NEURAL ACTIVITY MEASURES AND THEIR DYNAMICS

ELI SHLIZERMAN, KONRAD SCHRODER AND J. NATHAN KUTZ*

Abstract. We provide an asymptotically justified derivation of activity measure evolution equations (AMEE) for a finite size neural network. The approach takes into account the dynamics for each isolated neuron in the network being modeled by a biophysical model, i.e. Hodgkin-Huxley equations or their reductions. By representing the interacting network as self and pairwise interactions, we propose a general definition of spatial projections of the network, called activity measures, that quantify the activity of a network. We show that the evolution equations that govern the dynamics of the activity measure shadow the activity measure of the network (i.e. the two quantities stay close to each other for all times) for general interactions and various asymptotic dynamics. The AMEE effectively serve as a dimensionality reduction technique for the complex network when spatial synchrony and coherence are present and allow to *a priori* predict network dynamics that would not be guessed from individual neuron behavior. To demonstrate an explicit derivation of such a reduction, we consider the mean measure for a network of interacting FitzHugh-Nagumo neurons. Computational results comparing the full network dynamics with the mean AMEE model of identical and nonidentical FitzHugh-Nagumo and FitzHugh-Rinzel neurons validate the shadowing theorems and expose the various resulting AMEE models that allow to describe the mean of the network.

Key words. Network of neurons, Mean-field approach, Dynamics of networks, Synchronization, Coherent structures, Neural activity

AMS subject classifications. 92C20,92B20,34D06

1. Introduction. The development of mean-field theories arose naturally as a consequence of the difficulty of solving many-body systems typically associated with quantum mechanics. During the development of many-body theories in the 1950s and 1960s, computation was not even an option for exploring the underlying physics given the enormous computational complexity. Thus mean-field theory essentially provided a *dimensionality reduction* technique whereby the complex interaction dynamics of individual particles or atoms were replaced by a single spatiotemporal field variable (or a low number of field variables) that was significantly more amenable to analytic study and computation. In the best case scenario, the dimensionality reduction is justified in some physically reasonable asymptotic limit [1, 2], including the dilute gas limit where an accurate description of experimental findings of Bose-Einstein condensates (BEC) dynamics holds [3]. Although modern computing has revolutionized the approach to such complex systems, the mean-field approach and its dimensionality reduction still plays a key role in identifying the underlying mechanisms and parameter regimes that drive the fundamental dynamics of the complex system that is often obfuscated in direct, large-scale simulations.

Just as in atomic systems, networks of interacting neurons present an ideal application of such dimensionality reduction techniques. The complexity of a large number of interacting neurons represents a tremendous challenge for resolving the functionality and overall activity of neural systems as the coupling of neurons completely alters the dynamics of the underlying neural network, see [4, 5, 6, 7, 8, 9] and references therein for examples of such networks, resulting complex dynamics and possible analysis. This has led to a significant effort towards constructing detailed phenomenological models of specific neural systems [6, 10]. Such qualitative models highlight key biological mechanisms responsible for producing various dynamical phenomena observed in neural systems. Further, they are mathematically warranted

---

* Department of Applied Mathematics, University of Washington, Seattle, WA, 98195-2420.

given the large number of neurons (10^{11} in the human brain) and their connections to each other (a typical neuron has 10^4 connections), thus rendering even computational approaches intractable [11, 12].

Concurrently to the development of particular models, mean-field methods based on generic, qualitative interaction rules among nodes of the network were introduced. Specifically, two dominant models have emerged: the *coupled-oscillator/phase-field reduction* model and the *neural field* model. The first approach treats neural networks as systems of coupled identical oscillators [13, 14, 15, 16, 17]. The method is based on the assumption that the individual identical oscillators (neurons) have a normally hyperbolic attracting limit cycle, which is guaranteed to persist, by Fenichel's theorem, under generic weak interactions, see [7]. Along the attracting limit cycle, the dynamics can be described by a scalar phase variable for each oscillator such that a reduction of the network equations onto these variables results with a coupled system of PDEs. Since the perturbed limit cycle is close to the unperturbed one these equations can be evaluated on the uncoupled limit cycle resulting with a coupled system of ODEs for the phases. Time averaging applied to this system results with effective equations for the differences between phases. For networks of conductance based neurons, whose uncoupled dynamics possess a hyperbolic attracting limit cycle, application of a nonlinear transformation, if available, will bring the uncoupled network to a simple oscillator with linear phase evolution and will allow one to analyze such networks using the phase-reduction method. An alternative reduction approach makes use of a statistical physics viewpoint of the neural activity of a population [18, 19, 20]. In this formulation, the dynamics of a single neuron are cast in terms of neural activity (e.g. membrane potential, refractoriness or firing rate). Under the assumption of close spatial proximity and random connections, a continuum limit of the network ($N \to \infty$) is derived. The neural activity density of the whole population is then described by a scalar field variable $u(x, t)$ at position x and time t denoted as the *neural field* [21, 22, 19]. The dynamics of neural fields are usually modeled by a nonlocal linear PDE in which the linear term corresponds to the self neural activity (local population) and the nonlocal term corresponds to the interactions between the neural activities of the populations. Studies of these models reveal key spatio-temporal pattern formation in a large population given various interactions including localization [23, 24], spiral waves [25], various geometrical shapes [26], and traveling waves [27]. In addition to the deterministic approach, mean-field methods were developed for noisy populations of neurons, where each individual neuron in the population is approximately modeled as an independent Poisson process. Various methods allow to describe the dynamics on the macro-scale level, deriving equations for the probability distribution of the neural activity [28] or moments of the distribution [29, 20].

While the above described approaches result with qualitative mean-field models that demonstrate the general dynamical phenomenology, their relation to specific finite size networks of neurons is not immediate. Indeed, the phase-field approach assumes the neurons produce oscillatory behavior as opposed to the ubiquitous spiking behavior. Likewise, the neural field approach casts the entire dynamics in terms of a proxy (firing rates) and no longer retains information on the dynamics of the neuron itself. In general, two quantities characterize the dimensionally reduced mean-field: the individual neuron dynamics described by a nonlinear conductance based model (i.e. Hodgin-Huxley, FitzHugh-Nagumo, etc.) and their interactions (i.e. synaptic, gap, ephaptic, etc.). In a neural network architecture, the dynamics and bifurcation structure of the mean-field results in complex, nontrivial and potentially unexpected

dynamics. As an illustrative example, Fig. 1 considers three standard reductions of the Hodgkin-Huxley equations, all of which have nearly identical individual neuron response. Under identical interactions, their mean-fields produce markedly different behavior, including chaotic bursting, relaxation oscillations and sub-threshold oscillations. The simple example considered here clearly demonstrates that an intuitive guess based upon the individual neuron dynamics is insufficient to describe the observed behavior. The aim of this manuscript is to *a priori* derive the mean-field equations by combining the individual neuron dynamics and their interactions, thus providing a theoretical framework for understanding the diversity of behavior illustrated, for example, in Fig. 1.

The method developed here follows the philosophical principle of N-body quantum mechanics which dictates that individual atoms behave according to the first principle Schrödinger equation [1, 2, 3]. The approach can be viewed as a microscale to macroscale, i.e. a bottom-up approach, in which the individual neuron dynamics and their couplings are explicitly accounted for in the description of the dynamics of a spatial projection of the network denoted as the activity measure. An underlying assumption in the derivation is that there exists an underlying attractor for each neuron when it is a part of the network consortium and leads to the mean-filed equations termed as Activity Measure Evolution Equations (AMEE). For specific measures and attractors we show that these equations shadow, as defined in [30], the activity measure of the network. The AMEE equations are adjoint to the network and describe the dynamics of the activity measure. A reduction of the AMEE is achieved when spatial synchrony or coherence in the attractor dynamics is present such that the reduced equations serve as a low dimensional model for the dynamics on the attractor. Although the method is constructed for a network in which each node is a neuron, the method developed herein can be applied to any network of interacting nodes with prescribed dynamics.

In summary, to derive the bottom-up mean-field equations we accomplish the following steps: **(i)** Present the AMEE as an evolution equation for an empirical measure of an attracting orbit in a given nonlinear dynamical system. **(ii)** Assume that the dynamical system has an attracting set (attractor) and consider perturbations around orbits that belong to the attracting set. **(iii)** Present and prove "shadowing" results about how the AMEE is applicable in a neighborhood of the attractor. **(iv)** Discuss the applicability of dimension reduction to specific attractors. **(v)** Derive the AMEE for nonlinear dynamical systems that model networks of neurons. **(vi)** Do computations in specific cases that apply to these networks.

2. Construction of a network of interacting neurons. We formulate the network of interacting neurons by defining the self- and pairwise-interactions of the neurons. These assumptions are similar to the assumptions in the various models introduced for neural networks as described in [6, 7] and references therein. The self dynamics of an i-th neuron are described by an autonomous ordinary differential equation

$$\dot{V}_i = f_i(V_i) \tag{2.1}$$

where $V_i \in \mathbb{R}^n$ is a vector valued function of t, and $f_i : U \to \mathbb{R}^n$, is a smooth function for $U \subseteq \mathbb{R}^n$. This definition includes various descriptive models for neuron dynamics: the Hodgkin-Huxley (HH) model and reductions of the HH model, such as Morris-Lecar (ML) and FitzHugh-Nagumo (FHN) (for additional models see [11, 15]). Indeed, it implies that there are two indices involved: index i that enumerates the

A. Individual Dynamics **B. Network Dynamics**
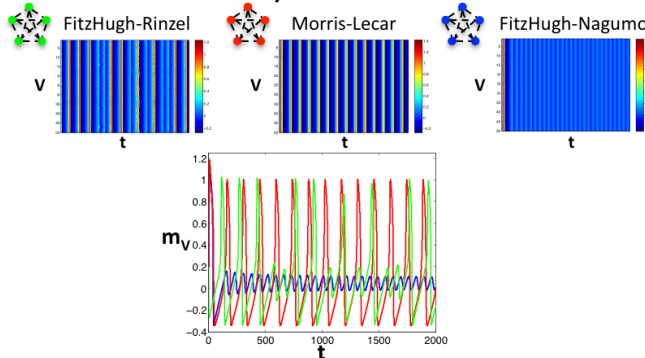
FitzHugh-Rinzel   Morris-Lecar   FitzHugh-Nagumo
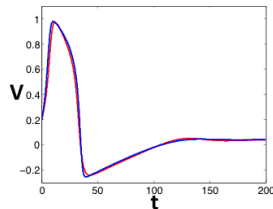
FIG. 1. *Individual dynamics vs. population dynamics.* **A:** *Three different models, FitzHugh-Rinzel (green), Morris-Lecar (red) and FitzHugh-Nagumo (blue) are fitted to produce a similar, once spiking trajectory for a given initial condition.* **B, top:** *Colored raster plots of the three different models placed into an all-to-all network with linear connections in the voltage variable. The initial condition for the network is chosen to be random perturbations around the fitted initial condition of A.* **B, bottom:** *The dynamics of the mean of each network; should be compared with A. See §2 for the definition of the individual models and §5 and §6 for analytic and computational description of synchronous network behavior.*

neurons ($i = 1, \ldots, N$) and index k that enumerates the dynamical variables that model each neuron ($k = 1, \ldots, n$). Using this notation Eq. (2.1) can be written as

$$\dot{v}_{ik} = f_{i,k}(v_{i1}, v_{i2}, \ldots, v_{in}).$$

In this notation the variables v_{ik} are represented as elements of a matrix $N \times n$. Rows of the matrix are copies of neurons and the columns are dynamical variables. The vectorfield $f_{i,k}$ is also a matrix $N \times n$ and defines the interactions between elements within each row. We distinguish between the variables and the elements of the vectorfield by using comma separating the indices of the vectorfield.

For the HH model $n = 4$, and the quadruple of $(v_{i1}, v_{i2}, v_{i3}, v_{i4}) = (v_i, n_i, m_i, h_i)$ are the gating variables that correspond to voltage, activation of potassium, activation and inactivation of sodium currents respectively. The FHN model will be of a special interest in this paper. For $V_i = (v_i, w_i)^T$ it is expressed by the following two dimensional nonlinear ODEs with parameters a, b, c

$$\dot{V}_i = f^{FHN}(V_i; a, b, c) = \begin{pmatrix} 0 & -1 \\ -c & b \end{pmatrix} V_i + \begin{pmatrix} f_1(v_i, a) \\ 0 \end{pmatrix}, \quad f_1(v_i, a) = v_i(v_i - a)(1 - v_i).$$

$$(2.2)$$

where the first term in the right hand side of the equation is linear and the second one is cubic. The variable $v_i(t)$ corresponds to the voltage of the neuron and the variable $w_i(t)$ is the recovery variable. The equation for v_i mimics the HH conductance equation. Unlike the HH equations, in which the conductance equation is coupled to three additional auxiliary variables, the FHN is a simplification such that it is coupled linearly to only one auxiliary variable w_i. The FHN model is of particular interest since it reproduces several important features of the HH equations, such as creation of action-potentials once the initial condition (initial stimulus) have passed a threshold, bi-stability and periodic solutions. Since this is a planar model, the phase plane study provides analytic parameters regimes at which each of these behavior

4

occurs, see [31, 11]. Furthermore, $f_1(v)$ is a cubic polynomial and it agrees with the assumption that $f(V_i)$ is smooth. Another plausible planar model for self interactions is the ML model [15, 32]

$$\dot{V}_i = f^{ML}(v_i, w_i; p) = \begin{pmatrix} \dot{v}_i \\ \dot{w}_i \end{pmatrix} = \begin{pmatrix} -g_{\mathrm{L}}(v_i - v_{\mathrm{L}}) - g_{\mathrm{Ca}} m_\infty(v_i)(v_i - v_{\mathrm{Ca}}) - g_{\mathrm{K}} w_i(v_i - v_{\mathrm{K}}) \\ \frac{1}{\tau}(w_\infty(v_i) - w_i) \end{pmatrix}$$
(2.3)

$$\text{where } m_\infty(v) = \frac{1}{1 + e^{(v_m - v)/\kappa_m}}, \ \ w_\infty(v) = \frac{1}{1 + e^{(v_w - v)/\kappa_w}}$$

where p denotes the vector of the parameters in the system. In this model the two components of the vectorfield are nonlinear and non-polynomial as expressed by the functions $m_\infty(v)$ and $w_\infty(v)$. Although qualitatively the phase planes of the ML and the FHN models are similar, their codimension 1 bifurcation structure is different with the DC input as a free parameter. For further details see [11, 15]. The ML and FHN models are used in Fig. 1 to demonstrate the importance of the choice of the individual neuron model and the effects of different bifurcation structure on the network dynamics.

