in red show adaptation in response to the activity of the red nodes, whereas gray edges do not. The two equations describe the membrane potential evolution of a QIF neuron for the cases of pre- and postsynaptic plasticity. Note that the adaptation variable A_i is specific for presynaptic source neurons for the former case, and specific to postsynaptic target neurons for the latter.

it triggers synaptic STP at every other neuron, which is the case in all-to-all coupled networks. Second, a single incoming spike triggers synaptic STP at all synapses of a neuron. Under those conditions, synaptic STP is no longer neuron specific and can simply be treated as a macroscopic variable driven by the mean-field activity of the network. This form of synaptic STP could be used to model forms of postsynaptic receptor desensitization, short-term changes in the number of available postsynaptic receptors, or resource depletion at the postsynaptic complex. Importantly, it cannot be considered to represent presynaptic forms of plasticity, such as vesicle depletion. While the first assumption would still hold for presynaptic STP in all-to-all coupled QIF networks, the second assumption would not. Presynaptic resource depletion cannot be assumed to affect all network connections, but only the efferent connections of a specific neuron (see Fig. 1).

A well-established model of presynaptic STP is the phenomenological model introduced in Ref. [39], which describes the dynamics of presynaptic facilitation and depression. We will discuss the derivation of mean-field equations for QIF networks with presynaptic STP with respect to this model, though we will discuss the implications of our findings for general descriptions of presynaptic STP dynamics as well. In the following section, we define the microscopic model under consideration. This will be followed by sections in which we discuss different approaches to derive equations for the low-dimensional network dynamics. While we do not find the exact mean-field equations for QIF networks with presynaptic STP, we provide two different approximations that match well with the QIF network dynamics. We point to the problems that would have to be solved in future attempts at an exact mean-field derivation and evaluate the accuracy of our approximate solutions via numerical simulations and bifurcation analysis.

II. LOW-DIMENSIONAL MANIFOLDS OF QIF NETWORKS WITH STP

We consider a network of N all-to-all coupled QIF neurons with presynaptic STP

$$\tau\dot{V}_i = V_i^2 + \eta_i + I(t) + \frac{J\tau}{N}\sum_{j=1}^{N} X_j^- U_j^+ S_j, \qquad (1a)$$

$$\tau_x\dot{X}_i = 1 - X_i - \alpha X_i^- U_i^+ S_i \tau_x, \qquad (1b)$$

$$\tau_u\dot{U}_i = U_0 - U_i + U_0(1 - U_i^-)S_i\tau_u, \qquad (1c)$$

$$S_i = \sum_{k\backslash t_i^k < t}\int_{-\infty}^{t} a(t - t')\delta(t' - t_i^k)dt', \qquad (1d)$$

where Eq. (1d) represents a convolution of the spiking activity of neuron i with a synaptic response kernel a, e.g., in the case of exponential synapses $a(t) = e^{-t/\tau_s}/\tau_s$ with synaptic timescale τ_s. A neuron i emits its kth spike at time t_i^k when it reaches a threshold V_θ on which V_i is reset to $V_r = -V_i$. Without loss of generality, we consider the limit $\tau_s \to 0$, such that S_i represents the spiking activity of neuron i. Equation (1b) and Eq. (1c) resemble the presynaptic STP mechanism described in Ref. [39]. We note here that $\cdot^-$ denotes a quantity just before a spike occurs (left limit), and $\cdot^+$ denotes a quantity just after the neuron spiked (right limit). This discontinuity accounts for the biological fact that a presynaptic spike triggers synaptic facilitation before it can affect the postsynaptic neuron, by moving vesicles closer to the membrane. Synaptic depression, however, results from the consumption of vesicles for the synaptic transmission process and is thus affected slightly later than synaptic facilitation. We assume neural spiking activity to affect all outgoing synapses of a neuron equally, hence X_i and U_i can be considered as neuron and not synapse specific. The adaptation dynamics are controlled by the depression and facilitation time constants τ_x and τ_u, a depression strength α, and a baseline synaptic efficacy U_0. Equation (1a) describes the evolution of the membrane potential V_i of neuron i, which depends on a background excitability parameter η_i, an extrinsic forcing term $I(t)$, the membrane time constant τ, and the coupling with the network activity. The latter is given by a sum over the output S_i of each neuron in the network, weighted by a global coupling strength J, and the neuron-specific synaptic depression X_i and facilitation U_i.

In the limit $V_\theta \to \infty$, the membrane potential V_i of a QIF neuron can be directly related to its phase via the transform $V_i = \tan\frac{\theta_j}{2}$. Under this transformation, (1a)–(1d) represents a network of theta neurons [40], which can be considered a network of globally coupled oscillators. Thus, the network satisfies the conditions for the existence of the Ott-Antonsen manifold, a low-dimensional manifold along which the network dynamics are guaranteed to evolve for $N \to \infty$ [25,41]. This manifold can be described for (1a)–(1d) by following the Lorentzian ansatz described in Ref. [29], i.e., by making the assumption that the state variables V_i are distributed according to a Lorentzian where the probability density of V for background excitability η at time t is given by

$$\rho(V|\eta, t) = \frac{1}{\pi}\frac{z(\eta, t)}{[V - y(\eta, t)]^2 + z(\eta, t)^2}. \qquad (2)$$

The center $y(\eta, t)$ and half-width-at-half-maximum (HWHM) $z(\eta, t)$ of Eq. (2) are associated with the mean firing rate $r(\eta, t)$ and the membrane potential average over all neurons $v(\eta, t)$ via $z(\eta, t) = \pi r(\eta, t)$, and $y(\eta, t) = v(\eta, t)$, respectively. Due to the conservation of the number of neurons, the network dynamics obey the following continuity equation:

$$\partial_t\rho + \partial_V\left[\left(\frac{V^2 + \eta + I}{\tau} + Jr_{\text{eff}}\right)\rho\right] = 0, \qquad (3)$$

where $r_{\text{eff}} = \frac{1}{N}\sum_{j=1}^{N} X_j^- U_j^+ S_j$ is the effective mean-field network activity that arrives at each neuron. By inserting Eq. (2) into Eq. (3) it can be shown that the dynamics of $z(\eta, t)$ and $y(\eta, t)$ obey

$$\partial_t w(\eta, t) = i\left[\frac{-w(\eta, t)^2 + \eta + I}{\tau} + Jr_{\text{eff}}\right], \qquad (4)$$

for any η, with $w(\eta, t) = z(\eta, t) + iy(\eta, t)$. Without synaptic STP, i.e., for $U(t) = X(t) = 1$, Eq. (4) can be solved for certain choices of the background excitability distribution. The most drastic reduction in the dimensionality of the system can be achieved by choosing a Lorentzian distribution with density function

$$g(\eta) = \frac{1}{\pi}\frac{\Delta}{(\eta - \bar{\eta})^2 + \Delta^2}, \qquad (5)$$

where $\bar{\eta}$ and Δ represent the center and HWHM of the distribution, respectively. This choice allows one to solve

$$\dot{w} = \int_{-\infty}^{\infty}\partial_t w(\eta, t)g(w)dw \qquad (6)$$

using the residue theorem of complex analysis, i.e., by evaluating the integral at the two poles of $g(w)$ given by $\bar{\eta} \pm i\Delta$. Subsequently, Eq. (4) can be solved for r and v, yielding

$$\tau\dot{r} = \frac{\Delta}{\pi\tau} + 2rv, \qquad (7a)$$

$$\tau\dot{v} = v^2 + \bar{\eta} + I(t) + Jr\tau - (\pi r\tau)^2, \qquad (7b)$$

where we additionally used $r_{\text{eff}} = \frac{1}{N}\sum_{j=1}^{N} S_j = r$.

However, for nonconstant X and U, solving Eq. (4) for r and v becomes a nontrivial problem. In this case, $r_{\text{eff}} = \frac{1}{N}\sum_{j=1}^{N} X_j^- U_j^+ S_j \neq r$ and, hence, r_{eff} must be calculated to arrive at closed-form equations for r and v. Two major problems have to be solved in this regard: (a) The effective network input r_{eff} has to be expressed via mean-field variables such as the average firing rate r and average depression and facilitation variables x and u. If this cannot be done, the mean-field equations would still contain neuron-specific variables, thus increasing their dimensionality dramatically. (b) The mean-field equations for the average depression $x = \frac{1}{N}\sum_{i=1}^{N} X_i$ and facilitation $u = \frac{1}{N}\sum_{i=1}^{N} U_i$ have to be solved. However, the evaluation of these sums requires one to solve the coupled, nonlinear differential equations (1b) and (1c), which only has been achieved for stationary network input so far [39]. In the following section, we will address problem (b) and compare our results with recently proposed mean-field equations for a similar synaptic STP model [42]. The remainder of this article will address different attempts to solve problem (a).

III. ANALYTICAL SOLUTIONS FOR MICROSCOPIC STP

As argued in the previous section, finding closed-form mean-field equations for the system given by equations (1) requires one to calculate the average depression $x = \frac{1}{N} \sum_{i=1}^{N} X_i$ and average facilitation $u = \frac{1}{N} \sum_{i=1}^{N} U_i$ across neurons. We start by considering neuron i that spikes periodically with a period T, thus producing a spike train $S_i(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_i)$. The interspike interval T_i corresponds to a firing rate of $1/T_i$. In this scenario, solutions for the microscopic STP variables can be obtained analytically [39]. The evolution equations for synaptic short-term depression X_i and short-term facilitation U_i are given by Eq. (1b) and Eq. (1c), respectively. For the remainder of this section, we will omit the neuron index i for brevity. The (relative) strength of a synapse is given by $0 < U^+ X^- < 1$. We denote U by U_n^- just before the corresponding neuron emitted its nth spike, and by U_n^+ just after the nth spike. Solving the homogeneous part of the model equation, we obtain

$$U_{n+1}^- = U_0 + (U_n^+ - U_0)\exp(-T/\tau_u), \qquad (8)$$

and the change of U due to a spike is found to be

$$U_{n+1}^+ = U_{n+1}^- + U_0(1 - U_{n+1}^-). \qquad (9)$$

These expressions can be reformulated into the following iteration scheme:

$$U_{n+1}^+ = U_0 + (1 - U_0)[U_0 + (U_n^+ - U_0)e^{-T/\tau_u}], \quad (10a)$$

$$U_{n+1}^- = U_0 + (1 - U_0)U_n^- e^{-T/\tau_u}. \qquad (10b)$$

For the depression variable X, we find the following set of equations:

$$X_{n+1}^+ = 1 + [(1 - \alpha U_n^+)X_n^- - 1]e^{-T/\tau_x}, \qquad (11a)$$

$$X_{n+1}^- = (1 - \alpha U_{n+1}^+)[1 + (X_n^+ - 1)e^{-T/\tau_x}]. \qquad (11b)$$