When additional variables that model the slow currents are introduced, they extend the planar dynamics to three dimensional phase space or higher. For example, adding the x_i variable to the FHN model extends it such that it is able to produce elliptic bursting solutions. This model extension is commonly known as the FitzHugh-Rinzel model (FHR) [33]

$$\dot{V}_i = f^{FHR}(v_i, w_i; p) = \begin{pmatrix} \dot{v}_i \\ \dot{w}_i \\ \dot{x}_i \end{pmatrix} = \begin{pmatrix} v_i(v_i - a)(1 - v_i) - w_i + x_i \\ -cv_i + bw_i \\ \varepsilon(-v_i + d - x_i) \end{pmatrix}.$$
(2.4)

Self interaction of the neurons can be also modeled by the integrate and fire models of different types. In such a case, the dynamics of f as given by (2.1) incorporates a threshold rule. The derivation of the AMEE for such a model of a neuron should be similar to the derivation presented here, when the threshold for such AMEE model should be adjusted by other methods/techniques or specified a priori [34, 35, 36, 37].

The interactions between the neurons in a network will be modeled by additional pairwise terms in the vectorfield, such that if we would like to model all interactions of the neuron V_i with all other neurons, it will be the sum of the neurons connected to V_i

$$\dot{V}_i = \sum_{j=1, j \neq i}^{N} e_{i,j}(V_i, V_j, t).$$
(2.5)

The assumption that the interactions are pairwise is in general not necessary. The adjoint evolution equations (AMEE) can be still derived for fairly general expression of the interactions. However, for the purpose of the analysis it is convenient to assume pairwise interaction terms that simplify both the structure of the network, and allow for simple expansion in terms of power series. Having formulated both the neuron's self and pairwise interactions, the general evolution equation for the network of interacting neurons can be constructed. Let $\mathbb{V}(t) \in (\mathbb{R}^N)^n$ be the vector of all variables that describe the states of all the neurons at a time t, i.e. $\mathbb{V}(t) = (V_1(t), V_2(t), \ldots, V_N(t))^T$. Then the full system of equations for the network is

$$\dot{\mathbb{V}} = G(\mathbb{V}, t), \qquad \mathbb{V}(0) = \mathbb{V}_0$$
(2.6)

where $G: U \to (\mathbb{R}^N)^n$, is a smooth function for $U \subseteq (\mathbb{R}^N)^n$ that consists of self and pairwise interaction terms, such that for each component v_{ik} of $\mathbb{V}$, its dynamics are described by

$$\dot{v}_{ik} = g_{i,k} = f_{i,k}(v_{i1}, v_{i2}, \ldots, v_{in}) + \sum_{j=1, j \neq i}^{N} e_{i,j,k}(v_{i1}, v_{i2}, \ldots, v_{in}, v_{j1}, v_{j2}, \ldots, v_{jn}, t).$$
(2.7)

The corresponding flow operator of the full system is defined as $\mathbb{V}(t) = \varphi^t(\mathbb{V}_0)$, see [30] for a definition. The full system of ODEs is assumed to have local existence and uniqueness of solutions, such that G is locally Lipshitz continuous. Such assumption includes piecewise continuous models for $f_{i,k}$ as well. Further, assuming that self and pairwise interactions, $f_{i,k}$ and $e_{i,j,k}$, are analytic and autonomous and that self interaction can be expressed only by pairwise interactions it is possible to approximate Eq. (2.7) as a finite sum of pairwise monomial terms which can be unified into one nested sum on the right hand side

$$\dot{v}_{ik} = \sum_{p'=1}^{p_1} \sum_{p''=1}^{p_2} \sum_{l=1}^{n} \sum_{m=1}^{n} \sum_{j=1}^{N} a_{ilmj}^{(p',p'')} (v_{il})^{p'} (v_{jm})^{p''},$$
(2.8)

where a denotes the coefficient of the pairwise monomials for self interactions, pairwise interactions and combined (both self and pairwise) interactions. Such a notation will not include the HH model since it has a triple interaction term (vm^3h), however it will include various two dimensional and three dimensional approximations to HH. The notation can be similarly extended to include the triple interactions terms.

3. Evolution of the measurable activity. Instead of solving for the entire network and resolving the activity of each neuron, one can consider a model of the activity of a population of neurons that is being measured, i.e. the *measurable activity*. This is effectively a spatial projection of the network dynamics. Such a model is constructed to approximate the dynamics of the activity of the whole network or part of the network. Being able to overcome the complexity of large networks and providing a succinct system of ODEs that describes the evolution of the measurable activity can simplify the analytic and numerical study of the dynamical properties of the functionality of the network. As a result, such a model will serve as a dimensionality reduction for the complex network. In many theoretical and experimental neural networks, the measured quantity is the averaged activity and hence the spatial mean is a natural choice for characterizing the measurable activity. Therefore we will consider the mean as a particular case of the measurable activity.

To construct the evolution equations for the measurable activity we make the following assumptions on the network structure and dynamics. First, the flow operator of the interacting network described by the system in Eq. (2.6) has a closed invariant, indecomposable set $\mathbb{X} \in (\mathbb{R}^N)^n$ and there exists a neighborhood U of positive Lebesgue measure such that $u \in U$ implies that the ω-limit set of u (the asymptotic flow of u, $\varphi^t(u)$, when $t \to \infty$) is contained in $\mathbb{X}$ and the forward orbit, $\varphi^t(u)$ from u is contained in U. This assumption implies that there exists an **attractor**, defined similarly to the definition in [30, §5.4], denoted by the set $\mathbb{X}$, in the network whose dynamics are described by Eq. (2.6). Further, the basin of attraction of the attractor, the set U, is of non zero measure. The attractor is not necessarily global and there may exist several local attractors with distinct basins of attractions. We parameterize the flow on the

6

attractor with the time t, such that $\varphi^t(\mathbb{X}(0))$ is the flow defined by G from a point $\mathbb{X}(0)$ to another point $\mathbb{X}(t)$. The invariance of $\mathbb{X}$ ensures that for any t, $\varphi^t(\mathbb{X}(0)) \in \mathbb{X}$. The attractor can be chaotic if it contains transversal homoclinic orbits.

To be more precise, the *activity measure* of the asymptotic dynamics is a mapping $M^t : \mathbb{X} \to \mathbb{R}^n$ that associates a number with N copies of each variable x_k, i.e.

$$M^t(\mathbb{X}) = (m_1(x_1(t)), \dots, m_n(x_n(t)))^T, \qquad m_k(x_k(t)) = \sum_{i=1}^{N} p_{ik}(x_{ik}(t)). \qquad (3.1)$$

The superscript t of M denotes that the activity measure implicitly depends on the parameterization t through $\mathbb{X}(t)$. The above definition arises from various statistical spatial measures that can be applied to the network to study its dynamics. The real valued smooth function $p_{ik}(x_{ik})$ intuitively represents the weight of each x_{ik} in the summation and can be nonlinear in x_{ik}. Here we choose a rather general definition to account for various measures that have unequal weights and functions for the variables both in k and in i. Usually in applications much simpler measures are chosen. For example, in many applications one of the frequently measured quantities is the mean activity or weighted mean activity where the weights w_{ik} are drawn from a probability distribution $\vec{w}_k \in \mathbb{R}^N$. Then the function $p(x_{ik})$ takes the form

$$p(x_{ik}) = x_{ik}(t) \, w_{ik}.$$

One can similarly define the measure to be the r-th central moment $p(x_{ik}) = x_{ik}^r(t) w_{ik}$, or other nonlinear function of x_{ik}. Notice that due to the finite number of neurons in the network, the measure is discrete in i. Indeed the activity measure is the projection of the dynamics on the attractor from $(\mathbb{R}^N)^n$ to $\mathbb{R}^n$, formulating the same dynamics but in a dual space of the measure, where usually in neural networks $N \gg n$. A successful choice of a measure is such that it captures the low-dimensional structure of the network. Similarly, an infinite version, when $N \to \infty$, can be formulated replacing a sum with an integral. Since we consider here networks of finite size we will use the finite version of the definition. Notice that in contrast to other methods where the limit of $N \to \infty$ is required, here N is left as a free parameter and therefore the effects of the individual neuron dynamics will play a significant role. In general, m_k is not restricted to be nonnegative due to the possibility of different signs of the measured variables x_{ik} or the weights w_{ik}, and thus the activity measure is a signed measure, i.e. has the additivity property. Indeed, in many neural systems the weights are taken to be signed such that positive weights correspond to excitatory neurons and negative for inhibitory neurons [29, 38]. The additivity property allows one to decompose the activity measure into the excitatory and inhibitory parts while the overall activity measure reflects the balance of the two populations; see §4.2 for a general decomposition of the activity measure. For particular activity measures such as even central moments with positive weights w_{ik}, the activity measure is a measure and possesses additional properties that result from non-negativity such as monotonicity.

Equipped with the definition of the measure of the attractor, we proceed to derive the adjoint evolution equations for the activity measure - the AMEE model. Specifically, the evolution of the activity measure is described by

$$\frac{dM}{dt} = \left(\sum_{i=1}^{N} g_{i,1} \frac{dp_{i1}}{dx_{i1}}, \dots, \sum_{i=1}^{N} g_{i,k} \frac{dp_{ik}}{dx_{ik}}, \dots, \sum_{l=1}^{N} g_{i,n} \frac{dp_{in}}{dx_{in}} \right)^T, \qquad M(0) = M^0(\mathbb{X}(0)).$$

$$(3.2)$$

where $g_{i,k}$ is the (i,k)-th component of the vectorfield G. This can be shown by considering the activity measure vector $M(t)$ and taking derivative with respect to t from both sides and using the chain rule. Note that for analytic $g_{i,k}$, approximated by Eq. (2.8), the adjoint equation has the form of

$$\frac{dm_k}{dt} = \sum_{p'=1}^{p_1} \sum_{p''=1}^{p_2} \sum_{l=1}^{n} \sum_{m=1}^{n} \sum_{j=1}^{N} \sum_{i=1}^{N} \frac{dp_{ik}}{dx_{ik}} a_{ilmj}^{(p',p'')} (x_{il})^{p'} (x_{jm})^{p''}.$$

When additionally p_{ik} is analytic, the approximate adjoint evolution equation can be expressed solely as a sum of pairwise monomials.

Similarly we can define the activity measure for trajectories that are approaching an attractor. When a trajectory is ε-close to the attractor for all $t > t^*$, in all variables, i.e. we can find a trajectory on the attractor $x_{ik}^A(t)$ such that $v_{ik}(t) = x_{ik}^A(t) + \varepsilon \xi_{ik}(t)$ with $\xi_{ik}(t) \sim \mathcal{O}(1)$. Such a perturbed trajectory $v_{ik}(t)$ is said to ε-shadow the trajectory $x_{ik}^A(t)$. The AMEE can be defined by expanding in ε (up to second order)

$$\frac{dm_k}{dt} = \sum_{i=1}^{N} \frac{dp_{ik}}{dx_{ik}}(x_{ik}^A) g_{i,k}(x_{qr}^A) +$$
$$\varepsilon \sum_{i=1}^{N} \left(\frac{dp_{ik}}{dx_{ik}}(x_{ik}^A) \sum_{j=1}^{N} \sum_{l=1}^{n} \xi_{jl} \frac{dg_{i,k}}{dx_{jl}}(x_{qr}^A) + \xi_{ik} \frac{d^2 p_{ik}}{d^2 x_{ik}}(x_{ik}^A) g_{i,k}(x_{qr}^A) \right). \quad (3.3)$$

The zero order right hand side (first row of Eq. (3.3)) corresponds to choosing for Eq. (3.2) an initial condition as a point on the attractor $x_{qr}^A(t^*)$ closest in terms of an Euclidian distance to $v_{qr}(t^*)$. Note that we use different indices for the arguments of $g_{i,k}$ and p_{ik} and their derivatives to distinguish the domains of the two functions. The vectorfield $g_{i,k}$ acts on all the coordinates (denoted by indices q and r) while the projection p_{ik} acts on a specific coordinate (denoted by i and k). See Appendix A for a detailed derivation.

We can summarize the closeness of the perturbed and the unperturbed solutions of the adjoint equations in the following theorem. The theorem is the primary analytic underpinning of the current work. Its proof is included in Appendix B. Notice that the same derivation as above will be correct if the solution x_{ik} will be replaced by v_{ik}, a transient solution which is not a part of the attractor. However, in such a case we will not be able to assure that the perturbed adjoint evolution equations is close to the unperturbed one. As a result it will not be expected to be generic for nearby initial conditions or perturbations.