In the stationary case, i.e., in the absence of transient dynamics, stationary solutions $U_\star^+ = U_n^+$, $U_\star^- = U_n^-$ and $X_\star^- = X_n^-$, $\forall n$ can be found:

$$U_\star^+ = \frac{U_0 + U_0(1 - U_0)[1 - \exp(-T/\tau_u)]}{1 - (1 - U_0)\exp(-T/\tau_u)}, \qquad (12a)$$

$$U_\star^- = \frac{U_0}{1 - (1 - U_0)\exp(-T/\tau_u)}, \qquad (12b)$$

$$X_\star^+ = \frac{(1 - \alpha U_\star^+)[1 - \exp(-T/\tau_x)]}{1 - (1 - \alpha U_\star^+)\exp(-T/\tau_x)}, \qquad (12c)$$

$$X_\star^- = \frac{1 - \exp(-T/\tau_x)}{1 - (1 - \alpha U_\star^+)\exp(-T/\tau_x)}. \qquad (12d)$$

It is interesting to note that these results differ from firing rate descriptions, where neural spiking activity is described by a firing rate. To arrive at a rate description, we replaced the spike train with interspike interval T by a firing rate $r_0 = 1/T$. Since neurons are assumed to spike periodically, the firing rate of each neuron is constant. Therefore, neurons receive constant synaptic inputs, and we can set $\dot{U} = \dot{X} = 0$ in Eq. (1) to obtain the simpler expressions

$$U_\star = \frac{U_0 + U_0 \tau_u r_0}{1 + U_0 \tau_u r_0}, \qquad (13a)$$

$$X_\star = \frac{1}{1 + \alpha \tau_x U_\star r_0}, \qquad (13b)$$

where we have made use of $U_\star^+ = U_\star^- = U_\star$, as well as $X_\star^+ = X_\star^- = X_\star$ since spike times are irrelevant in the rate description.

In Fig. 2 we compare these solutions for varying firing rates. As can be seen, the results for constant firing rates r_0 are more closely related to the adaptation variables before spikes than after spikes. This shows that it does matter for microscopic STP whether exact spike timings and the time of evaluation of U and X are considered or not, a finding which we expect to hold for nonstationary firing rates $S(t)$ as well.

The expressions derived above can be used to evaluate the mean-field quantities x and u, if the spike times or firing rates of all neurons are known. Alternatively, they can be used to evaluate r_{eff} directly. In the following sections, we will address the problem of evaluating r_{eff} to derive the mean-field equations for equations (1). We will derive two different mean-field models, for which the results of this section will be used to refine the mean-field descriptions of the presynaptic STP dynamics. In this context, we will evaluate how Eq. (12) vs. Eq. (13) affect the mean-field dynamics of the QIF network.

IV. MEAN-FIELD DERIVATION UNDER A POISSONIAN ASSUMPTION OF NEURAL DYNAMICS

Recently, an approach for the derivation of a mean-field model for a QIF network has been presented in Ref. [37] for the following set of equations:

$$\tau \dot{V}_i = V_i^2 + \eta_i + I(t) + \frac{J\tau}{N} \sum_{j=1}^{N} X_j^- U_j^- S_j, \qquad (14a)$$

$$\tau_x \dot{X}_i = 1 - X_i - \alpha X_i^- U_i^- S_i \tau_x, \qquad (14b)$$

$$\tau_u \dot{U}_i = U_0 - U_i + U_0(1 - U_i^-)S_i \tau_u, \qquad (14c)$$

$$S_i = \sum_{k \setminus t_i^k < t} \int_{-\infty}^{t} a(t - t')\delta(t' - t_i^k)dt'. \qquad (14d)$$

The authors used a mean-field approximation of macroscopic quantities x and u, averaged over all neurons in the network, that has been proposed in Ref. [42]. In this article, a mean-field approximation of the effective network input

$$r_{\text{eff}}(t) = \frac{1}{N} \sum_{j=1}^{N} U_j^- X_j^- s_j, \qquad (15)$$

is derived, where X_i^- and U_i^- are given by Eq. (14b) and Eq. (14c), respectively. Note that the original STP model formulation described in Ref. [39] uses $U_j^+ X_j^-$ as the effective weight of a synapse at the time of an incoming spike, whereas $U_j^- X_j^-$ is used in Ref. [42]. As shown in Fig. 2(c), these two choices can lead to substantial differences of the synaptic weight for small input rates. Since an effective synaptic weight of $U_j^- X_j^-$ is also used in Ref. [37], we will discuss the validity of their mean-field description for both the spiking neural network given by Eq. (1) and the spiking neural network considered in Ref. [37]. Henceforth, we will refer to the former as SNN$_{\text{pre}}$ and to the latter as SNN$_{\text{pre}}$ II. Under the
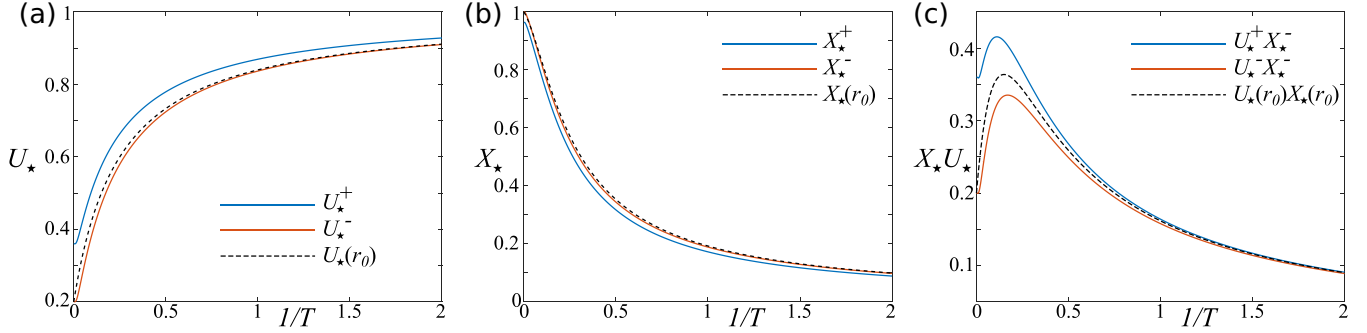
FIG. 2. Comparison of the microscopic adaptation variables before and after spikes for discrete spikes and for constant firing rates r_0. The interspike interval T is varied. The constant firing rate is expressed as $r_0 = 1/T$. Parameters: $\alpha = 0.1$, $U_0 = 0.2$, $\tau_x = 50.0$, and $\tau_u = 20.0$.

assumption that all S_i follow independent Poisson processes, the effective network input in SNN_{pre} II is approximated by $r_{\text{eff}} \approx u(t)x(t)r(t)$, where $r(t)$ is the average firing rate across neurons at time t. As explained in Ref. [42], this mean-field approximation rests on two assumptions: (i) Synapse indices can be randomized, i.e., the spike times matter, but not the synapses at which those spikes occur. (ii) The average impact of a spike on X_i and U_i, $\forall i$ can be approximated by sampling from Gaussian distributions around the current values of x and u. A first-order mean-field approximation is then given by

$$\tau_x \dot{x} = 1 - x - \alpha \tau_x x u r, \tag{16a}$$

$$\tau_u \dot{u} = U_0 - u + U_0 \tau_u (1 - u) r. \tag{16b}$$

As can be seen from these equations, both x and u are driven by the average firing rate $r = \frac{1}{N} \sum_{j=1}^{N} S_j$ of the QIF network. This allows to one to apply the Lorentzian ansatz in the same way as demonstrated for postsynaptic depression in Ref. [34]. The dynamics of the complex variable $w(\eta, t)$ can be expressed as

$$\partial_t w(\eta, t) = i \left[\frac{-w(\eta, t)^2 + \eta + I(t)}{\tau} + Jxur \right] \tag{17}$$

and by evaluating Eq. (17) at $\pi r(t) + iv(t) = w(\bar{\eta} - i\Delta, t)$ one finds that the dynamics of r and v follow:

$$\tau \dot{r} = \frac{\Delta}{\pi \tau} + 2rv, \tag{18a}$$

$$\tau \dot{v} = v^2 + \bar{\eta} + I(t) + Jxur\tau - (\pi r\tau)^2. \tag{18b}$$

We will refer to the set of mean-field equations given by (16) and (18) as $\text{FRE}_{\text{Poisson}}$ where FRE stands for firing rate equations. Short descriptions and equation references for each of the SNNs and mean-field models derived throughout this article can be found in Table I.

It is important to notice that $\text{FRE}_{\text{Poisson}}$ cannot be considered exact. While assumption (I) holds for a network of independent, homogeneous Poisson neurons (hence called Poissonian assumption), it does not hold in general [42]. Therefore, the mean-field derivation essentially approximates a heterogeneous network of deterministic QIF neurons by a homogeneous network of stochastic Poisson neurons. Furthermore, the first-order approximation given by Eq. (16a) and Eq. (16b) ignores the nonlinear interaction between X_i and U_i in Eq. (1b). As shown in Ref. [42], considering second-order

dynamics can improve the accuracy of the mean-field approximation, especially in the vicinity of transient inputs to the network. Adding second-order dynamics would involve sampling from a multivariate Gaussian distribution over (x, u), however. This means that the mean-field derivation could not be considered deterministic and, hence, also not exact anymore.

Still, it has been shown in Ref. [37] that $\text{FRE}_{\text{Poisson}}$ can accurately describe the mean-field dynamics of SNN_{pre} II under certain conditions. To test whether this holds in general, we compared the dynamics of the two models for three different STP parametrizations, leading to synapses that are either depressing, facilitating, or depressing and facilitating. We solved the initial value problem of both sets of equations via an explicit Euler formalism with an integration step-size of $dt = 0.0001$. This step size was sufficiently small to capture the dynamics of the network and was used for all subsequent numerical integration problems as well. We then applied rectangular input pulses to the models and observed their dynamic responses around these inputs. The resulting time series can be observed in Fig. 3. For purely depressing synapses, we find that there is a substantial mismatch between the mean-field dynamics of SNN_{pre} II and $\text{FRE}_{\text{Poisson}}$. As can be seen in Fig. 3(a) for the average depression x, there is a considerable offset between the mean-field model (orange) and the average of X_i evaluated across neurons in the QIF network (black). With respect to purely facilitating synapses, we find that the mean-field model provides a reasonable approximation of the QIF network. Even though offsets can be observed between the mean-field model and the QIF network [see dynamics of v in Fig. 3(b)], the qualitative behavior of the QIF network is captured well by the mean-field model. This holds both in the steady-state regimes and during transient behavior around the on- and offsets of the input $I(t)$. In the case of synapses with short-term depression and facilitation, the mean-field model expresses a substantial mismatch to the QIF network dynamics again. For example, Fig. 3(c) shows that the dynamics of the average firing rate r express focus dynamics for $\text{FRE}_{\text{Poisson}}$ after the onset of the first stimulus, whereas the average firing inside SNN_{pre} II does not show such behavior. In the upper row of Fig. 3, we show the evolution of the distribution over the combined synaptic state $X_i U_i$ in the microscopic model. We find that this distribution tends to express multimodalities in regions with a strong mismatch between mean-field and microscopic model. These results suggest that the mean-field
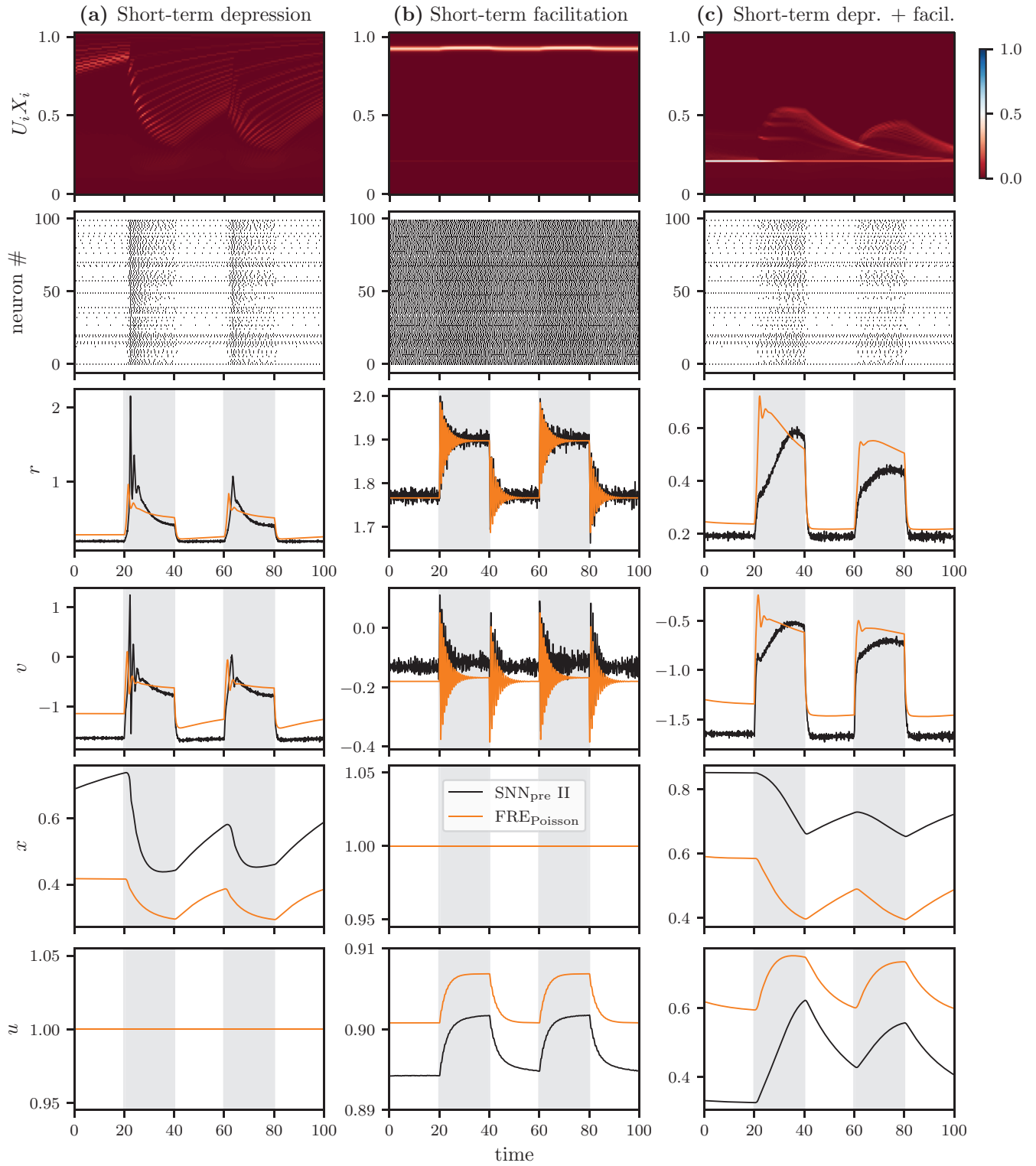
FIG. 3. Evolution of the state variables of a QIF network and a mean-field approximation thereof for three different types of synaptic short-term plasticity [(a) depression, (b) facilitation, and combined (c) depression and facilitation]. The first two rows show the distribution over the synaptic state $X_j U_j$ and the spiking activity of 100 randomly selected neurons, respectively. The last four rows show a comparison between the spiking neural network (black) and the mean-field approximation (orange) for the average firing rate r, the average membrane potential v, the average depression x, and the average facilitation u. In the SNN, averages were calculated across neurons i. Gray-shaded areas depict time intervals in which a rectangular input of $I(t) = 2.0$ was applied to the model [the effects of $I(t)$ and $\bar{\eta}$ on the QIF model are identical]. Color bars depict the probability density inside a given bin of the distribution over $X_i U_i$. Parameters for (a) $U_0 = 1.0$, $\alpha = 0.1$. Parameters for (b) $U_0 = 0.2$, $\alpha = 0.0$. Parameters for (c) $U_0 = 0.2$, $\alpha = 0.1$. Other model parameters: $\tau = 1.0$, $\Delta = 2.0$, $\bar{\eta} = -3.0$, $J = 15.0\sqrt{\Delta}$, $\tau_x = 50.0$, $\tau_u = 20.0$, $N = 10000$.

TABLE I. Model definitions.

Abbreviations	Descriptions	Equations
SNN$_{pre}$	Spiking neural network model with synaptic STP given by Tsodyks model	(1)
SNN$_{pre}$ II	Spiking neural network model with synaptic STP given by a simplified Tsodyks model	(14)
FRE$_{Poisson}$	Firing rate equations derived via the Poissonian assumption	(18)
FRE$_{MPA}$	Firing rate equations representing multiple coupled FRE$_{Poisson}$ populations	(21)
FRE$_{aa1}$	Firing rate equations derived via the adiabatic assumption, neglecting spike timings in the STP mean-field approximation	(26), (36)
FRE$_{aa2}$	Firing rate equations derived via the adiabatic assumption, accounting for spike timings in the STP mean-field approximation	(26), (37)

model can approximate the low-dimensional dynamics of the QIF network only if X_i and U_i express unimodal, narrow distributions. This finding makes intuitive sense, since the mean-field approximation of the dynamics of U_i and X_i given by Eqs. (16) represents a first-order approximation. Our results confirm that this approximation only performs well if the mean over X_i and U_i contains much information about the actual underlying distributions. Thus, by providing these counter examples, we have shown that the mean-field model resulting from the Poisson assumption does not provide an exact mean-field description of the QIF network.

Since we are actually interested in the mean-field equations for SNN$_{pre}$ given by Eqs. (1), we now examine whether FRE$_{Poisson}$ can nonetheless provide an approximation of SNN$_{pre}$ under some conditions. To gain further insight into the relationship between the mean-field equations and the QIF network, we asked whether there exists a QIF network description for which the mean-field model given by (16a), (16b), (18a), and (18b) can be considered exact. Indeed, such a network exists and is easy to find. Since x and u are only driven by the mean-field firing rate r, we can just introduce microscopic variables U_i and X_i that enter the microscopic evolution equation for v_i in the same was as the macroscopic evolution equation for v [Eq. (18b)] and are also driven by the mean-field activity of the QIF network:

$$\tau \dot{V}_i = V_i^2 + \eta_i + I(t) + \frac{J\tau}{N} U_i X_i s, \qquad (19a)$$

$$\tau_x \dot{X}_i = 1 - X_i - \alpha X_i U_i s \tau_x, \qquad (19b)$$

$$\tau_u \dot{U}_i = U_0 - U_i + U_0(1 - U_i)s\tau_u, \qquad (19c)$$

$$s = \sum_{j=1}^{N} \sum_{k \backslash t_j^k < t} \int_{-\infty}^{t} \delta(t' - t_j^k) dt', \qquad (19d)$$

where $s = r$ is the mean firing rate across all neurons in the network. Apart from the description of the STP dynamics, this network description is equivalent to the one used in Ref. [34] for a QIF network with postsynaptic depression. Indeed, under a first-order approximation of the dynamics of x and u via the Poissonian assumption, the system given by Eqs. (1), a QIF network with presynaptic STP, is essentially approximated by Eqs. (19), a QIF network with postsynaptic STP (see Fig. 1 for

a visualization of the differences between the two). Hence, we will refer to the network given by Eqs. (19) as SNN$_{post}$.

Next, we compared the behavior of the two different QIF network descriptions (SNN$_{pre}$ and SNN$_{post}$) to the mean-field model dynamics. This was done to verify that FRE$_{Poisson}$ is indeed an exact mean-field model of SNN$_{post}$ and to see under which conditions pre- and postsynaptic STP have similar or different effects on the QIF network dynamics. To this end, we used bifurcation analysis to identify phase transitions in the mean-field model around which we compared the behavior of the three models. This way, we were able to set up stimulation paradigms that induce strong changes in the dynamic behavior of the mean-field model and evaluate whether the QIF networks express qualitatively similar phase transitions or not. Bifurcation analysis was performed numerically, using the Python software PyRates [43], which provides an interface to the parameter continuation software Auto-07p [44]. We initialized the mean-field model with either purely depressing synapses ($U_0 = 1.0$, $\alpha = 0.04$) or purely facilitating synapses ($U_0 = 0.2$, $\alpha = 0.0$). In each case, we performed a parameter continuation in the background excitability $\bar{\eta}$ for two different values of $\Delta \in 0.01, 0.4$. The latter introduces two different levels of firing rate heterogeneity to the QIF network. We expected this firing rate heterogeneity to directly affect the broadness of the distributions over X_i and U_i. If that is indeed the case, the mean-field model should provide a better description of the SNN$_{pre}$ dynamics for $\Delta = 0.01$ than for $\Delta = 0.4$.