THEOREM 1 (Shadowing). *Consider a perturbed trajectory $v_{ik}(t) = x_{ik}^A(t) + \varepsilon \xi_{ik}(t)$ for $t \geq t^*$ with $\xi_{ik}(t) \sim \mathcal{O}(1), \varepsilon \ll 1$ and $x_{ik}^A(t)$ is a trajectory on the attractor $\mathbb{X}$.*
***(a):** If the measure $p(x_{ik})$ is a linear measure, $p(x_{ik}) = w_{ik} x_{ik}$, with normalized weights $\sum_{i}^{N} |w_{ik}| = 1$, then $m(v_{ik})$ ε-shadows $m(x_{ik}^A)$.*
***(b):** If the measure $p(x_{ik})$ is (nonlinear) analytic and the attractor $\mathbb{X}$ is compact, then $m(v_{ik}) - m(x_{ik}^A)$ is bounded ($m(v_{ik})$ shadows $m(x_{ik}^A)$). If additionally the convergence rate is fast enough such that for all $t > t^*$, $|\xi_{ik}(t)| < 1/\left(N p_{ik}^{(1)}(x_{ik}^A(t)) \right)$ and $|\xi_{ik}(t)|^r < r!/\left(N \varepsilon^{r-2} p_{ik}^{(r)}(x_{ik}^A(t)) \right)$ then $m(v_{ik})$ ε-shadows $m(x_{ik}^A)$ and the derivation*

of Eq. (3.3) *holds.*

We conclude that the evolution equations in Eq. (3.3) can be treated as the adjoint equation for the trajectory on the attractor and denote the asymptotic dynamics of the activity measure, a spatial projection of the network. It retains its structure both from the full network and the expression of the measure. Being the adjoint equation means that to solve the equation (find the solution $m_k(t)$ at time t), the trajectory on the attractor from $X(0)$ to $X(t)$ (or close by trajectory) should be given. Therefore, Eqs. (3.2) or (3.3) as-is, do not simplify the characterization of the network dynamics, since first we still need to solve Eq. (2.6) and for the purpose of finding $M(t)$ at each time t we could use Eq. (3.1).

4. Dimension reduction of the AMEE. In particular cases of the measure M and the vectorfield G, it is possible to formulate the adjoint equations (or parts of it) as independent equations of the attractor trajectory, only in terms of m, hence providing the effective equation for the dynamics of the evolution. For example, if G is linear and all p_{ik} are linear (which corresponds to the mean) the AMEE can be formulated as an independent equations of variables $m_1, \ldots, m_n$. Such an equation is sometimes referred to as a linear mean-field model. When p_{ik} is nonlinear and a monomial in x_{ik}, only the terms in $g_{i,k}$ that are linear in x_{ik} can be formulated in terms of m_k. A knowledge of the dynamics of some of the variables on the attractor can help in designing a measure that will result with an independent equation. However, it again requires us to resolve some part of the dynamics in order to find a suitable measure. This is, in general, a difficult task, however some generic approaches for dimension reduction are possible as we describe next.

4.1. Synchronous attractor. Formulating a genuine low dimensional adjoint equations is possible for the case when the dynamics of a subset of variables is similar, such that they can be related via simple scaling relation. Thus after rescaling, the solutions are identical up to small fluctuations. We formulate this case as the "synchronous" state assumption. Such type of behavior has been observed in networks of general oscillators and defined as the sync state [39, 40]. Existence of a synchronous attractor was shown to exist in general network architectures with sufficiently strong connections. Examples include a network of oscillators with linear connections [41], neural piecewise linear models connected by gap junctions [42, 43] and in general models with gap junctions interactions that obey the semi-diffusivity property [44]. Here, Fig. 1 serves as an example for various synchronous behavior.

In particular, for fixed k, a subset $\{v_{ik}\}$ of $\tilde{N}$ variables, where $\tilde{N} \leq N$, is called *synchronous* if each of them can be expressed as

$$v_{ik}(t) = \nu_i \bar{x}_k(t) + \varepsilon \xi_{ik}(t) \tag{4.1}$$

where $\bar{x}_k(t) \in \mathbb{R}$ is the synchronized solution and ε is a small parameter such that $\varepsilon \ll \nu_i$ and $\bar{x}_k$ and ξ_{ik} are of $O(1)$. For a completely synchronous network, such that for all k the synchronized subset is the set of all N variables, $\tilde{N} = N$, and $\bar{x}_k$ belongs to the attractor, the AMEE up to second order in ε is given by

$$\frac{dm_k}{dt} = \sum_{i=1}^{N} \frac{dp_{ik}}{dx_{ik}}(\bar{x}) g_{i,k}(\bar{x}) + \varepsilon \sum_{i=1}^{N} \left(\frac{dp_{ik}}{dx_{ik}}(\bar{x}) \sum_{j=1}^{N} \sum_{l=1}^{n} \xi_{jl} \frac{dg_{i,k}}{dx_{jl}}(\bar{x}) + \xi_{ik} \frac{d^2 p_{ik}}{d^2 x_{ik}}(\bar{x}) g_{i,k}(\bar{x}) \right). \tag{4.2}$$

where $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$. In this special case, the adjoint evolution equation for a linear measure (the weighted mean $m_k(x_k) = \sum_{i=1}^{N} \nu_i x_{ik}$) results in a simplified equation for the leading order and its correction

$$\frac{dm_k(\bar{x})}{dt} = \sum_{i=1}^{N} \nu_i \ g_{i,k}(\bar{x}) \tag{4.3}$$

$$\frac{dm_k(\xi_k)}{dt} = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{l=1}^{n} \nu_i \xi_{jl} \frac{dg_{i,k}}{dx_{jl}}(\bar{x}) \tag{4.4}$$

This result indicates that for a network of identical neurons that has a synchronized solution $\bar{x}_k$, the evolution of the mean measure is inherited from the structure of the network $g_{i,k}$. The mean of the fluctuations around the solution will inherit its structure from the Jacobian of $g_{i,k}$. The structure of Eqs. (4.2), i.e. the nonlinear terms of the effective vectorfield and the resulting symmetries carried from the topology of the interactions in the network, provide a classification of the possible dynamics of the synchronous attractor without solving the network. When parameters are being varied the classification will result with particular bifurcation sequences. See §5 and §6 for an example of a study for the FHN and FHR networks of identical neurons. For more complex symmetries of the vectorfield the structural study in Refs. [45, 46, 47] can reveal the possible bifurcations in the resulting AMEE model, despite the AMEE itself not being derived.

4.2. Synchronous subregions. Several neuronal networks possess the structure of distinct levels / subregions of neurons interacting with each other. Each subregion is responsible for a specific functionality. Within each subregion there are similar neurons connected densely to each other. The interaction between neurons that belong to different subregions can be modeled by connecting only a subset of neurons (output neurons) or dense connections but with weaker strength than the connections within a subregion. The connections are usually directed and model inhibitory/excitatory interactions between the regions. Examples of such networks include interactions of the amacrine and the ganglion cells in the retina [48, 49], hypercolumns in the visual cortex [50, 51] and glomeruli interactions in the olfactory bulb [52, 53].

For such networks, existence of a complete synchronous attractor is usually replaced by the existence of (spatial) multiple attractors associated with each distinct neural subregion. Resolving the functionality of each subregion, i.e. the dynamics of an activity measure for each attractor, can reveal how signals propagate through different subregions [54], mechanisms of formation of waves [48, 55], delays and advances in neural activity, see §7.2 and [29]. On the global level, resolution of the total activity measure can provide information on the presence and absence of balance between the regions.

Existence of multiple synchronous attractors results with an extension of the dimension reduction for a complete synchronous attractor in §4.1. For each region $j = 1, \ldots, K$, where K is the number of subregions, the total activity measure is the sum of activity measures m_k^j taken over each subregion, where $\tilde{N}_j$ is the number of nodes in subregion j

$$m_k = m_k^1 + \ldots + m_k^K = \sum_{i=1}^{\tilde{N}_1} p_{ik}(x_{ik}(t)) + \ldots + \sum_{i=1}^{\tilde{N}_K} p_{ik}(x_{ik}(t)).$$

Assuming that for each subregion there is a corresponding attractor $v_{ik}^j = \bar{x}_k^j + \varepsilon \xi_{ik}^j$ the evolution equations for m_k^j, that are an extension to an AMEE in Eq. (4.2) are derived

$$\frac{dm_k^j}{dt} = \sum_{i=1}^{\tilde{N}_j} \frac{dp_{ik}}{dx_{ik}}(\bar{x}^j) g_{i,k}(\bar{x}^1, \ldots, \bar{x}^K).$$

These equations are the $\mathcal{O}(1)$ terms and similarly AMEE in Eq. (4.2) are extended for $\mathcal{O}(\varepsilon)$ terms. The construction of the network in §2 imposes that each node dynamics are governed by a sum of vectorfields of individual terms and pairwise interaction terms and thus allows to write the $g_{i,k}$ term in the above equation as a sum of terms

$$\frac{dm_k^j}{dt} = \sum_{i=1}^{\tilde{N}_j} \frac{dp_{ik}}{dx_{ik}}(\bar{x}^j)\left(g_{i,k}^j(\bar{x}^j) + e_{i,k}(\bar{x}^j, \bar{x}^1) + \ldots + e_{i,k}(\bar{x}^j, \bar{x}^K)\right).$$

The first term in the parentheses corresponds to individual terms and interactions within the j-th subregion. The remaining terms correspond to effective interactions between the distinct regions. In summary, in the presence of multiple spatial attractor the AMEE can be reduced to nK coupled equations whose solution will shadow the attractor in the full network. In practice, dynamics of a network composed of several regions may exhibit complex spatio-temporal behavior, while the reduced AMEE equations will reveal the mechanism (coupling terms and coefficients) that is responsible for such an attractor and the variety of attractors that may exist in such a network. Furthermore, resolution of $m_k^j(t)$'s implies the resolution of the overall activity measure $m_k(t)$. Here again the properties of the coupled AMEE equations such as symmetries between the $m_k^j(t)$'s will result with restricted overall activity measure such as balanced or unbalanced activity and can be determined from the structural properties of the resulting equations [45]. In §7.2 we show an example of FHN network with two distinct populations where one population exhibits relaxation oscillations in delay to another one.

4.3. Coherent structures. When the dynamics of the network do not result with a synchronous attractor but with coherent structures, and the network is not clustered into subregions, an alternative approach for dimension reduction is to assume that the attractor is composed of $K \leq N$ coherent structures, or orthogonal modes, see [56, 57, 58]. In the network formalism as defined in §2, the trajectory of the k-th variable on the attractor $\vec{x}_k^A(t) = (x_{1k}^A, \ldots, x_{Nk}^A)$, in a vector notation, can be decomposed into

$$\vec{x}_k^A(t) = \sum_{j=1}^K r_k^j(t)\vec{y}_k^j, \quad \text{such that} \quad <\vec{y}_k^j, \vec{y}_k^j> = 1, \quad \text{and} \quad <\vec{y}_k^j, \vec{y}_k^l> = 0 \qquad (4.5)$$

where $< \cdot, \cdot >$ is the dot product of two vectors. Then the AMEE are written as

$$\frac{dm_k}{dt} = \sum_{i=1}^N \frac{dp_{ik}}{dx_{ik}} \sum_{j=1}^K \dot{r}_k^j(t)\vec{y}_k^j \qquad (4.6)$$

and effectively represent the equations for the activity measure of the coefficients of the orthogonal modes. The equations for the coefficients are obtained by substitution of the expansion $\vec{x}_k^A(t) = \sum_{j=1}^K r_k^j(t)\vec{y}_k^j$ into the vectorfield G such that its k-th

11

component is expressed as

$$\dot{\vec{x}}_k = \vec{g}_k(x_{qr}, t) = \vec{g}_k(r_l^j \vec{y}_l^j, t), \quad l = 1, ..., n, \quad j = 1, ..., K.$$

Taking the dot product with each $\vec{y}_k^j$ and using the orthogonality in Eq. (4.5), dynamical equations for the coefficients are received

$$\dot{r}_k^j = \; < \vec{g}_k(r_l^j \vec{y}_l^j, t), \vec{y}_k^j > \; .$$

If g is analytic, the right hand side can be expanded in power series and will take the form of polynomials multiplication of the variables r_l^j. The coefficients of these polynomials will be the inner product of the same polynomial multiplication of the orthogonal modes associated with the variables r_l^j and $\vec{y}_k^j$. For an example of three coherent patterns dynamics in a network with gap junctions interaction, see §8.1.