As can be seen in Figs. 4(a) and 4(b), we identified fold bifurcations for facilitating synapses for $\Delta = 0.4$ as well as $\Delta = 0.01$. These fold bifurcations mark the outer limits of a bistable regime in which a stable high-activity focus and a stable low-activity node can coexist, separated by a saddle focus. Indeed, we find that the steady-state behavior of the mean-field model and SNN$_{post}$ can be forced toward either of the two stable equilibria via extrinsic stimulation. As shown for $\Delta = 0.4$ and $\Delta = 0.01$ in Figs. 4(a) and 4(b), respectively, there is always a very good agreement between those two models. Regarding SNN$_{pre}$, we failed to identify the bistable regime for $\Delta = 0.4$. In Fig. 4(a), it can be seen that the system behavior is only governed by a high-activity focus, even though the mean-field model predicts the coexistence of a low-activity stable node for $\bar{\eta} = -0.6$. Thus, the mean-field model fails to predict the behavior of the QIF network with
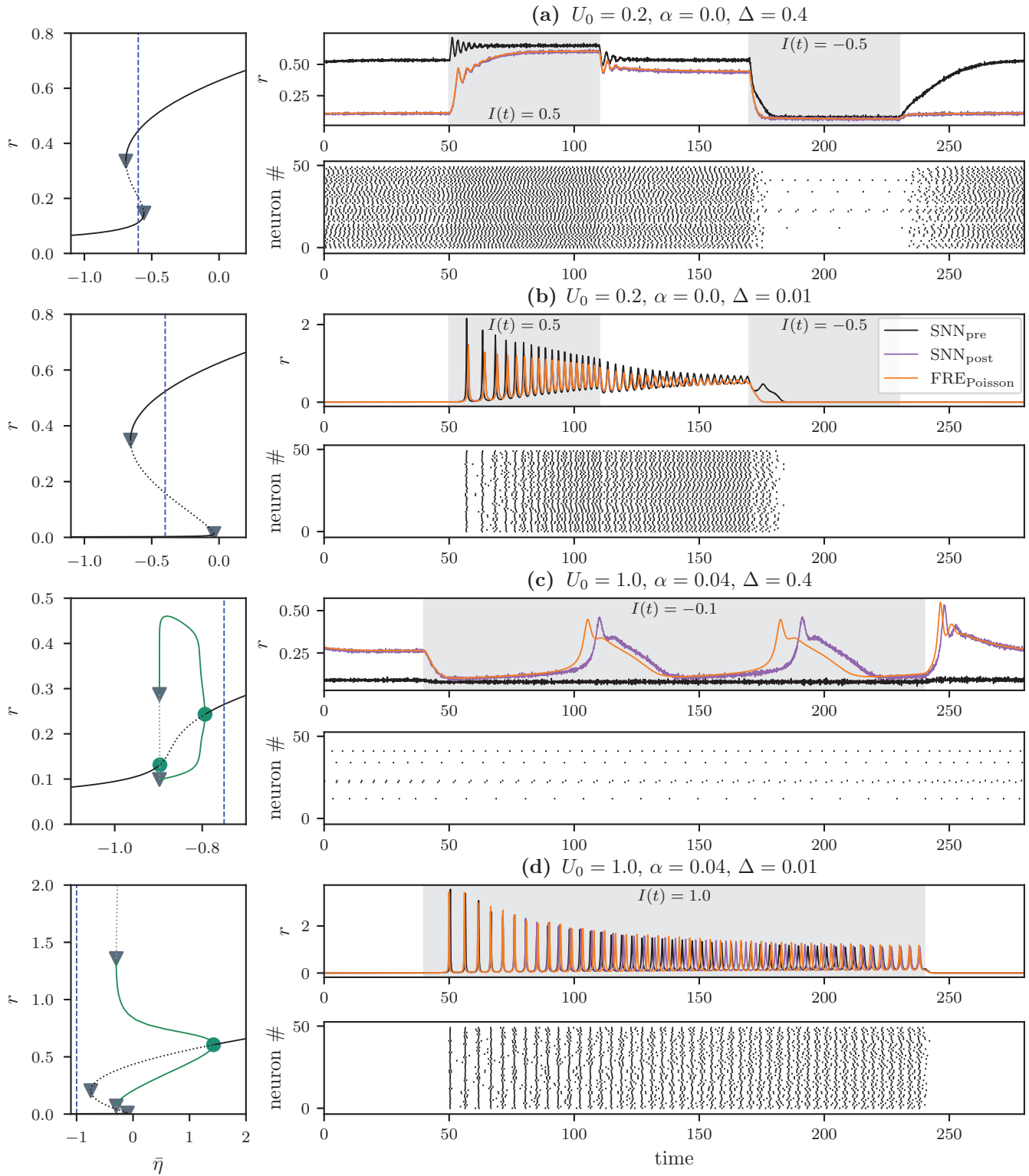
FIG. 4. Comparison between FRE$_{\text{Poisson}}$ (orange), SNN$_{\text{pre}}$ (black), and SNN$_{\text{post}}$ (purple) for four different parameter sets [(a)–(d)]. The first column shows 1D bifurcation diagrams in $\bar{\eta}$. Gray triangles represent fold bifurcations and green circles represent Andronov-Hopf bifurcations. Blue dashed lines mark the value of $\bar{\eta}$ that was used for the firing rate and spike raster plots in the second column. Spike raster plots show the spiking activity of 50 randomly selected neurons of SNN$_{\text{pre}}$. Gray areas represent time intervals during which an extrinsic input $I(t)$ was applied to the models [the effects of $I(t)$ and $\bar{\eta}$ on the QIF model are identical]. Remaining model parameters: $J = 8.0$, $\tau_u = 20.0$, $\tau_x = 50.0$, $\tau = 1.0$, and $N = 10\,000$.

presynaptic STP in this case. However, in the case of very low heterogeneity, we identified both stable states in SNN_{pre} and found a good agreement with the mean-field model [see Fig. 4(b)].

For depressing synapses, we found regimes of synchronized oscillations that emerge via Andronov-Hopf bifurcations for small as well as for high firing rate heterogeneity [see Figs. 4(c) and 4(d)]. Again, these oscillations could be induced in $\text{FRE}_{\text{Poisson}}$ as well as in SNN_{post} with a very good match between the two. Consistent with our findings for facilitating synapses, SNN_{pre} expressed oscillations only for $\Delta = 0.01$ [see Fig. 4(d)]. For higher firing rate heterogeneity ($\Delta = 0.4$), the network did not show any tendency to oscillate at all, even though the mean-field model predicted oscillations to be present at $\bar{\eta} = -0.85$ [see Fig. 4(c)].

Thus, our results confirm that $\text{FRE}_{\text{Poisson}}$ is indeed an exact mean-field equation of SNN_{post}. Furthermore, they demonstrate that SNN_{pre} and SNN_{post} can behave both very differently and very similarly, depending on the firing rate heterogeneity inside the network. In our simulations, we were able to control this heterogeneity successfully via the parameter Δ. In regimes of low firing rate heterogeneity, SNN_{pre} and SNN_{post} expressed similar behavior, thus allowing for a good approximation of the mean-field dynamics of SNN_{pre} via $\text{FRE}_{\text{Poisson}}$. In regimes of high firing rates heterogeneity, the opposite was the case. In the next sections, we investigate whether more accurate mean-field models of QIF networks with presynaptic STP can be derived and, if so, how they perform near the parameter regimes described in this section.

V. MULTIPOPULATION APPROXIMATION OF DISTRIBUTED PARAMETERS IN THE QIF NETWORK

In the previous section, we have found that $\text{FRE}_{\text{Poisson}}$ is in good agreement with the dynamics of SNN_{pre}, when the distribution of η_i is particularly narrow, i.e., when $\Delta \ll 1$. Here we exploit this fact and approximate the mean-field dynamics by dividing the microscopic network into subnetworks with narrow distributions in η_i. In other words, the Lorentzian distribution with $\{\bar{\eta}, \Delta\}$ is divided into a set of M Lorentzian distributions with $\{\bar{\eta}_m, \Delta_m\}$, $m = 1, \ldots, M$, such that

$$\frac{\Delta/\pi}{(\eta - \bar{\eta})^2 + \Delta^2} \approx \frac{1}{M} \sum_{m=1}^{M} \frac{\Delta_m/\pi}{(\eta - \bar{\eta}_m)^2 + \Delta_m^2}. \quad (20)$$

The resulting set of equations for the evolution of the mean-field variables is then given by

$$\tau \dot{r}_m = \frac{\Delta_m}{\pi \tau} + 2 r_m v_m, \quad (21a)$$

$$\tau \dot{v}_m = v_m^2 + \bar{\eta}_m + I(t) + \frac{J\tau}{M} \sum_{n=1}^{M} x_n u_n r_n - (\pi r_m \tau)^2, \quad (21b)$$

$$\dot{x}_m = \frac{1 - x_m}{\tau_x} - \alpha u_m x_m r_m, \quad (21c)$$

$$\dot{u}_m = \frac{U_0 - u}{\tau_u} + U_0(1 - u_m)r_m. \quad (21d)$$

We will refer to this set of mean-field equations as FRE_{mpa} for multipopulation approximation. One assumption we make

here is that each subnetwork contains the same number of neurons, which means that the weights for each subnetwork are the same, and the mean-field variables can be obtained by computing the mean $y = (1/M) \sum_{m=1}^{M} y_m$, where y represents the mean-field variable under consideration. The parameters $\bar{\eta}_m$ and Δ_m are chosen as follows:

$$\bar{\eta}_m = \bar{\eta} + \Delta \tan \frac{\pi(2m - M - 1)}{2(M + 1)}, \quad (22a)$$

$$\Delta_m = \Delta \left[\tan \frac{\pi(2m - M - 1/2)}{2(M + 1)} \right.$$
$$\left. - \tan \frac{\pi(2m - M - 3/2)}{2(M + 1)} \right]. \quad (22b)$$

The density of the parameters η_m follows the Lorentzian distribution, and the Δ_m are chosen such that the half-widths approximately match the distances between the centers of the distributions of the subnetworks, i.e., $\bar{\eta}_{m+1} - \bar{\eta}_m \approx \Delta_{m+1} + \Delta_m$. The results are shown in Fig. 5(a). As can be seen, even at large M the adaptation variables still show a small discrepancy with the result obtained from the spiking neural network SNN_{pre}. We hypothesize that this difference is due to different results for the adaptation variables when the firing rate is assumed constant, and when it is assumed to be a spike train with constant ISI, as shown in Fig. 2. In other words, we expect that accounting for the fact that $\text{FRE}_{\text{Poisson}}$ was derived for SNN_{pre} II instead of SNN_{pre} will reduce the difference. As the adaptation variables are in essence time-averaged quantities, the adaptation variables could be posed as $x = (X^- + X^+)/2$ and $u = (U^- + U^+)/2$. However, with the update rules $U^+ = U^- + U_0(1 - U^-)$ and $X^+ = X^- - \alpha U^+ X^-$, this would yield out-of-bound values for X^- at $x = 1$, and U^- at $u = 0$. The results shown in Fig. 2 suggest that the mean-field variables are closest to X^- and U^-, which is why we set $X^- \approx x$, and $U^- \approx u$. The update rule for U^+ gives the following correction term:

$$U^+(u) \approx u + U_0(1 - u). \quad (23)$$

Inserting this term into the mean-field equations for FRE_{mpa} produces a closer match of the mean-field variables with the results of the microscopic model SNN_{pre}, see Fig. 5(b).

As a final test of the predictive accuracy of FRE_{MPA}, we examined how well the model can predict the onset of oscillations in the QIF network. Using bifurcation analysis, we identified the Hopf bifurcation leading to the oscillations in Fig. 4(c) and investigated the locus of that Hopf bifurcation in the 2D parameter space spanned by $\bar{\eta}$ and Δ. This, we did for both $\text{FRE}_{\text{Poisson}}$ and FRE_{MPA} with $M = 100$ mean-field populations. As shown in Fig. 6(a), we found that the Hopf curves emerged from a Bogdanov-Takens bifurcation in both FRE models. This represents the same bifurcation structure as has already been identified for QIF networks with SD (see Figs. 2 and 4 in Ref. [34] for the corresponding 1D and 2D bifurcation diagrams, respectively). Furthermore, we have shown the corresponding 1D bifurcation diagrams for the $\text{FRE}_{\text{Poisson}}$ model for $\Delta = 0.4$ and $\Delta = 0.01$ in Fig. 4(c) and 4(d), respectively. Thus, we expect stable oscillations to exist in the regions enclosed by the Hopf curves. As shown in Fig. 6(a), the difference between the Hopf curves predicted
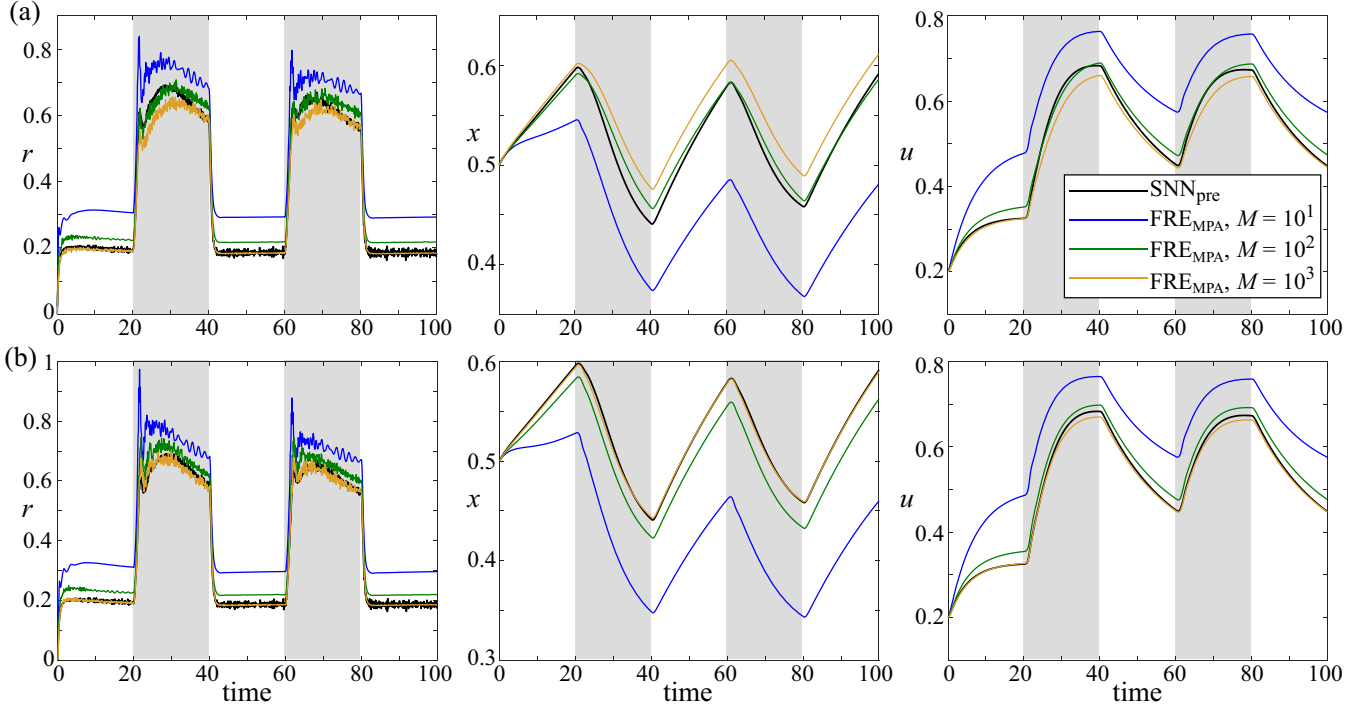
FIG. 5. Comparison of the mean-field variables of the microscopic spiking neural network, and the mean-field model of the spiking neural network divided into M subnetworks with narrow distribution [multipopulation approximation (MPA)]. Gray areas indicate time intervals with $I(t) = 3.0$ [the effects of $I(t)$ and $\bar{\eta}$ on the QIF model are identical]. (a) MPA with standard mean-field description and (b) MPA with correction term for U^+. Parameters: $\alpha = 0.1$, $\tau = 1.0$, $\Delta = 2.0$, $\bar{\eta} = -3.0$, $J = 15.0\sqrt{\Delta}$, $\tau_x = 50.0$, $\tau_u = 20.0$, and $N = 10\,000$.

by $\mathrm{FRE_{Poisson}}$ and $\mathrm{FRE_{MPA}}$ becomes larger when Δ increases. For $\Delta = 0.4$, $\mathrm{FRE_{Poisson}}$ predicts stable oscillations to exist at $\bar{\eta} = -0.85$, which we already failed to find in the QIF network in Fig. 4(d). $\mathrm{FRE_{MPA}}$ predicts the existence of a stable node at $\bar{\eta} = -0.85$, however, and the existence of stable oscillations for $-0.66 < \bar{\eta} < -0.6$. To see whether the oscillations
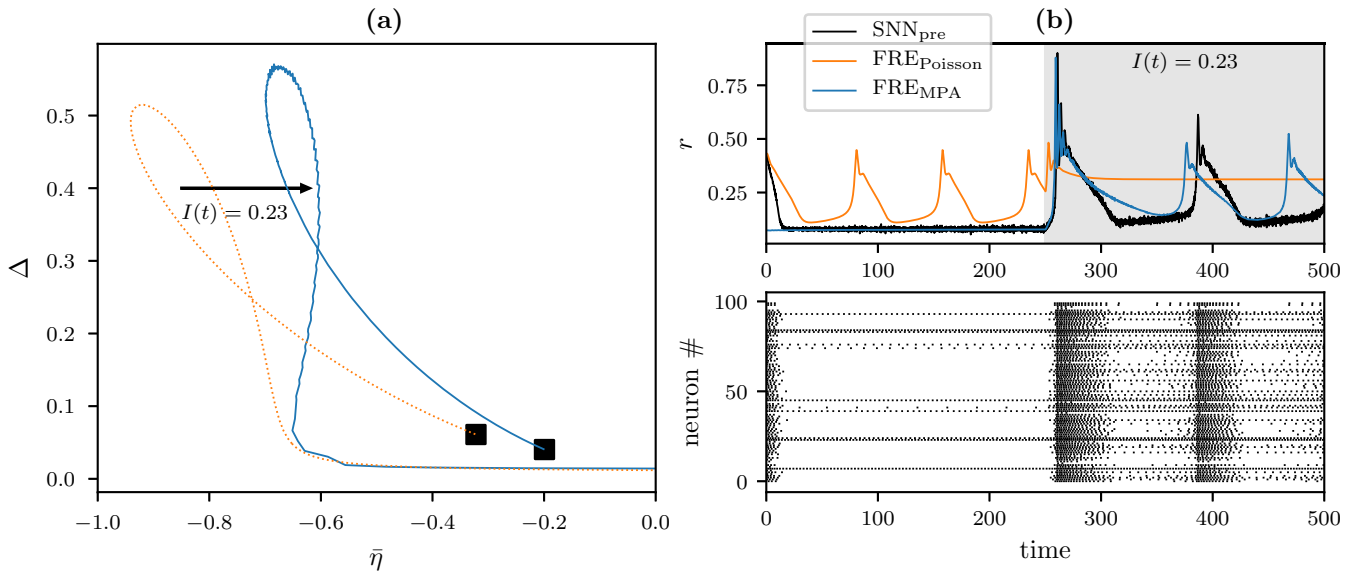