5. AMEE for a synchronous attractor in a network of identical FHN neurons. Using the formulation of §3, we derive the AMEE for the mean measure, for an explicit example network. This network is described by the FHN model with pairwise interactions

$$\dot{V}_i = f^{FHN}(V_i; a_i, b_i, c_i) + \sum_{j=1, j \neq i}^{N} \sum_{m=1}^{p_1} \sum_{n=1}^{p_2} a_{ijmn} v_i^m v_j^n. \tag{5.1}$$

The uncoupled system of ODEs, representing only self interactions, consists of identical neurons

$$f^{FHN}(V_i; \alpha) = \begin{cases} a_3 v_i^3 + a_2 v_i^2 + a_1 v_i - w_i + I_{i,ext}(t), \\ bw_i - cv_i. \end{cases}$$

where $\alpha = \{a_1, a_2, a_3, b, c\}$ is a set of fixed parameters independent of i. The interaction between the neurons will be pairwise, such that it indeed can be formulated as in Eq. (2.5). To simplify the interaction terms, the interaction will be assumed to be only via the voltage variable, i.e. the first components of $\vec{V}_i = (v_i, w_i)$ and $\vec{V}_j = (v_j, w_j)$, as in Eq. (5.1). To proceed and simplify the above system of equations we assume that all the neurons eventually will have similar common dynamics, i.e. we assume that there is a synchronous attractor in the full network, as discussed in §4.1:

The interacting network is locked into a common state ("synchronous" attractor) such that the dynamics of each neuron can be approximated by the attractor solution and fluctuations around it $V_i(t) = \begin{bmatrix} \bar{v}(t) \\ \bar{w}(t) \end{bmatrix} + \varepsilon \begin{bmatrix} \xi_{iv}(t) \\ \xi_{iw}(t) \end{bmatrix}$ where $\bar{V}(t) \in \mathbb{R}^n$ is the synchronized solution and ε is a small parameter. To derive the AMEE model we use Eq. (4.3) with the uniform mean,

$$m_v(v) = \frac{1}{N} \sum_{i=1}^{N} v_i, \; m_w(w) = \frac{1}{N} \sum_{i=1}^{N} w_i, \; m_v(\xi_v) = \frac{1}{N} \sum_{i=1}^{N} \xi_{iv}, \; m_w(\xi_w) = \frac{1}{N} \sum_{i=1}^{N} \xi_{iw} \tag{5.2}$$

which results in the following evolution system for the mean over the attractor and average fluctuations.

The averaged model for a network of the type of FHN interacting neurons that satisfy the above assumptions is given by

$$O(1): \dot{\bar{v}}(t) = -\bar{w} + \frac{1}{N} \sum_{m=0}^{p_1} \sum_{n=0}^{p_2} \left[A \right]_{mn} \bar{v}^{m+n}, \qquad \dot{\bar{w}}(t) = -c\bar{v} + b\bar{w} \qquad (5.3)$$

$$O(\varepsilon): \dot{\bar{\xi}}_v(t) = -\bar{\xi}_w + \frac{1}{N} \sum_{m=0}^{p_1} \sum_{n=0}^{p_2} \left(m \left[A\vec{\xi}_v \right]\bigg|_{mn} + n \left| A\vec{\xi}_v \right]_{mn} \right) \bar{v}^{m+n-1}, \; \dot{\bar{\xi}}_w(t) = -c\bar{\xi}_v + b\bar{\xi}_w$$

$$(5.4)$$

where $[A] \in \mathbb{R}$ is the sum of all elements in a matrix $\{a_{ij}\}$ such that

$$\left[A \right]_{mn} = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ijmn} = \vec{1} A_{mn} \vec{1}^T, \quad \text{where} \quad \vec{1} = (1, 1, ..., 1),$$

and $[A\vec{z}] \in \mathbb{R}$ and $|A\vec{z}| \in \mathbb{R}$ are

$$\left[A\vec{z} \right]\bigg|_{mn} = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ijmn} z_i \text{ and } \left| A\vec{z} \right]_{mn} = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ijmn} z_j \text{ where } \vec{z} = (z_1, z_2, \ldots, z_N).$$

The derivation of this averaged model is provided in Appendix C. Note that the resulting model in Eq. (5.3) is similar to the model that describes the self dynamics of each neuron with additional effective terms due to pairwise interactions, as expected from Eq. (3.3). These terms are in general nonlinear, for $p_1 + p_2 > 1$, expressed by monomials of $\bar{v}$. The $\frac{1}{N}$ scaling of the effective terms implies that for consistency with respect to uniform mean measure, for all m and n, the coefficients $[A]_{mn}$ should be of $O(N)$. The condition restricts the choice of interactions to be either sparse or weak in magnitude. If the condition does not hold and N is large, then for $[A]_{mn} \sim O(1)$ some of these terms can be neglected or considered as a perturbation of the self-dynamics as analyzed in [7]. When the terms are very large, i.e. $[A]_{mn} \gg O(N)$ then these are expected to dominate the dynamics. As a consequence, the following corollaries are noteworthy:

When interactions are modeled by polynomials of v_i and v_j up to third order (i.e. $m + n \leq 3$) the system in Eq. (5.3) is of the form

$$\dot{\bar{v}} = -\bar{w} + \frac{1}{N}([A_{10}] + [A_{01}])\bar{v} + \frac{1}{N}([A_{11}] + [A_{02}] + [A_{20}])\bar{v}^2 \quad \bigg| \quad \dot{\bar{w}} = b\bar{v} - c\bar{w}$$

$$+ \frac{1}{N}([A_{21}] + [A_{12}] + [A_{30}] + [A_{03}])\bar{v}^3 \qquad (5.5)$$

and for the averaged synchronized correction, Eq. (5.7) becomes

$$\dot{\bar{\xi}}_v = -\bar{\xi}_w + \frac{1}{N}([A_{10}] + [A_{01}])\bar{\xi}_v + \frac{2}{N}([A_{11}] + [A_{02}] + [A_{20}])\bar{v}\bar{\xi}_v \quad \bigg| \quad \dot{\bar{\xi}}_w = b\bar{\xi}_w - c\bar{\xi}_v.$$

$$+ \frac{3}{N}([A_{21}] + [A_{12}] + [A_{30}] + [A_{03}])\bar{v}^2\bar{\xi}_v \qquad (5.6)$$

Assuming that the fluctuations are synchronized such that for $\xi_{vi}(t) = \bar{\xi}(t) + O(\varepsilon)$, Eq. (5.4) becomes

$$O(\varepsilon): \; \dot{\bar{\xi}} = \frac{1}{N} \sum_{m=0}^{p_1} \sum_{n=0}^{p_2} (m+n) \left[A \right]_{mn} \bar{v}^{m+n-1}\bar{\xi} \qquad (5.7)$$
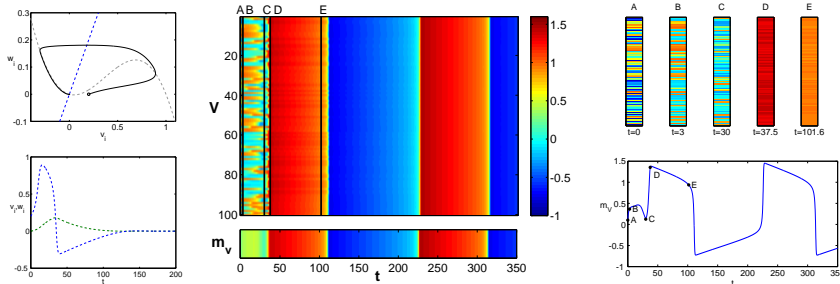
13

FIG. 2. *(Color online) Left, top: The phase plane and the action potential trajectory of the identical i-th FHN neuron. The (blue) dashed line corresponds to the $\dot{w}_i = 0$ nullcline and the (gray) dashed cubic curve corresponds to the $\dot{v}_i = 0$ nullcline. The (black) thick trajectory is an action potential trajectory initiated by an initial condition $(v_i(0), w_i(0)) = (0.2, 0)$, denoted by a circle. Left, bottom: The trajectory of the voltage v_i and the refractory variable w_i in (blue) and (green) respectively. Middle: Synchronization in a network of identical FHN neurons, with parameters $\alpha = (a_1, a_2, a_3, b, c) = (-0.1, 1.1, -1, 10^{-2}, 10^{-2})$ and initial data $(v(0), w(0)) = (0.05 + U(-1, 1), U(-0.05, 0.05))$ coupled linearly at $t = 30$ via excitatory connections. Raster plot of the voltage of the network $v(i, t)$ and its spatial average $m_v(t)$ below. Right, top: five snapshots of the network activity: A,B: no-interactions in the network, C: interactions introduced, D,E: network dynamics settle to a synchronous attractor. Right, bottom: the spatial average $m_v(t)$ of the network with dots correspond to the times at which time snapshots were taken.*

6. Computational study of a network of identical FHN neurons. For linear, bi-linear and synaptic interactions we show numerically for several examples that there are various attractors in the full network of similar FHN neurons. These attractors can be determined using a phase-plane analysis of the AMEE model and it will not necessarily be the attractor that is expected for each individual neuron. This again demonstrates the effectiveness of the AMEE in predicting *a priori* network dynamics that would not be guessed from individual neuron behavior. We show that indeed trajectories of the AMEE model in Eq. (5.3) shadow the trajectory of the full network. Notice that initial conditions of the network and the AMEE should be chosen close to each other, within $O(\varepsilon)$, such that they will be comparable for the transient behavior as well.

To simulate the activity of the network we implemented a network of identical FHN neurons with parameters as $\alpha = (-0.1, 1.1, -1, 10^{-2}, 10^{-2})$ for most of the cases. For these parameters the phase plane of each neuron is depicted in Fig. 2 where spiking behavior is supported. Such a phase plane contains a single fixed point at the origin (intersection of the nullclines $\dot{v}_i = 0$ and $\dot{w}_i = 0$). The interaction matrices A_{mn} are chosen as full interaction matrices, thus implementing all-to-all connections with the same coupling coefficient. In Fig. 2 we demonstrate an activity of a network with $N = 100$ vertices. From $t = 0$ to $t = 30$ the nodes evolve from random initial conditions. At $t = 30$, linear, uniform all-to-all connections such that A_{01} in Eq. (5.5) is nonzero, are turned on. To graphically plot the activity of the network we use a color 'raster plot' of the $v_i(t)$ variables (voltage of all neurons in the network vs. time), left column of Fig. 2. We also compute the average voltage activity of the network m_v, defined in Eq. (5.2), and display it below the raster plot with the same color code as the raster plot and additionally as a curve in the bottom of the right column of Fig. 2. To further demonstrate the synchronization, on the top of the right column of Fig. 2 we show snapshots of v_i at five different times.

Since the initial state of the network is highly variable as shown in the snapshot of

14

$t = 0$, it can be observed that there is no synchronization from $t = 0$ to $t = 30$, however when interactions are turned on the network rapidly synchronizes into a common state that exhibits relaxation-oscillation dynamics. Recall that each individual neuron left to its own produces spiking behavior. From snapshots at times $t = 37.5$ and $t = 101.6$ it is seen that indeed the activity of all nodes in the network is of the same value up to small variations and exhibits relaxation oscillations in contrast to individual isolated neuron single spiking behavior. Indeed, these simulations suggest that for the parameters chosen in Fig. 2, the relaxation-oscillation solution has a large basin of attraction and the mean AMEE model developed here is expected to shadow the mean solution shown in Fig. 2, as proved in Theorem 1. Therefore, in the following parts of this section we study different network setups, parameter regimes and interactions, to demonstrate various solutions of the mean AMEE model and the averaged activity.

6.1. Linear interaction. Linear form of interactions can correspond to a current injected into the membrane of the i-th neuron whose strength is proportional to the voltage of the j-th neuron - electrical interaction [10]. The equations for the i-th neuron can be written as

$$\dot{V}_i = f^{FHN}(V_i; \alpha) + \sum_{j=1, j \neq i}^{N} \frac{a_{int}}{N-1} v_j \tag{6.1}$$

and the nonzero coefficients in the corresponding first order mean AMEE model are $[A_{10}] = Na_1$, $[A_{01}] = Na_{int}$, $[A_{20}] = Na_2$, $[A_{30}] = Na_3$, implying that, in effect, the linear term has been modified to $a_1 + a_{int}$. Linearization around the fixed point at the origin yields the Jacobian J with eigenvalues $\lambda_{1,2}$

$$J = \begin{pmatrix} a_1 + a_{int} & b \\ -1 & c \end{pmatrix}, \quad \lambda_{1,2} = \frac{1}{2}\Big[c - (a_1 + a_{int}) \pm \sqrt{(c + a_1 + a_{int})^2 - 4b}\Big].$$

See [11, 31] for the analysis of the FHN phase plane and the dynamics.