FIG. 6. Phase transitions between steady-state and oscillatory regimes in $\mathrm{FRE_{Poisson}}$ and $\mathrm{FRE_{MPA}}$. (a) 2D bifurcation diagram of the Hopf curve in $\mathrm{FRE_{Poisson}}$ (orange) and $\mathrm{FRE_{MPA}}$ (blue). The arrow represents the phase transition introduced by $I(t)$ in either model [the effects of $I(t)$ and $\bar{\eta}$ on the QIF model are identical]. The black square represents the Bogdanov-Takens bifurcation from which the Hopf bifurcations emerge. (b) The first row shows the simulated firing dynamics of the spiking neural network and both mean-field models. The second row shows the corresponding spiking activity of 100 randomly selected neurons of $\mathrm{SNN_{pre}}$. Parameters: $\alpha = 0.04$, $U_0 = 1.0$, $\tau = 1.0$, $\Delta = 0.4$, $\bar{\eta} = -0.85$, $J = 8.0$, $\tau_x = 50.0$, $\tau_u = 20.0$, $N = 10\,000$, $M = 100$, and $I(t) = 0.23$ for $t > 250$ and $I(t) = 0.0$ otherwise.

predicted by FRE$_{MPA}$ indeed exist in SNN$_{pre}$, we performed numerical simulations where we initialized the QIF network at $\bar{\eta} = -0.85$ and then forced it toward $\bar{\eta} = -0.62$ via extrinsic stimulation. As can be seen in Fig. 6(b), the QIF network expressed steady-state behavior for $\bar{\eta} = -0.85$ and started to oscillate when pushed to $\bar{\eta} = -0.62$. Hence, FRE$_{MPA}$ correctly predicted the existence of oscillatory bursts in the QIF network for $M = 100$, but not for $M = 1$, for which FRE$_{MPA}$ reduces to FRE$_{Poisson}$. The bursts have similar properties as the ones found in QIF networks with postsynaptic plasticity [34] and can be expected to result from the interaction between synaptic short-term depression and recurrent excitation via the network. Comparing the firing rate dynamics of FRE$_{MPA}$ and SNN$_{pre}$ in Fig. 6 reveals a slight difference between the oscillation period of the mean-field model and the QIF network. This difference shows that FRE$_{MPA}$ can not be considered an exact mean-field model, even for $M = 100$. Still, we find that it captures the phase transitions inside SNN$_{pre}$ well and thus provides a reasonable trade-off between accuracy and computational complexity.

VI. ADIABATIC APPROXIMATION OF STP DYNAMICS

For simplification, we will consider synapses with mere short-term depression in this section, since we showed in Sec. IV that the mismatch between the mean-field model FRE$_{Poisson}$ and the QIF networks SNN$_{pre}$ and SNN$_{pre}$ II could be reproduced in this simpler case as well. We thus consider the microscopic system given by

$$\tau \dot{V}_i = V_i^2 + \eta_i + I(t) + \frac{J\tau}{N} \sum_{j=1}^{N} X_j^- S_j, \qquad (24a)$$

$$\tau_x \dot{X}_i = 1 - X_i - \alpha X_i^- S_i \tau_x, \qquad (24b)$$

$$S_i = \sum_{k \setminus t_j^k < t} \int_{-\infty}^{t} a(t - t') \delta(t' - t_j^k) dt'. \qquad (24c)$$

In this system, we approximate the STP dynamics given by Eq. (24b) via a linear differential operator L, i.e., $LX_i(t) = S_i(t)$. In such a linear case, a Green's function $G(t)$ exists that allows one to express the dynamics of X_i via a convolution of $G(t)$ with the spiking activity of neuron i:

$$X_i(t) = \int_{-\infty}^{t} G(t - t') S_i(t') dt' = G * S_i. \qquad (25)$$

Then, since S_i is related to $z(\eta_i, t)$ via $S_i \pi = z(\eta_i, t)$, Eq. (4) can be written as

$$\partial_t w(\eta, t) = i \left[\frac{-w(\eta, t)^2 + \eta + I(t)}{\tau} \right. $$
$$\left. + J \left(G * \frac{\mathrm{Re}(w)}{\pi} \right) \mathrm{Re}(w) \right]. \qquad (26)$$

To solve Eq. (26) for r and v, the effective firing rate $r_{\mathrm{eff}} = \int_{-\infty}^{\infty} (G * r(\eta)) r(\eta) g(\eta) d\eta$ must be determined, which requires one to evaluate the product between the single cell firing rate and a convolution of itself. This makes it difficult to find a closed-form solution for r and v, since the synaptic depression kernel G cannot simply be pulled out from the convolution integral. The simplest approximation of this

problem is to replace the convolution integral by a mean synaptic depression, as is done for the Poissonian assumption. Alternatively, we assume that the dynamics of X_i are slow in comparison to the dynamics of v_i. For the relaxation dynamics of X_i, this assumption is met if $\tau_x \gg \tau$. We note here, however, that the spiking activity of the neuron also introduces a relatively fast timescale to Eq. (24b), which may violate our assumption. Still, under this assumption, we can apply an adiabatic approximation to the system and consider the dynamics of the fast subsystem for effectively constant adaptation (see Refs. [34,45] for a similar approach):

$$\tau \dot{V}_i = V_i^2 + \eta_i + I(t) + \frac{J\tau}{N} \sum_{j=1}^{N} X_j^- S_j, \qquad (27a)$$

$$S_i = \sum_{k \setminus t_j^k < t} \int_{-\infty}^{t} \delta(t' - t_j^k) dt', \qquad (27b)$$

where X_j is approximated as neuron-specific constant. Due to the Lorentzian distribution of the background excitabilities η_i and the resulting heterogeneity of single cell firing rates in the network, X_i cannot be assumed as homogeneous across neurons. Instead, it must be considered a distributed quantity, governed by a probability density function $h(X_i)$. Then, the main difficulty in developing the mean-field description lies in the fact that $h(X_i)$ is generally unknown if a mean-field variable is considered. More precisely, if we consider the mean-field variable x that describes the average synaptic depression across the network, little is known about the distribution of the microscopic variables X_i, which is required to determine the effective firing rate r_{eff}. By using the adiabatic approximation, we argue that an approximation of r_{eff} can be obtained by estimating the distributions $X(\eta)$ and $r(\eta)$ from the mean-field variables in the stationary case, and solving

$$r_{\mathrm{eff}} = \int_0^1 \int_{-\infty}^{\infty} X r(\eta) h(X|\eta) g(\eta) d\eta dX. \qquad (28)$$

Assuming independent Lorentzian density functions for h and g, i.e., $h(X|\eta) g(\eta) = h(X) g(\eta)$, Eq. (26) would only need to be evaluated at the poles in the lower half-planes $\pi r(t) + iv(t) = w(\bar{\eta} - i\Delta, \bar{X} - i\Delta_X, t)$, where $\bar{X}$ and Δ_X would represent the center and HWHM of the Lorentzian distribution over X, respectively. Then, the effect of presynaptic STP on the network dynamics would effectively reduce to a distribution over the coupling parameter J. For the mean-field equations of a QIF network with distributed coupling parameters see Ref. [29]. However, h and g cannot be assumed to be independent, since η_i controls the firing rate of neuron i, which in turn controls its synaptic depression X_i. Furthermore, X is bound between $[0, 1]$ and hence a Lorentzian distribution cannot be assumed. In the upper row of Fig. 3, we show the evolution of the distribution over $X_i U_i$ for three different parametrizations, corresponding to a purely depressing synapse, a purely facilitating synapse, and a synapse with facilitation and depression acting on different timescales. Importantly, the evolution of the distribution reveals that it is not always unimodal. For purely depressing synapses, it clearly expresses an at least bimodal distribution over the whole time course. Thus, finding an appropriate form of h that holds in

general is a highly nontrivial problem that we did not find a solution for.

To further simplify the problem, we assume that the depression of a neuron's efferent synapses X_i is merely a function of the firing rate r_i of the same neuron. The stationary firing rate of a QIF neuron in response to an external input I_{in} is $\sqrt{I_{in}}/\pi$ if $I_{in} > 0$, and zero otherwise. Hence, the distribution of firing rates for a given input is (in the stationary case) given by

$$r(\eta; I_{in}) = H(\eta + I_{in})\sqrt{\eta + I_{in}}/\pi, \qquad (29)$$

where H is the Heaviside step function. Therefore, for any given mean-field firing rate r one can find a unique constant I_r for which

$$r = \int_{-\infty}^{\infty} r(\eta; I_r)g(\eta)d\eta, \qquad (30)$$

which allows us to translate the mean-field variable r into the distribution $r(\eta; I_r)$.

Similarly, we can use the assumption that X_i is a function of r_i to translate the mean-field variable for synaptic depression, x, into the distribution $X(\eta; I_x)$. First, we use the rate relationship given by Eq. (13) to approximate

$$x(\eta; I_x) = 1/[1 + \alpha\tau_x r(\eta; I_x)], \qquad (31)$$

for any given input I_x, and then define

$$x_1 = \int_{-\infty}^{\infty} g(\eta)/[1 + \alpha\tau_x r(\eta; I_x)]d\eta. \qquad (32)$$

Alternatively, we can use Eq. (12) to approximate the distribution $x(\eta)$ in the spiking scenario:

$$x(\eta; I_x) = \frac{1 - \exp[-1/\tau_x r(\eta; I_x)]}{1 - (1 - \alpha)\exp[-1/\tau_x r(\eta; I_x)]}, \qquad (33)$$

which yields

$$x_2 = \int_{-\infty}^{\infty} \frac{\{1 - \exp[-1/\tau_x r(\eta; I_x)]\}g(\eta)}{1 - (1 - \alpha)\exp[-1/\tau_x r(\eta; I_x)]}d\eta. \qquad (34)$$

Having obtained I_r and I_x, we can ultimately compute

$$r_{eff} = \int_{-\infty}^{\infty} r(\eta; I_r)x(\eta; I_x)g(\eta)d\eta, \qquad (35)$$

where $x(\eta; I_x)$ is either chosen for the rate scenario [Eq. (31)], or in the spike scenario [Eq. (33)]. This requires one to solve

$$r_{eff} = \frac{\Delta}{\pi^2} \int_{\min(-I_x, -I_r)}^{\infty} \frac{1}{1 + \alpha\tau_x\sqrt{\eta + I_x}} \frac{\sqrt{\eta + I_r}}{(\eta - \bar{\eta})^2 + \Delta^2}d\eta, \qquad (36)$$

in the rate scenario, and

$$r_{eff} = \frac{\Delta}{\pi^2} \int_{\min(-I_x, -I_r)}^{\infty} \frac{\exp\left(\frac{\pi}{\tau_x\sqrt{\eta + I_x}}\right) - 1}{\exp\left(\frac{\pi}{\tau_x\sqrt{\eta + I_x}}\right) - (1 - \alpha)}$$
$$\times \frac{\sqrt{\eta + I_r}}{(\eta - \bar{\eta})^2 + \Delta^2}d\eta, \qquad (37)$$

in the spiking scenario. We refer to this mean-field model as FRE$_{aa}$ for adiabatic approximation, with FRE$_{aa1}$ and FRE$_{aa2}$ denoting the mean-field model considering the rate and spike scenario, respectively.