In our numerical simulations, an individual non-interacting neuron ($a_{int} = 0$) produces spiking behavior as demonstrated in Fig. 2. When interactions are introduced, fixing α and varying a_{int}, the mean AMEE model indicates that we can change the behavior of the network as a synchronized collection of neurons from its single base behavior which is a single action potential. These variations in behavior correspond to changes in the characteristics of the nullclines and stability of fixed points. To study the effect of the interaction we demonstrate the network activity with three plots: a projection of the average onto the phase plane, the average profile of the voltage/refractory and three randomly selected voltage/refractory profiles. These correspond to left, middle and right plots in Fig. 3. In the phase plane plot, the nullclines of a single neuron are depicted by dashed lines and the nullclines of the AMEE model (effective nullclines) are depicted by solid lines. The solution of the AMEE model ($\bar{v}(t)$ and $\bar{w}(t)$) is depicted on top of these nullclines with a solid curve. The average of the voltage and the refractory variable, of the interacting network, is sampled and depicted using the asterisks symbol. The average voltage/refractory profile plot, includes the single neuron voltage and refractory variable trajectory depicted by dashed blue and green color respectively. The $\bar{v}(t)$ and $\bar{w}(t)$ solutions of the AMEE is depicted by a solid blue and green color respectively. The sampled averages are depicted using the asterisk symbol. In the right plot, that shows randomly selected voltage and refractory profiles for three neurons, the blue solid line corresponds to
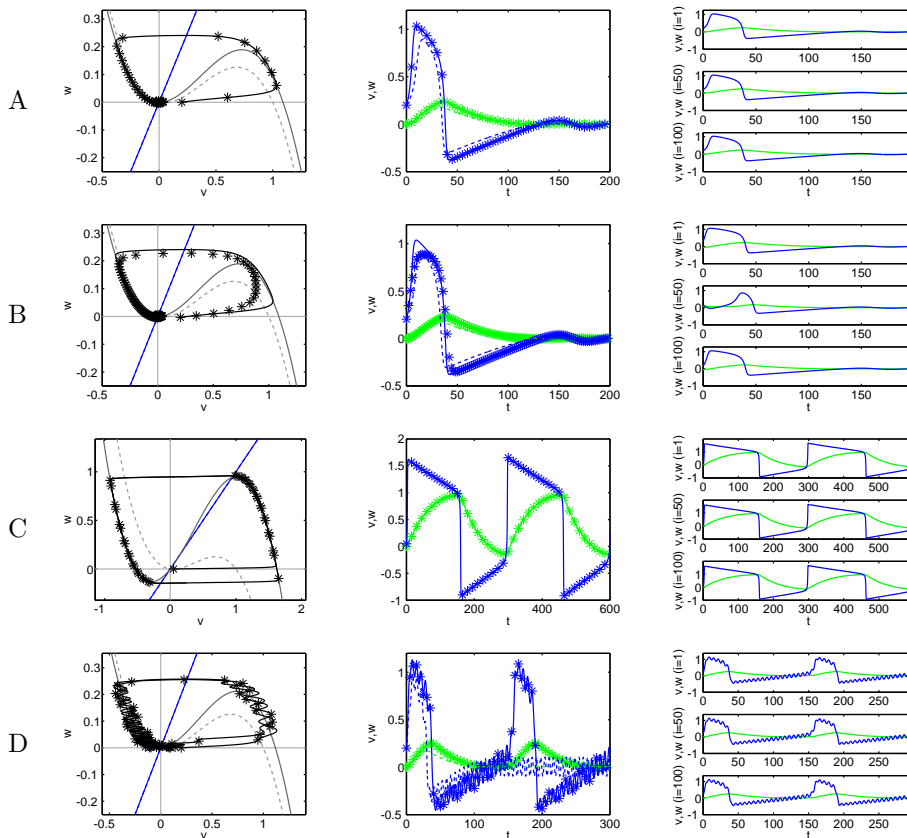
Fig. 3. *(Color online) FHN network with linear interactions. See the text of §6.1 for the explanation of the plots.*

the voltage and the green solid line corresponds to refractory variable of a particular neuron.

In Fig. 3A, by setting the pairwise interaction between each pair of neurons to be approximately 10^{-3}, the average behavior of the network exhibits an action potential for identical initial conditions $(v_i(0), w_i(0)) = (0.2, 0)$. Such interaction is very small, almost negligible in terms of voltage values that the scaled neuron can produce ($v_i^{min} \approx -0.1$ and $v_i^{max} \approx 1.1$), however since these are all-to-all interactions it creates a significant effective change in the dynamics; the equations that describe the dynamics of the average activity will be modified by $a_{int} = 0.089$. From the voltage/refractory plot (middle), one can observe that the average behavior of the network has deformed from the single neuron behavior. The action potential is more elongated and has decaying subthreshold oscillations around the resting state. Noticing that the projected average solution and the solution of the AMEE model are very close to each other demonstrates that the AMEE solution shadows the average activity of the network. Indeed, the plots of random samples of activity (right) strengthen that conclusion, since it is observed that the network seem to reach rapidly a synchronized state solution. Hence, the deformed behavior can be easily understood by studying the phase plane of the AMEE model (left). The cubic nullcline of the AMEE

16

model (solid) is deformed from the single neuron cubic nullcline such that its maximum point is located higher and its intersection with the horizontal axis, $w = 0$, is at higher value of v. Indeed, the effect of the maxima point being elevated is that the trajectory will follow the slow cubic nullcline for a longer time and that the action potential will reach a higher peak value in the voltage. Furthermore, the stability calculation indicates that the resting fixed point becomes a sink while for the single neuron it was a stable node, explaining the subthreshold oscillations after the recovery from the action potential.

In Fig. 3B the same interaction is chosen as in 3A, but the initial conditions are nonidentical and drawn from a uniform distribution. For a wide distribution of the initial conditions as chosen in this example, for a transient time from initiation the error between the solution of the AMEE model and the average activity is observed. As expected, the trajectories do not seem to be close to each other. However beyond the transient time, the error decreases and eventually the average and the AMEE trajectories become close to each other.

To further investigate the closeness of the average and the AMEE model we choose the interaction such that the phase plane of the AMEE model will bifurcate and the network is then expected to exhibit a completely different dynamics that corresponds to that modified phase plane. From geometrical arguments, we can conclude that decreasing a_{int} such that $a_{int} < -\frac{1}{4}(a-1)^2$, the $\dot{v}_i = 0$ nullcline will become monotone decreasing. For the parameters chosen here, this threshold corresponds to inhibitory interactions of $a_{int}^{inh} \approx -0.2025$. In this situation, no action potential is possible which means that choosing the interaction lower than a_{int}^{inh} will result with a completely inhibited network. On the other hand, increasing a_{int} by choosing $a_{int} > \frac{1}{4}(a-1)^2 - \frac{b}{c}$, the two nullclines have multiple points of intersection, resulting in the creation of a second stable fixed point. The threshold for this condition, with our chosen parameters, is $a_{int}^{bi} \approx 0.7975$. In that case, the interacting network is expected to exhibit bi-stability, such that some initial conditions will lead to the polarized network and some to the rest state. Indeed numerical simulations indicate that the network exhibits such solutions.

Further increasing a_{int}, by moving a_{int} outside of the range $|a_{int}| < \frac{b-a_1 c}{c}$ causes the fixed point at the origin to become unstable. For our choice of parameters as described above, that means $a_{int}^u = 0.9$. For such interaction and small enough values of b, c the network can exhibit relaxation-oscillation dynamics. These solutions are shown in Fig. 3C and 2. For initial conditions near the unstable resting fixed point, the network almost instantly is locked into large oscillations that are the solutions of the AMEE model with similar initial condition.

Linear stability analysis of the resting fixed point indicates that the fixed point undergoes Andronov-Hopf bifurcation when $trJ = a_1 + a_{int} + c = 0$. The bifurcation is of a supercritical type such that when $trJ > 0$ an attracting limit cycle is created. Therefore, a network with interactions $a_{int} > a_{int}^{AH} = a_1 + c$ and initial conditions close to the resting fixed point is expected by the AMEE model prediction to possess indefinite non-decaying "low amplitude" subthreshold oscillations of the size of the limit cycle. Our numerical simulations, shown in Fig. 3D validate this prediction.

Aside from internal interactions, many networks have external input to the neurons. Such an external input is supposed to excite or inhibit the network and models the external stimulus applied to the network to receive a specific functionality. To study the mean AMEE approach to a network subject to such an input we have added a periodic time-dependant term to Eq. (6.1), $I_{i,ext} = 0.1 \sin(t)$. Such an input to the

single neuron is not expected to change its qualitative behavior since the amplitude of the input is below the spiking threshold for the chosen single neuron. For initial conditions $(v_i(0), w_i(0)) = (0.2, 0)$ the neuron is expected to exhibit one action potential (with low amplitude oscillations on top of the voltage profile) and then settle onto a limit cycle near the rest voltage, created by the periodic input, as demonstrated in Fig. 3E. On the population level, however, when interactions are introduced, the network repeatedly exhibits action potentials. Indeed, from Fig. 3E it can be observed that choosing the interaction in the regime of a single elongated action potential and then low amplitude oscillations, $a_{int} = 0.12$, the external input causes the network to produce periodic elongated action potentials. The AMEE model for such a non-autonomous network is similar to the AMEE of an autonomous network - it is the autonomous mean AMEE model with the additional term $0.1 \sin(t)$ in the equation for $\dot{\bar{v}}$ in Eq. (5.5). The phase plane analysis of the autonomous AMEE model explains the source of the periodic firing - for $a_{int} = 0.12$ the rest fixed point becomes unstable and the Andronov-Hopf bifurcation occurs. Periodic perturbation from the limit cycle, created by the bifurcation, causes the trajectory to cross the threshold for the action potential and to fire repeatedly.

6.2. Bi-linear interaction. Since the mean AMEE reduction approach can be applied to any analytical pairwise interaction, in this subsection we simulate a network of bi-linear all-to-all interactions. Such an interaction can be expressed with the following equations for the i-th neuron

$$\dot{V}_i = f^{FHN}(V_i; \alpha) + \sum_{j=1, j \neq i}^{N} \frac{a_{int}}{N-1} v_i v_j \tag{6.2}$$

Although we choose such an interaction mainly for the purpose of simulating the network with the simplest nonlinear interaction, it has some biological inspiration. It can be suitable for a simple modeling of neural connections that depend both on the voltage of the i-th neuron and the voltage of the interacting j-th neuron. Intuitively, the voltage level of the i-th neuron determines its excitability. For example, if the i-th neuron near the rest state at some t then the interacting j-th neurons should have much larger voltage levels in order to excite it. However, when the voltage of the i-th neuron is near the threshold for an action potential, even small interactions can excite it. Bi-linear interaction term can be also considered as a simplest approximation of fast, almost instantaneous, chemical (synaptic) interactions. When these interactions are modeled by a sigmoid function of the presynaptic voltage v_j multiplying the postsynaptic voltage v_i [59, 60, 61], then linearization of the sigmoid term results in the interaction term being bi-linear: $v_i v_j$.

The mean AMEE model for the bi-linear interaction results with the following nonzero coefficients $[A_{10}] = Na_1$, $[A_{11}] = Na_{int}$, $[A_{20}] = Na_2$, $[A_{30}] = Na_3$. Effectively the interaction term modifies the coefficient of $\bar{v}^2$ from a_2 to $a_2 + a_{int}$. Such an interaction does not modify the stability of the rest fixed point (due to the form of J at $(0,0)$) however it can modify the cubic nullcline of the single neuron such that multiple fixed points exist and the network is expected to exhibit bi-stability. The analysis of the phase plane determines that this value is $a_{int}^{bi} = 0.9976$. Notice that this value, as expected, is greater than the bi-stability criterion for linear interaction. In practice, choosing interaction coefficient above a_{int}^{bi}, there exist initial conditions for which the network will saturate. Similarly, the critical value of the interaction coefficients for which the cubic nullcline is monotonically decreasing can be calculated.
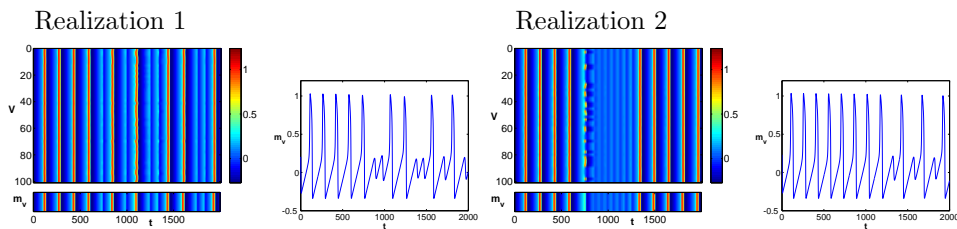
Fig. 4. *(Color online) Chaotic bursting attractor. Two realizations of the synchronous chaotic attractor in the FHR model. In each realization, Left: Raster plot of the interacting network and the spatial average below. Right: Solution of the AMEE initiated with random i.c. around the $(v(0), w(0), x(0))$. See text of §6.3 for an explanation.*

This value is $a_{int}^{inh} = -0.7166$, such that choosing a_{int} below that value will result with an "inhibited" network, in which for all initial conditions the network rapidly return back to the rest fixed point without exhibiting an increase in voltage and then a decrease (action potential like behavior).

6.3. Chaotic synchronous attractor. For a network of FHN neurons, the resulting AMEE are planar, thus restricting the attractor to be non-chaotic (a fixed point or a limit cycle). However, simplicity of the dynamics on the attractor is not required. By increasing the dimension of the governing model of each neuron it is indeed possible that the synchronous attractor will be chaotic. To explore such attractors, we consider the FHR model, Eq. (2.4), with parameters $(a, b, c, d, \varepsilon) = (0.19, 10^{-2}, 10^{-2}, 0.138, 10^{-4})$ that for initial condition $(v(0), w(0), x(0)) = (0.2, 0, 0.047)$ exhibits a similar once spiking voltage profile as in Fig. 2. Connecting identical FHR neurons with linear interactions of strength $a_{int} = 0.099$ brings the AMEE to a regime of elliptic bursting shown to be chaotic (a slow passage of subcritical Andronov-Hopf bifurcation and a fold of a limit cycle) [33].

In Fig. 4 we show two realizations of the FHR network initiated from random initial conditions near the initial condition $(v(0), w(0), x(0))$. It can be observed that for most parts of the evolution the synchronous solution dominates. Furthermore, two different realizations of the network indicate that the synchronous attractor is chaotic, i.e. it is sensitive to initial conditions and exhibits irregular intermittency of relaxations and subthreshold oscillations. In Fig. 4 to the right of the raster plots we show the respective trajectories of the AMEE. The qualitative properties of the trajectories of the averaged activity (colored plots below the raster plots) and the solutions of AMEE are similar, however they cannot be compared quantitatively. This demonstrates that indeed for chaotic synchronous attractors we can assure the closeness of AMEE and activity measure only in terms of shadowing as proved in Theorem 1. Notice that this example also demonstrates the significance of the choice of self-dynamics. While the network of FHN was not capable of producing time chaotic synchronous solutions, the network of FHR could produce synchronous solutions with more complex evolution.

6.4. Synaptic interaction. Time scales of chemical interactions (synaptic) can typically be longer than the time scales of the neural spiking. Such interactions can be modeled by specifying an auxiliary variable that represents the synaptic activity of the pre-synaptic neuron [59, 62, 15]. The variable is activated when the neuron produces an action potential and after activation, synaptic activity decays to zero with a decay rate characteristic to the specific synaptic connection. As an example

for such interactions, we model a network of FHN neurons coupled via a single channel continuous synaptic conductance-based model [15]. The connections are expressed as the sum of the auxiliary variables of the pre-synaptic neurons s_j multiplied by the voltage of the post-synaptic neuron v_i

$$\dot{V}_i = f^{FHN}(V_i; \alpha) + \sum_{j=1, j \neq i}^{N} \frac{a_{int}}{N-1} v_i s_j, \qquad \tilde{V}_i = (v_i, w_i, s_i) \qquad (6.3)$$

$$\dot{s}_i = a_r \phi(v_i; \kappa, v_{th})(1 - s_i) - a_d s_i.$$

The parameters a_r and a_d correspond to the rise and decay times of the synaptic activity and are usually chosen such that $a_r \gg a_d$. The function ϕ is the activation of the synaptic activity and typically chosen as a sigmoid function $\phi(v_i; \kappa, v_{th}) = 1/(1 + \exp(-\kappa(v_i - v_{th})))$, where the parameters κ and v_{th} are responsible for tuning the maximal amplitude of the synaptic variable and the voltage level above which the synaptic variable is activated [15].

Here we set the parameters to $(\kappa, v_{th}, a_r, a_d) = (50, 0.95, 20, 0.1)$ that correspond to synaptic interactions with long decay times such that the synaptic variable cannot be approximated by a voltage function and it is a part of the AMEE model. Indeed, for identical neurons, synaptic interaction can lead to synchronous attractors with dynamical features different than the isolated individual neurons. For such synchronous solutions, the resulting AMEE model of $\mathcal{O}(1)$ for uniform mean is similar to Eq. (5.3), with the monomial effective terms being replaced by coupling to the effective synaptic variable producing a system of three equations.

For individual neurons as in Fig. 2, we demonstrate in Fig. 5 that slow decay rate interactions cause the network to exhibit subthreshold oscillations or periodic firing depending on the amplitude of the interactions or decay time. The green curve in the middle plot depicts the synaptic variable approximation by the AMEE model (solid) and the sampled average of the synaptic variable (asterisk). The projections onto the plane of $(\bar{v}, \bar{w})$ remain useful for understanding the network dynamics, however the prediction of the effective nullclines/stable and unstable manifolds in the three dimensional system is not straightforward. Geometrical methods such as singular perturbation theory [63] can provide such a prediction for all-to-all connections or for specific symmetries in the connection matrix. In general, for complex interaction topologies and possibly for nonlinear activity measures, the resulting AMEE is an extremely nontrivial system to analyze.

In contrast to limitations of the analytic study, computational methods such as numerical continuation can be applied efficiently by exploiting the low dimensionality of the reduced AMEE model. Such studies allow one to explain and predict dynamics of neural activity when a few control parameters are being varied. As an illustrative example we study the firing rate of the synchronous periodic solution of the FHN network (as in Fig. 5B) connected with inhibitory or excitatory synaptic connections. We fix $(b, c) = (10^{-3}, 10^{-3})$ parameters in the FHN neurons and vary a. In the range of $a \in (-0.5, 0)$ the model exhibits periodic firing of relaxation-oscillations type (for negative values outside this range linear nullcline intersects the cubic nullcline such that stable fixed point is created at and for positive values outside this range the rest fixed point is stable). The firing rate (FR) curve is depicted in Fig. 6 (left). Connecting the neurons via synaptic connections can deform and reposition significantly the FR as shown in Fig. 6 (middle,right). For small values of a, the excitatory synaptic connections with low amplitude position the effective FR curve of the synchronous
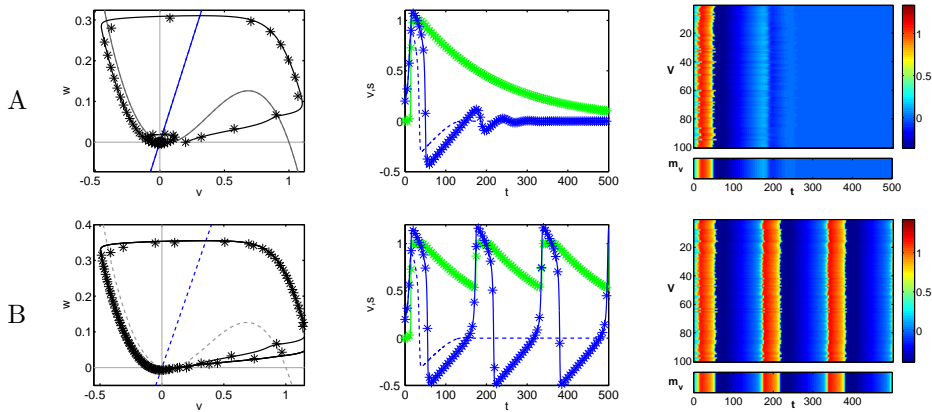
FIG. 5. *(Color online) Synaptic interactions. A: A network of identical neurons interacting synaptically exhibiting subthreshold oscillations induced by the interaction. The interaction coefficient is $a_{int} = 0.1$ and synaptic rise and decay parameters are $(a_d, a_r) = (20, 0.1)$ and the initial condition is $(v_i(0), w_i(0)) = (0.2, 0)$. B: Larger amplitude interactions, $a_{int} = 0.25$, resulting with periodic spiking for initial conditions and rise and decay parameters as in the top row.*

solution (marked by a green asterisk curve) above the FR curve of the isolated neuron. As a is increased, the FR curve it crosses the isolated neuron's curve resulting with slower FR than the isolated neuron (although being excitatory connections). Increasing the amplitude of the synaptic connection places the effective FR curve (marked by green triangles) above the isolated neuron's curve for the whole range of a. Keeping the large amplitude, but decreasing the decay time of the synaptic variable, the effective FR curve is positioned below the isolated neuron's FR curve for the whole range of a (marked by a dashed green curve), significantly delaying the firing rate (cf. Fig. 6 middle). Similarly, inhibitory interactions result with nontrivial, unexpected firing rates. Low amplitude fast interactions position the effective FR curve above the isolated neuron's curve, while high amplitude fast and slow interactions increase the firing rate and position the effective curve above the isolated neuron's curve (cf. Fig. 6 right).
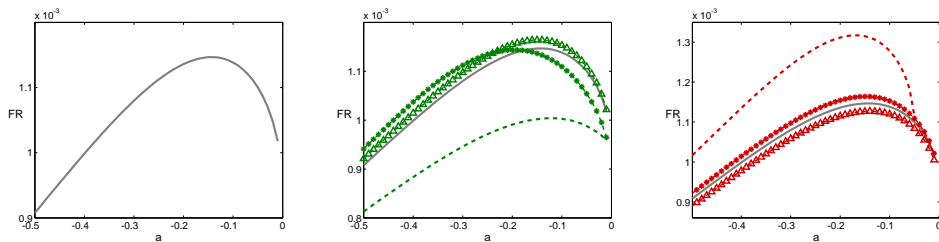

FIG. 6. *(Color online) Firing rates as a function of the parameter a with excitatory and inhibitory synaptic interactions. Left: Firing rates for isolated identical neurons $\alpha = (a, a - 1, 1, 10^{-3}, 10^{-3})$. Middle: Excitatory fast and slow synaptic interactions with parameters being varied $(a_d, a_{int}) = (0.1, 0.05)$ (starred green) and $(a_d, a_{int}) = (0.1, 0.1)$ (triangles green) for fast interactions. Parameters $(a_d, a_{int}) = (0.001, 0.1)$ (dashed green) chosen for slow interactions. Right: Inhibitory fast and slow synaptic interactions with parameters being varied $(a_d, a_{int}) = (0.1, -0.05)$ (starred red) and $(a_d, a_{int}) = (0.1, -0.1)$ (triangles red) for fast interactions. Parameters $(a_d, a_{int}) = (0.001, -0.1)$ (dashed red) for slow interactions.*
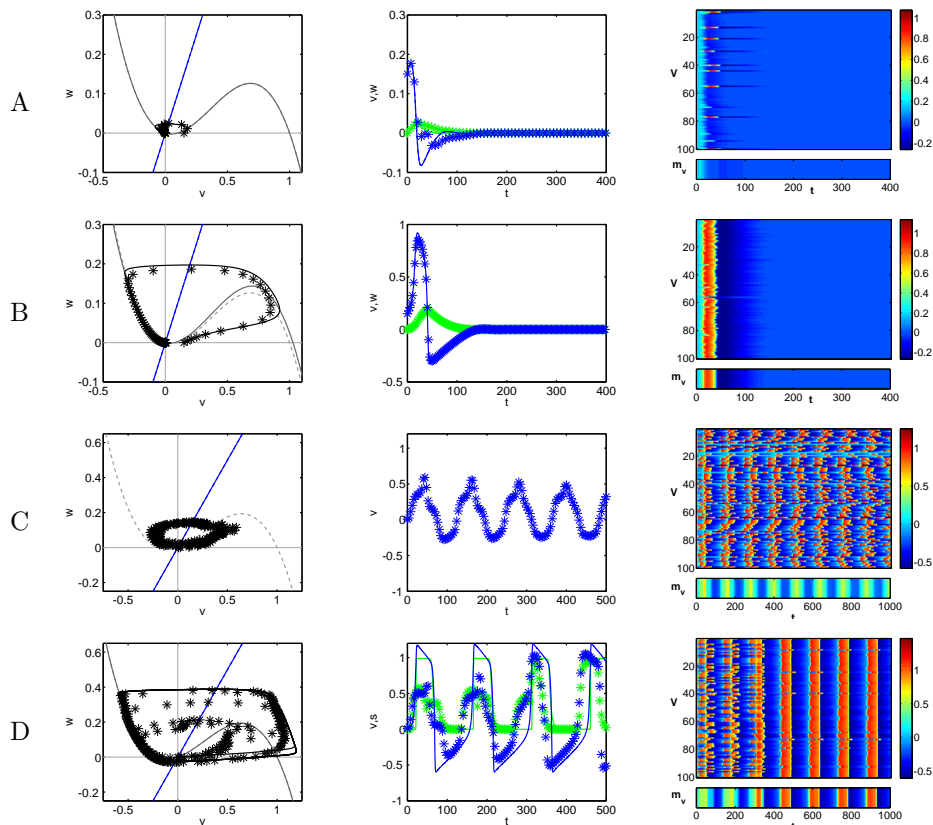
FIG. 7. *(Color online) A network of near-identical neurons. A: Non-interacting neurons for initial conditions $(v_i(0), w_i(0)) = (0.15, 0)$. B: Interacting neurons, $a_{int} = 0.025$, for initial conditions as in the top row. C: Non-interacting neurons for initial conditions $(v_i(0), w_i(0)) = (0.01, 0) + U(-0.05, 0.05)$. D: Synaptic interaction induces synchronization, $a_{int} = 0.3$, for initial conditions as in the third row and synaptic timescale parameters $(a_r, a_d) = (20, 0.2)$.*

7. Network of non-identical neurons. In previous numerical simulations we have chosen identical neurons and tested the validity of the synchronous attractor assumption. Figures 2-4 illustrate that for identical linearly interacting FHN neurons there exist a synchronized attractor and hence for large parameter range the synchronous attractor assumption is valid.

7.1. Almost identical neurons. Here we simulate an interacting network of almost identical FHN neurons. The parameters were drawn from uniform distribution with zero expectation around the parameters for the identical neuron such that $(a_1, a_2, a_3) = (-0.1, 1.1, -1) + U(-0.05, 0.05)$, $(b, c) = (10^{-2}, 10^{-2}) + 5 \times U(-10^{-3}, 10^{-3})$. Since the uniform distribution has zero expectation, the average parameters are the parameters of the single neuron as chosen in Fig. 2. However the variation especially in the a_1 parameter, that determines the stability of the fixed point and the threshold value, changes significantly the dynamics of each individual neuron. Therefore, we expect the collection of non-interacting single neurons to have various behavior for initial conditions $(v_i(0), w_i(0)) = (0.15, 0)$ chosen close to the threshold value of the single average neuron, as demonstrated in the raster plot in

Fig. 7A, where only a small fraction of neurons pass the threshold and produce and action potential. In such a case the AMEE model is not guaranteed to follow the asynchronous solution which is also clearly observed from the numerical simulation.

However, when interaction is introduced, numerical simulations show that even very small interaction ($a_{int} = 0.025$) that does not cause a bifurcation in the phase plane structure of the average single neuron, the network exhibits a synchronized solution in the form of a single action potential, as demonstrated in Fig. 7B. The AMEE model constructed for the identical (average) neurons case, can still be a good approximation for the average dynamics. The projection to a phase plane in Fig. 7 illustrates the closeness of the two solutions and the raster plot indicates that there is synchronization. These numerical simulations lead to a conjecture of the existence of an underlying attractor for nearly identical neurons as well and implies that the mean AMEE model is a valid approximation for the average dynamics even for almost identical network of linearly interacting FHN neurons.