The integrals involved in this approximation are hard to evaluate analytically, therefore we solve these integrals numerically for a range of values of I_r and I_x and create look-up tables for I_r, I_x, and r_{eff} in order to be able to integrate the resulting model equations numerically. In Fig. 7 we compare the results of the mean-field model FRE$_{aa}$ with the dynamics of the spiking neural network SNN$_{pre}$, and the mean-field model FRE$_{Poisson}$. We find that FRE$_{aa}$ is closer to the microscopic dynamics of SNN$_{pre}$ than FRE$_{Poisson}$.

VII. CONCLUSION

In this work, we examined whether spiking neural networks with presynaptic short-term plasticity allow for the derivation of low-dimensional mean-field equations via the Lorentzian ansatz described in Ref. [29]. To this end, we considered heterogeneous, all-to-all coupled QIF networks with presynaptic STP dynamics, described by a well-known phenomenological model of synaptic short-term depression and facilitation [39]. For such QIF networks, other forms of STP have already been shown to be compatible with the Lorentzian ansatz [34]. In the case of presynaptic STP, we identified the evaluation of the effective network input r_{eff} as the central problem for a mean-field derivation via the Lorentzian ansatz. This effective network input represents a weighted sum of incoming spikes, where the weights are given by the presynaptic depression and facilitation terms. We presented three different approaches to express r_{eff} and thus find the mean-field equations: First, a mean-field description of the STP dynamics via the Poissonian assumption used in Ref. [37]; second, a multipopulation approximation that approximates distributed parameters inside the QIF network via a set of coupled subpopulations with different parametrizations; and third, an adiabatic approximation of the STP timescales.

For the first approach, the effective network input r_{eff} is approximated by a modulation of the mean-field firing rate with an average depression and an average facilitation. Our analysis revealed that this approach essentially approximates presynaptic STP with postsynaptic STP. We compared the behavior of QIF networks with pre- vs. postsynaptic STP and found that they can express substantial qualitative differences in their dynamics, especially when SNN$_{pre}$ expresses a high firing rate heterogeneity across neurons. Near such regimes, FRE$_{Poisson}$ follows the dynamics of SNN$_{post}$, and thus fails to capture the behavior of SNN$_{pre}$. It is worth noticing that the mean-field derivation via the Poissonian assumption works well for networks of homogeneous Poisson neurons with independent noise [42]. In such networks, single cell firing rates can differ momentarily due to noise, but approach the same rate when averaged over increasing time intervals. This is a very different scenario compared to the QIF network considered here, where the Lorentzian distribution over η_i causes substantial heterogeneity in the single cell firing rates. Hence, the Poissonian approximation becomes worse the stronger the heterogeneity of single cell firing rates inside the QIF network is. In Ref. [37], where the Poissonian approximation was first applied to a QIF network with presynaptic STP, the authors chose QIF networks with relatively low firing rate heterogeneity, leading to a good correspondence with the mean-field model. Here we clarified that this correspondence does not
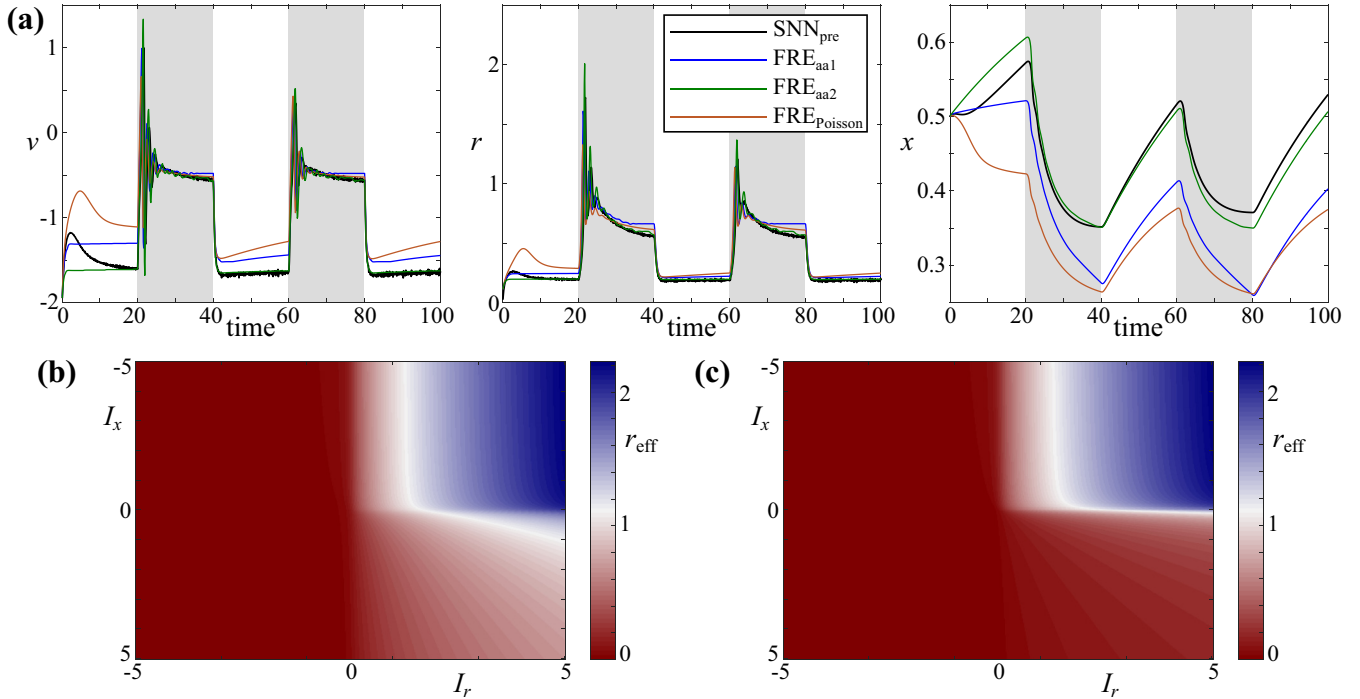
FIG. 7. (a) Comparison of the mean-field variables of the microscopic spiking neural network, the mean-field model using the Poissonian assumption, and the mean-field model with approximation of the effective firing rate. Gray areas indicate time intervals with $I(t) = 3.0$ [the effects of $I(t)$ and $\bar{\eta}$ on the QIF model are identical]. (b) Look-up table for the conversion of I_r and I_x into the effective firing rate r_{eff} in the rate scenario [Eq. (36); FRE$_{\text{aa1}}$], and (c) in the spiking scenario [Eq. (37); FRE$_{\text{aa2}}$]. Parameters: $\alpha = 0.1$, $\tau = 1.0$, $\Delta = 1.0$, $\bar{\eta} = -2.0$, $J = 15.0$, $\tau_x = 50.0$, $\tau_u = 20.0$, and $N = 10\,000$.

generalize to regimes where the QIF network expresses more heterogeneous firing rates.

Populations of neurons that naturally express heterogeneous firing rates exist in subcortical structures, for example. Single cell firing rates in the globus pallidus have been shown to differ substantially across neurons [46,47]. This firing rate heterogeneity has been suggested as an important desynchronization mechanism of pallidal activity [48,49]. Our results suggest that studying the mean-field dynamics in such a population via FRE$_{\text{Poisson}}$ comes at the risk of substantial errors. We thus developed a mean-field model that addresses the issue of high firing rate heterogeneities. Since the distribution over η_i is the source of heterogeneity in the QIF network, we attempted to improve the mean-field model by considering a set of coupled subnetworks with distinct, but narrow distributions over η_i. This way, the neurons inside each subpopulation are parametrized such that they express a considerably lower firing rate heterogeneity than the overall network. We found that, by increasing the number of subpopulations, the mean-field model converges to the QIF network behavior. Of course, this approach leads to mean-field models of relatively high dimensionality. Still, we found that a mean-field model with 100 subpopulations (i.e., a 400-dimensional model), accurately predicted phase transitions of the QIF network from steady-state to oscillatory behavior in a regime where FRE$_{\text{Poisson}}$ failed to do so. Importantly, FRE$_{\text{MPA}}$ can be investigated by means of bifurcation analysis, whereas the corresponding SNN cannot. Furthermore, the multipopulation approach allows to consider other parameter distributions than the Lorentzian distribution considered in this manuscript. In

Ref. [50] it is shown that a similar approach can be used for Gaussian parameter distributions in QIF networks, leading to a considerable reduction in the number of required subpopulations to achieve a good approximation of the QIF mean-field dynamics. Thus, we argue that the multipopulation approximation provides a useful tool for mean-field analyses, the dimensionality of which can be freely chosen to achieve a desired level of approximation accuracy.

As an alternative to the Poissonian approximation, we applied an adiabatic approximation to the QIF network, assuming slow STP dynamics in comparison to the QIF dynamics. This assumption is supported by experimental results that suggest depression and facilitation recovery timescales that are at least 10 times slower than typical membrane potential timescales [37,39,51]. Previously, this approach has been used successfully for the derivation of mean-field equations for QIF networks with spike-frequency adaptation [34]. By approximating the presynaptic STP dynamics as slow, they can be considered as constant, distributed quantities in the fast subsystem. This way, the STP dynamics do not have to be considered for the evaluation of r_{eff}. Instead, appropriate distributions over the STP constants have to be chosen. In our work, we derived analytical solutions of the microscopic STP dynamics in the stationary case and used these solutions to approximate the STP distributions. This approach can be considered exact for the description of steady-state solutions but not for transient dynamics. That is, the network must have converged to an equilibrium for our approximation to be accurate. Still, we find that our adiabatic approximation provides a more accurate approximation of the mean-field

dynamics of the QIF network dynamics than the Poissonian approximation, even for transient dynamics. A disadvantage of this method is, however, that we had to approximate the integrals over the STP distribution numerically and calculate r_{eff} via look-up tables. This makes it more difficult to implement the model equations and perform parameter continuations.