Note that existence of a synchronous solution implies that for networks with general random neurons, not necessarily almost identical, and individual dynamics modeled by a polynomial vectorfield, the mean AMEE model will shadow the dynamics of the spatial average on the synchronous attractor. To demonstrate how a synchronous solution is formed and how the AMEE models the synchronized solution, we simulate a non-identical network with $a = -0.3 + U(-0.05, 0.05)$ of periodic FHN neurons. The uncoupled network exhibits an asynchronous solution where each neuron is firing periodically with a different firing rate, as shown in Fig. 7C. When synaptic interaction is introduced, the network gradually approaches a synchronous solution. Indeed, the dynamics of the synchronous solution are predicted by the AMEE model as shown in Fig. 7D.

7.2. Two interacting populations of neurons. As described in §4.2 it is often the case that the network of neurons can be modeled as being composed of several distinct subpopulations, where within each subpopulation the nodes are identical or almost identical. Here we study the case of two subpopulations of FHN neurons. The first subpopulation is chosen with parameters $\alpha^{(1)} = (-0.03, 1.03, -1, 10^{-2}, 10^{-2})$ such that for initial condition $(0.1, 0)$ the voltage spikes once and settles to subthreshold oscillations (a stable limit cycle). The second population is chosen with parameters $\alpha^{(2)} = (0.01, 0.99, -1, 10^{-2}, 10^{-2})$ such that for the same initial condition it does not cross the threshold for action potential and settles to the stable resting fixed point. Interactions within each subpopulation (positive linear, all-to-all, $a_{int} = 0.02$) cause the first population to exhibit relaxations oscillations and the second to spike and settle to a stable resting fixed point, as depicted in Fig. 8A.

The choice of such parameters represents two subpopulations where subpopulation (1) is more sensitive than the other (2). Connecting these populations can create an attractor solution that is different than the attractor solution of each subpopulation connected within. In the particular example here, with the same linear connections, we expect that for positive excitatory connections the more sensitive subpopulation will drive the less sensitive subpopulation. As demonstrated in Fig. 8B this is indeed the case. Both populations exhibit relaxation oscillations, where the second population is delayed after the first one. Following §4.2 by extending Eqs. (5.5) we derive the
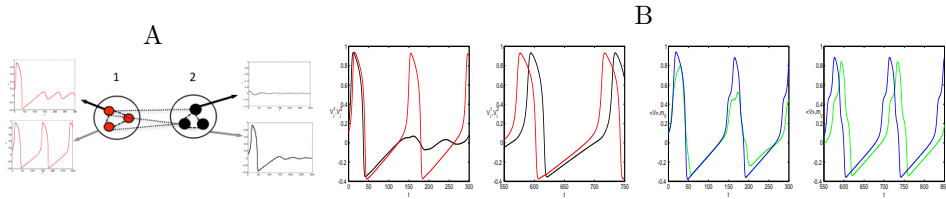
FIG. 8. *(Color online) Dynamics of a network of FHN neurons composed of two distinct populations. A: a schematic chart of the network. The plots indicate the average voltage of the network for non-interacting (top) and interactions within each population (bottom). B: Left to right: Dynamics of the average voltage of each population for $t = [0, 300]$ and $t = [550, 850]$, overall spatial average, dark (blue) curve, along with the reconstructed overall average from the solution of the AMEE, bright (green) curve for $t = [0, 300]$ and $t = [550, 850]$.*

reduced first order mean AMEE as four coupled odes

$$\dot{\bar{v}}_1 = -\bar{w}_1 + (a_1^{(1)} + a_{int})\bar{v}_1 + a_2^{(1)}\bar{v}_1^2 + a_3^{(1)}\bar{v}_1^3 + \frac{\tilde{N}_1}{\tilde{N}_2}a_{int}\bar{v}_2, \quad \dot{\bar{w}}_1 = b\bar{v}_1 - c\bar{w}_1, \quad (7.1)$$

$$\dot{\bar{v}}_2 = -\bar{w}_2 + (a_1^{(2)} + a_{int})\bar{v}_2 + a_2^{(2)}\bar{v}_2^2 + a_3^{(2)}\bar{v}_2^3 + \frac{\tilde{N}_2}{\tilde{N}_1}a_{int}\bar{v}_1, \quad \dot{\bar{w}}_2 = b\bar{v}_2 - c\bar{w}_2.$$

To demonstrate the closeness of the AMEE solution and the mean, in Fig. 8B we compare for two time intervals of identical duration, the initial time interval (left) and a later interval (right), the mean over each population (top row), the voltage components $\bar{v}_1, \bar{v}_2$ of the AMEE model (middle row) and the total mean along with $\bar{v}_1 + \bar{v}_2$ (bottom row). As these plots demonstrate, after a transient time the solution of AMEE reproduces the observed dynamics and captures accurately the delay between the oscillations of the two subpopulations. As a result, the nontrivial profile of the total mean is being recovered accurately.

8. Coherent structures. Non existence of a completely synchronous attractor, especially for random initial conditions, does not imply a spatial decoherence of the attractor. In fact, typical dynamics on such attractors exhibit partial synchronization expressed by interaction of spatially coherent structures. The form of the coherent patterns depends on the symmetries in the network and is especially expected in networks in which there is a synchronous attractor for some parameter regime or initial conditions. As shown in §4.3, coherence can be used to reduce the AMEE and resolve the dynamics of the interactions of the coherent structures and to construct the activity measure for the partially synchronous network. Here we show a construction of AMEE for gap junctions interactions where the random initial condition is not in the basin of attraction of the completely synchronous attractor. Instead, this initial condition is attracted to a periodic coherent pattern.

8.1. Gap junctions. As an example of AMEE for partially synchronous attractors we consider a FHN network interacting by all-to-all symmetric gap junctions

$$\dot{v}_i = f^{FHN}(V_i; \alpha) + \frac{a_{int}}{N-1} \sum_{j=1, j \neq i}^{N} (v_i - v_j). \qquad (8.1)$$

with the parameters α for the individual neurons chosen as in Fig. 2 and $a_{int} = 0.2$. Such an interaction imposes the dynamics on the complete synchronous attractor to
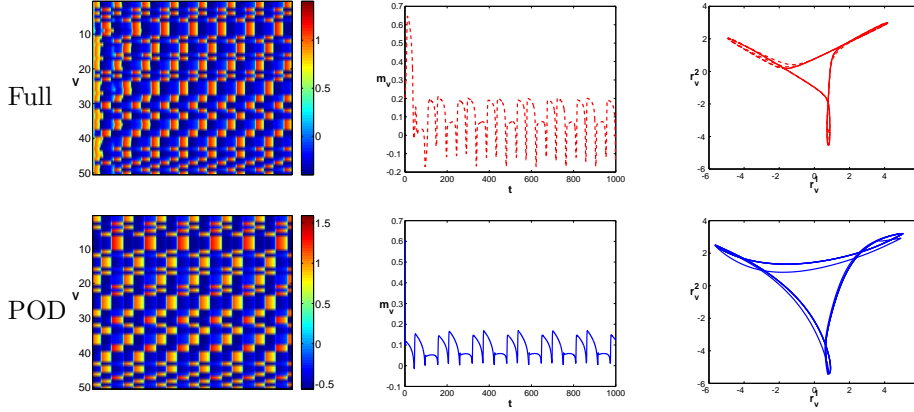
FIG. 9. *(Color online) Non-synchronous coherent regime for gap junctions interactions. Top: Simulation of the full network. Bottom: Simulation of the projected system onto POD modes. Left: Raster plot. Middle: The overall spatial average of the network and the reconstructed network. Right: Projection of the dynamics onto two modes demonstrating a limit cycle.*

be the same as the dynamics of the individual neuron - a single action potential. For identical initial conditions, as in Fig. 2, the network exhibits such dynamics. However, randomly perturbed initial data does not belong to the basin of attraction of the synchronous solution, see [43] for a stability study. Instead, coherent structures in the form of displaced relaxations oscillations appear as demonstrated in Fig. 9. Such dynamics can be explained by decomposing the interaction in Eq. (8.1) into the self-interaction term (v_i) and non self-interaction term (v_j) resulting in $a_{int}v_i - \frac{a_{int}}{N-1}\sum_{j=1,j\neq i}^{N} v_j$. Including only the self-interaction term will result in a FHN model in the regime of an unstable resting state fixed point and a stable limit cycle such that an initial condition as in Fig. 2 results with relaxation oscillations dynamics. The non self-interaction term, when neurons are perfectly synchronized, is supposed to balance the self-interaction term and bring the resting state to stability. For random initial data, the balance is imperfect which in turn creates a delay between the relaxation oscillations of the neurons.

To derive the AMEE model that will resolve the activity measures of this solution we perform a singular value decomposition on the network matrix $v(i,t) = U_v S_v Y_v^T$, $w(i,t) = U_w S_w Y_w^T$. Then the column vectors of Y^T, $\vec{y}_v^m$ and $\vec{y}_w^m$, are the proper orthogonal modes such that the dynamics on the attractor can be approximately represented as a finite decomposition of these vectors multiplied by time varying coefficients, $r_v^m(t)$ and $r_w^m(t)$. The resulting decomposition of the attractor is then $\vec{x}_v^A(t) = \sum_{m=1}^K r_v^m(t)\vec{y}_v^m$ and $\vec{x}_w^A(t) = \sum_{m=1}^K r_w^m(t)\vec{y}_w^m$. Following §4.3, to derive the equations for $r_k^m(t)$ we project Eqs. (8.1) onto three modes using the Galerkin projection. Using Newton's monomial formula the self terms in Eqs. (8.1) result with

$$\dot{r}_v^m(t) = -\sum_{s=1}^3 <\vec{y}_w^s, \vec{y}_v^j> r_w^s + (a_1 + a_{int})r_v^m + \sum_{p_1+p_2+p_3=l} a_l \prod_{q=1}^3 d_{p_1p_2p_3}^{l,m}(r_v^q)^{p_q}, \quad (8.2)$$

$$\dot{r}_w^m(t) = b\sum_{s=1}^3 <\vec{y}_v^s, \vec{y}_w^j> r_v^s - cr_w^m$$

where the coefficients $d_{p_1p_2p_3}^{l,m}$ are the inner product of monomials of $\vec{y}_v^q$ with $\vec{y}_v^m$ and

multiplied by the appropriate monomial coefficient such that $l = p_1 + p_2 + p_3$ and $p_1, p_2, p_3 \in \mathbb{Z}^+$

$$d^{l,j}_{p_1 p_2 p_3} = \binom{l}{p_1, p_2, p_3} < (\vec{y}_v^1)^{p_1} (\vec{y}_v^2)^{p_2} (\vec{y}_v^3)^{p_3}, \vec{y}_v^m > .$$

The non-self interaction term is projected by substituting into v_j the orthogonal decomposition resulting with circular shifts of the vectors $\vec{y}_v^q$ by $h = 1, ..., N-1$ denoted by $\vec{y}_v^{q,h}$

$$\dot{r}_v^m = \frac{a_{int}}{N-1} \sum_{q=1}^{3} \sum_{h=1}^{N-1} < \vec{y}^{q,h}, \vec{y}_v^m > r_v^q. \tag{8.3}$$

The combination of Eqs. (8.2) and (8.3) is the projected model onto three orthogonal modes, resulting with a system of six coupled odes with the coefficients r_v^m, r_w^m being the variables in these equations. In a similar manner the model can be extended to a projection onto K modes.

Fig. 9 demonstrates the full network dynamics (top) and the reduced model (bottom). Indeed, the raster plot (left) and the projection onto the first two modes (right) show similar patterns and that their dynamics are periodic both in the full network and in the projected POD model. The spatial mean of the network (middle, top) and the mean activity computed according to Eq. (4.6) (middle,bottom) indicate that the overall average is fluctuating around the zero with discrete jumps at times corresponding to switching between patterns. The reduced model is able to capture these transitions and leads to a qualitative reconstruction of such overall activity. As in previous examples, the reduced AMEE model provides a model for such dynamics. A detailed study of its structure, symmetries and bifurcations, should reveal the conditions for balance/unbalance of such a network [45, 46, 64].

9. Discussion. In this paper we introduced a new theoretical framework for modeling neural activity of a network of interacting neurons based upon the first principle dynamics of individual neurons. Our bottom-up approach gives a self-consistent, analytically constructive, and asymptotically justified "mean-field" reduction (AMEE model) for *quantitatively* describing collective neural activity. The analysis characterizes network dynamics as a function of the individual neuron dynamics and their connection type and strength, often producing collective dynamics that cannot be guessed from individual neuron behavior. Our analysis should be contrasted with previous top-down mean-field descriptions where the mean-field is specified with some generic, qualitative interaction rules among nodes of the network. The derived AMEE model provides a dimensionality reduction of the phase space of the complete network, especially when the asymptotic dynamics of the network exhibit synchrony and the measure is simple. For such networks, the solution of the reduced set of equations shadows the activity measure of the network and hence provides a good approximation for the dynamics of the complete network. Applicability of additional reduction methods to the AMEE model was shown as well. These methods allow for a reduced model even for partial synchrony.

For the mean measure of the neural field, such an approach gives rise to models that retain the structure of the underlying individual neuron whose dynamics are well known. Prescribed neuron-to-neuron interactions introduce additional "mean-field" effects that have a profound effect on the overall network dynamics. The explicit construction of the AMEE model for identical interacting FHN neurons carried out in

Section 5 demonstrates that the resulting evolution equations for the mean measure can be studied by analytic phase-plane analysis. Such a study accurately predicts the mean functionality of the network for a wide range of parameters and initial conditions of the neurons. Indeed, the derived AMEE description is validated for linear and simple nonlinear (bi-linear) network interactions as well as synaptic interactions. Unlike for the mean measure, the theory shows that the dynamics of nonlinear measures results in a mean-field description where the structure of the individual neuron is modified together with the interaction terms. Moreover, it is able to produce a completely modified description at the mean-field level detached from the original single neuron model despite being based on the single neuron structure.