In conclusion, we performed a thorough analysis of the problems that arise when attempting to derive the mean-field equations for QIF networks with synaptic short-term plasticity. Though we did not find a set of exact, closed-form mean-field equations, we provided two different mean-field approximations that we found to be more accurate than a previously proposed mean-field model. Both of these

mean-field approximations can capture the qualitative dynamics of the QIF network and can thus be used for future investigations of its macroscopic dynamics. Finally, our work provides insight into the distinct effects that pre- vs. postsynaptic STP can have on the mean-field dynamics of spiking neural networks.

[1] E. Başar, *Chaos in Brain Function* (Springer Science & Business Media, New York, 2012).

[2] D. R. Chialvo, Emergent complex neural dynamics, Nat. Phys. **6**, 744 (2010).

[3] G. Deco, V. K. Jirsa, and A. R. McIntosh, Emerging concepts for the dynamical organization of resting-state activity in the brain, Nat. Rev. Neurosci. **12**, 43 (2011).

[4] A. K. Engel, P. Fries, and W. Singer, Dynamic predictions: Oscillations and synchrony in top–down processing, Nat. Rev. Neurosci. **2**, 704 (2001).

[5] T. R. Knösche, C. Neuhaus, J. Haueisen, K. Alter, B. Maess, O. W. Witte, and A. D. Friederici, Perception of phrase structure in music, Hum. Brain Map. **24**, 259 (2005).

[6] T. Kujala, M. Tervaniemi, and E. Schröger, The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations, Biol. Psychol. **74**, 1 (2007).

[7] V. K. Jirsa, W. C. Stacey, P. P. Quilichini, A. I. Ivanov, and C. Bernard, On the nature of seizure dynamics, Brain **137**, 2210 (2014).

[8] P. T. Sadtler, K. M. Quick, M. D. Golub, S. M. Chase, S. I. Ryu, E. C. Tyler-Kabara, B. M. Yu, and A. P. Batista, Neural constraints on learning, Nature (Lond.) **512**, 423 (2014).

[9] J. D. Murray, A. Bernacchia, N. A. Roy, C. Constantinidis, R. Romo, and X.-J. Wang, Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex, Proc. Natl. Acad. Sci. USA **114**, 394 (2017).

[10] A. Babloyantz and A. Destexhe, Low-dimensional chaos in an instance of epilepsy, Proc. Natl. Acad. Sci. USA **83**, 3513 (1986).

[11] J. A. S. Kelso, *Dynamic Patterns: The Self-organization of Brain and Behavior* (MIT Press, Cambridge, UK, 1995).

[12] A. Celletti and A. E. P. Villa, Low-dimensional chaotic attractors in the rat brain, Biol. Cybern. **74**, 387 (1996).

[13] A. Bollimunta, Y. Chen, C. E. Schroeder, and M. Ding, Neuronal mechanisms of cortical alpha oscillations in awake-behaving macaques, J. Neurosci. **28**, 9976 (2008).

[14] A. Spiegler, T. R. Knösche, K. Schwab, J. Haueisen, and F. M. Atay, Modeling brain resonance phenomena using a neural mass model, PLoS Comput. Biol. **7**, e1002298 (2011).

[15] G. Deco and V. K. Jirsa, Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors, J. Neurosci. **32**, 3366 (2012).

[16] G. Deco, V. K. Jirsa, P. A. Robinson, M. Breakspear, and K. Friston, The dynamic brain: From spiking neurons to neural

masses and cortical fields, PLoS Comput. Biol. **4**, e1000092 (2008).

[17] H. R. Wilson and J. D. Cowan, Excitatory and inhibitory interactions in localized populations of model neurons, Biophys. J. **12**, 1 (1972).

[18] F. H. Lopes da Silva, A. Hoeks, H. Smits, and L. H. Zetterberg, Model of brain rhythmic activity, Kybernetik **15**, 27 (1974).

[19] W. Freeman, Models of the dynamics of neural populations, Electroencephalogr. Clin. Neurophysiol. Suppl. **34**, 9 (1978).

[20] B. H. Jansen and V. G. Rit, Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns, Biol. Cybern. **73**, 357 (1995).

[21] P. A. Robinson, C. J. Rennie, and J. J. Wright, Propagation and stability of waves of electrical activity in the cerebral cortex, Phys. Rev. E **56**, 826 (1997).

[22] S. El Boustani and A. Destexhe, A master dquation formalism for macroscopic modeling of asynchronous irregular activity states, Neural Comput. **21**, 46 (2009).

[23] M. A. Buice, J. D. Cowan, and C. C. Chow, Systematic fluctuation expansion for neural network activity equations, Neural Comput. **22**, 377 (2009).

[24] T. Schwalger, M. Deger, and W. Gerstner, Towards a theory of cortical columns: From spiking neurons to interacting neural populations of finite size, PLoS Comput. Biol. **13**, e1005507 (2017).

[25] E. Ott and T. M. Antonsen, Low dimensional behavior of large systems of globally coupled oscillators, Chaos **18**, 037113 (2008).

[26] Y. Kuramoto, Collective synchronization of pulse-coupled oscillators and excitable units, Physica D **50**, 15 (1991).

[27] T. B. Luke, E. Barreto, and P. So, Complete classification of the macroscopic behavior of a heterogeneous network of theta neurons, Neural Comput. **25**, 3207 (2013).

[28] S. Coombes and A. Byrne, Next generation neural mass models, in *Nonlinear Dynamics in Computational Neuroscience*, PoliTO Springer Series, edited by F. Corinto and A. Torcini (Springer International Publishing, Cham, 2019), pp. 1–16.

[29] E. Montbrió, D. Pazó, and A. Roxin, Macroscopic Description for Networks of Spiking Neurons, Phys. Rev. X **5**, 021028 (2015).

[30] I. Ratas and K. Pyragas, Macroscopic self-oscillations and aging transition in a network of synaptically coupled quadratic integrate-and-fire neurons, Phys. Rev. E **94**, 032215 (2016).

[31] A. Byrne, M. J. Brookes, and S. Coombes, A mean field model for movement induced changes in the beta rhythm, J. Comput. Neurosci. **43**, 143 (2017).

[32] M. di Volo and A. Torcini, Transition from Asynchronous to Oscillatory Dynamics in Balanced Spiking Networks with Instantaneous Synapses, Phys. Rev. Lett. **121**, 128301 (2018).

[33] B. Pietras, F. Devalle, A. Roxin, A. Daffertshofer, and E. Montbrió, Exact firing rate model reveals the differential effects of chemical versus electrical synapses in spiking networks, Phys. Rev. E **100**, 042412 (2019).

[34] R. Gast, H. Schmidt, and T. R. Knösche, A mean-field description of bursting dynamics in spiking neural networks with short-t adaptation, Neural Comput. **32**, 1615 (2020).

[35] P. Suffczynski, S. Kalitzin, G. Pfurtscheller, and F. Lopes da Silva, Computational model of thalamo-cortical networks: Dynamical control of alpha rhythms in relation to focal attention, Int. J. Psychophysiol. **43**, 25 (2001).

[36] R. Moran, S. Kiebel, K. Stephan, R. Reilly, J. Daunizeau, and K. Friston, A neural mass model of spectral responses in electrophysiology, NeuroImage **37**, 706 (2007).

[37] H. Taher, A. Torcini, and S. Olmi, Exact neural mass model for synaptic-based working memory, PLoS Comput. Biol. **16**, e1008533 (2020).

[38] A. Levina, J. M. Herrmann, and T. Geisel, Dynamical synapses causing self-organized criticality in neural networks, Nat. Phys. **3**, 857 (2007).

[39] M. Tsodyks, K. Pawelzik, and H. Markram, Neural networks with dynamic synapses, Neural Comput. **10**, 821 (1998).

[40] G. B. Ermentrout and N. Kopell, Parabolic bursting in an excitable system coupled with a slow oscillation, SIAM J. Appl. Math. **46**, 233 (1986).

[41] B. Pietras and A. Daffertshofer, Ott-Antonsen attractiveness for parameter-dependent oscillatory systems, Chaos **26**, 103101 (2016).

[42] V. Schmutz, W. Gerstner, and T. Schwalger, Mesoscopic population equations for spiking neural networks with synaptic short-term plasticity, J. Math. Neurosci. **10**, 5 (2020).

[43] R. Gast, D. Rose, C. Salomon, H. E. Möller, N. Weiskopf, and T. R. Knösche, PyRates—A Python framework for rate-based neural simulations, PLoS One **14**, e0225900 (2019).

[44] E. J. Doedel, T. F. Fairgrieve, B. Sandstede, A. R. Champneys, Y. A. Kuznetsov, and X. Wang, AUTO-07P: Continuation and Bifurcation Software for Ordinary Differential Equations, Technical Report (2007).

[45] G. Gigante, M. Mattia, and P. D. Giudice, Diverse Population-b Modes of Adapting Spiking Neurons, Phys. Rev. Lett. **98**, 148101 (2007).

[46] H. Kita, A. Nambu, K. Kaneda, Y. Tachibana, and M. Takada, Role of ionotropic glutamatergic and GABAergic inputs on the firing activity of neurons in the external pallidum in awake monkeys, J. Neurophysiol. **92**, 3069 (2004).

[47] J. N. Mercer, C. S. Chan, T. Tkatch, J. Held, and D. J. Surmeier, Nav1.6 sodium channels are critical to pacemaking and fast spiking in globus pallidus neurons, J. Neurosci. **27**, 13552 (2007).

[48] C. J. Wilson, Active decorrelation in the basal ganglia, Neuroscience **250**, 467 (2013).

[49] R. Gast, R. Gong, H. Schmidt, H. G. E. Meijer, and T. R. Knösche, On the role of arkypallidal and prototypical neurons for phase transitions in the external pallidum, J. Neurosci. **41**, 6673 (2021).

[50] V. Klinshov, S. Kirillov, and V. Nekorkin, Reduction of the collective dynamics of neural populations with realistic forms of heterogeneity, Phys. Rev. E **103**, L040302 (2021).

[51] M. V. Tsodyks and H. Markram, The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability, Proc. Natl. Acad. Sci. USA **94**, 719 (1997).