The AMEE also brings a new perspective to a common intuition associated with neuron dynamics. Specifically, it is typically understood that neurons operate near bifurcation points so that small perturbations can alter the neuron behavior substantially [11, 7], i.e. small voltage perturbations can cause the neuron to spike and fire. This concept is the basis for computationally efficient phenomenological models of single neuron responses [65, 66, 11, 67]. In the analysis presented here of the AMEE, it is not the individual neuron that matters. Indeed, dynamically speaking, individual neurons may be far from any bifurcation behavior. However, when coupled together in a network, even very weakly, the coupling can significantly shift the parameter space of the effective neuron dynamics as given by the AMEE. In particular, the coupling can shift the AMEE so that the network is near a bifurcation point. Ultimately, it is the network of neurons that should be near the bifurcation point so that the network can react to perturbations to the neural system. Thus when connected in a network, dynamically uninteresting individual neurons are capable of exhibiting key dynamical behaviors associated with network functionality. Although we do not consider noise driven dynamics, which are certainly relevant to networks of neurons, the AMEE identifies the deterministic parameter regimes that are potentially sensitive to such stochastic fluctuations.

For illustrative purposes, in many of the examples we have chosen a simple interaction between the neurons, via the voltage variable, in order to derive a succinct mean-field model. The success of this first step in modeling the bottom-up dynamics can lead to host of further considerations. To faithfully model common neural interactions one should take into account the synaptic variables that model the slow and fast synaptic receptors. We extend and explain the derivation of the AMEE to single channel synaptic interaction and demonstrate a possible resulting bifurcation study using such a model. Similarly, a detailed model consisting of neural coupling via several synaptic variables with HH model for individual dynamics may introduce further nontrivial evolution equations that exhibit effective complicated dynamics and are analytically intractable. Numerical studies of such a model can reveal the parameters and the mechanisms that lead to such behaviors of the network. Further, other hybrid models, stochastic systems and spatial dynamics may be potentially implemented within this mean-field approach. Such extensions to the current model are currently under consideration.

To contrast our approach with previous mean-field descriptions, we emphasize that the constructive assumption here is that the interacting network asymptotically converges to an attractor with no restriction on the topology of the attractor in phase space. In the phase reduction methodology, however, it is assumed that the individual neurons possess a limit cycle attractor and interactions are weak. The generality of our assumption results with being able to guarantee shadowing of the activity measure

27

dynamics by the solution of AMEE in contrast to an exact bound on the closeness between them in terms of the small parameter that multiplies the interaction terms. Furthermore, in the derivation of the AMEE, the size of the network is left as a free parameter such that both small and large networks can be considered. The small parameter that is used to guarantee shadowing is the size of the fluctuations around the underlying attractor which together with spatial coherence provides a low dimensional AMEE model even for a very large system. In the neural fields approach the small parameter is $1/N$, and the passage to the continuum limit provides the reduced model in terms of PDEs.

As a final comment, fundamental studies of neural network dynamics noted that the performance of circuits in the brain can change dramatically when the neurons in these circuits synchronize and thus it is important to understand the mechanisms that lead to synchronization [4, 38, 29]. These studies and following works investigate under what conditions populations of interacting neurons, evolve to a state in which there is an appreciable amount of synchrony. The theory developed here provides a potential methodology to address such issues in biophysically relevant settings.

Appendix A. Expansion of the AMEE. Expanding the right hand side of Eq. (3.1) with $v_{ik}(t) = x_{ik}^A(t) + \varepsilon \xi_{ik}(t)$ results in

$$\frac{dm_k(v_{ik})}{dt} = \sum_{i=1}^N \frac{d}{dt}\left(p_{ik}(x_{ik}^A) + \varepsilon \xi_{ik} \frac{dp_{ik}}{dx_{ik}}(x_{ik}^A) + O(\varepsilon^2) \right)$$

$$= \sum_{i=1}^N \left(\frac{dp_{ik}}{dx_{ik}}(x_{ik}^A)\frac{dx_{ik}^A}{dt} + \varepsilon \frac{d\xi_{ik}}{dt}\frac{dp_{ik}}{dx_{ik}}(x_{ik}^A) + \varepsilon \xi_{ik}\frac{d^2 p_{ik}}{d^2 x_{ik}}(x_{ik}^A)\frac{dx_{ik}^A}{dt} + O(\varepsilon^2) \right).$$

Perturbation around x_{ik}^A yields the following ODE system

$$\frac{dx_{ik}^A}{dt} + \varepsilon \frac{d\xi_{ik}}{dt} = g_{i,k}(x_{qr}^A) + \varepsilon \sum_{j=1}^N \sum_{l=1}^n \xi_{jl} \frac{dg_{i,k}}{dx_{jl}}(x_{qr}^A) + O(\varepsilon^2)$$

so that

$$\frac{dm_k}{dt} = \sum_{i=1}^N \left(\frac{dp_{ik}}{dx_{ik}}(x_{ik}^A)g_{i,k}(x_{qr}^A) + \varepsilon \frac{dp_{ik}}{dx_{ik}}(x_{ik}^A) \sum_{j=1}^N \sum_{l=1}^n \xi_{jl}\frac{dg_{i,k}}{dx_{jl}}(x_{qr}^A) \right.$$

$$\left. + \varepsilon \xi_{ik}\frac{d^2 p_{ik}}{d^2 x_{ik}}(x_{ik}^A)g_{i,k}(x_{qr}^A) + O(\varepsilon^2) \right).$$

Appendix B. Proof of Theorem 1. (a): Let $\eta_k(t) = m_k(v_{ik}(t)) - m_k(x_{ik}^A(t))$. Expansion of $m_k(v_{ik})$ results in

$$m_k(v_{ik}(t)) = m(x_{ik}^A(t) + \varepsilon \xi_{ik}(t)) = \sum_{i=1}^N \left[p_{ik}(x_{ik}^A(t)) + \varepsilon \xi_{ik}(t)\frac{dp_{ik}}{dx_{ik}}(x_{ik}^A(t)) \right] + h.o.t$$

since p_{ik} is linear, $dp_{ik}^j/d^jx_{ik} = 0$ for $j \geq 2$ and the h.o.t vanish. The remaining terms are

$$m_k(v_{ik}(t)) = \sum_{i=1}^{N} \left[p_{ik}(x_{ik}^A(t)) + \varepsilon\xi_{ik}(t)w_{ik} \right] = m(x_{ik}^A(t)) + \sum_{i=1}^{N} \varepsilon\xi_{ik}(t)w_{ik}.$$

Then $\eta_k(t)$ is of the form $\eta_k(t) = \varepsilon \sum_{i=1}^{N} \xi_{ik}(t)w_{ik}$. Since the weights are normalized and $\xi_{ik}(t) \sim \mathcal{O}(1)$, then $\eta_k(t)$ is guaranteed to be bounded by ε and we have proved that $m_k(v_{ik}(t))$ ε-shadows $m_k(x_{ik}^A(t))$.

(b): Similarly to the proof in part (a), $m_k(v_{ik})$ can be expanded using Taylor series such that

$$\eta_k(t) = \sum_{r=1}^{\infty} \sum_{i=1}^{N} \frac{\varepsilon^r\xi_{ik}(t)^r}{r!} \frac{d^r p_{ik}}{dx_{ik}^r}(x_{ik}^A(t))$$

The terms $\xi_{ik}(t)^r$ are bounded by $\mathcal{O}(1)$. Analyticity of p_{ik} and compactness of $\mathbb{X}$ ensure that the series on the right hand side will be convergent and hence can be bounded by a constant C_k. Therefore, $m(v_{ik}(t))$ is said to shadow $m(x_{ik}^A(t))$.

To prove ε-shadowing and the validity of the expansion in Eq. (3.3) we need to estimate the constant C_k to ensure that the distance between the two orbits is always bounded by ε and higher order terms. From the expansion, we conclude that if the convergence toward the attracting orbit $x_{ik}^A(t)$ is fast enough, in particular $\xi_{ik}(t)^r$ decays faster than the growth of the derivatives of p_{ik} ensuring that the sum over N for $r = 1$ is bounded by $\mathcal{O}(\varepsilon)$ and for $r \geq 1$ is bounded by $\mathcal{O}(\varepsilon^2)$, then for all $t > t^*$, $m(v_{ik}(t))$ ε-shadows $m(x_{ik}^A(t))$. Thus we decompose the expansion of $\eta_k(t)$ into a term expected to be of $\mathcal{O}(\varepsilon)$ and higher order terms in ε :

$$\eta_k(t) = \varepsilon \sum_{i=1}^{N} \xi_{ik}(t)\, p_{ik}^{(1)}\big(x_{ik}^A(t)\big) + \sum_{r=2}^{\infty} \frac{\varepsilon^r}{r!} \sum_{i=1}^{N} \xi_{ik}(t)^r p_{ik}^{(r)}\big(x_{ik}^A(t)\big)$$

where $p_{ik}^{(r)}\big(x_{ik}^A(t)\big) = \frac{d^r p_{ik}}{dx_{ik}^r}\big(x_{ik}^A(t)\big)$. For the validity of Eq. (3.3), the first term has to be bounded by ε: $\varepsilon \sum_{i=1}^{N} \xi_{ik}(t)\, p_{ik}^{(1)}\big(x_{ik}^A(t)\big) \sim \mathcal{O}(\varepsilon)$ resulting in the condition on $\xi_{ik}(t)$

$$\big|\xi_{ik}(t)\big| < \frac{1}{N\big|p_{ik}^{(1)}\big(x_{ik}^A(t)\big)\big|}.$$

that has to be satisfied for all t. In addition, we would like the remaining terms to be bounded by ε^2: $\sum_{r=2}^{\infty} \frac{\varepsilon^r}{r!} \sum_{i=1}^{N} \xi_{ik}(t)^r p_{ik}^{(r)}\big(x_{ik}^A(t)\big) \sim \mathcal{O}(\varepsilon^2)$ that will be satisfied with the additional conditions on $\xi_{ik}(t)$

$$\big|\xi_{ik}(t)\big|^r < \frac{r!}{N\varepsilon^{r-2}\big|p_{ik}^{(r)}\big(x_{ik}^A(t)\big)\big|}.$$

Appendix C. AMEE for a synchronous attractor for FHN network with identical neurons. From Eq. (4.3), the mean measure with equal weights $\nu_i = 1/N$ and $g_{i,k}$ as in Eq. (5.1) the following equations are received for the first order

$$\frac{dm_v(\bar{v})}{dt} = -\bar{w} + \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{m=0}^{p_1} \sum_{n=0}^{p_2} a_{ijmn}v_i^m v_j^n, \qquad \frac{dm_w(\bar{w})}{dt} = \dot{\bar{w}}(t) = -c\bar{v} + b\bar{w}.$$

Assumption of a synchronous attractor implies that for all i, j the variables v_i and v_j can be replaced by $\bar{v}$ and hence the resulting averaged model becomes

$$\frac{dm_v(\bar{v})}{dt} = -\bar{w} + \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{m=0}^{p_1}\sum_{n=0}^{p_2} a_{ijmn}\bar{v}^{m+n}, \qquad \frac{dm_w(\bar{w})}{dt} = -c\bar{v} + b\bar{w}.$$

Denoting $[A]_{mn}$ as $\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ijmn}$ brings the above equations to the form of Eqs. (5.3). For the second order, Eq. (4.4) results with the following set of equations

$$\frac{dm_v(\xi_v)}{dt} = -\bar{\xi}_w + \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}\xi_{lv}\frac{d}{dv_l}\left(\sum_{m=0}^{p_1}\sum_{n=0}^{p_2} a_{ijmn}v_i^m v_j^n\right),$$

$$\frac{dm_w(\xi_w)}{dt} = \dot{\bar{\xi}}_w(t) = -c\bar{\xi}_v + b\bar{\xi}_w$$

such that the the sum over l consists of two terms, $l = i$ and $l = j$

$$\frac{dm_v(\xi_v)}{dt} = -\bar{\xi}_w + \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{m=0}^{p_1}\sum_{n=0}^{p_2} a_{ijmn}(mv_i^{m-1}v_j^n\xi_{iv} + nv_i^m v_j^{n-1}\xi_{jv}),$$

$$\frac{dm_w(\xi_w)}{dt} = -c\bar{\xi}_v + b\bar{\xi}_w$$

replacing v_j and v_i by $\bar{v}$ results with

$$\frac{dm_v(\xi_v)}{dt} = -\bar{\xi}_w + \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{m=0}^{p_1}\sum_{n=0}^{p_2} a_{ijmn}(m\xi_{iv} + n\xi_{jv})\bar{v}^{m+n-1},$$

$$\frac{dm_w(\xi_w)}{dt} = -c\bar{\xi}_v + b\bar{\xi}_w.$$

Denoting $[A\vec{\xi}_v]_{mn} = \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ijmn}\xi_{iv}$ and $|A\vec{\xi}_v]_{mn} = \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ijmn}\xi_{jv}$ brings the above equations to the form of Eqs. (5.4).

REFERENCES

[1] G Baym. *Lectures on Quantum Mechanics*. Lecture Notes and Supplements in Physics. Perseus Books, Reading, Massachusetts, 1990.
[2] E H Lieb and R Seiringer. Proof of Bose-Einstein Condensation for Dilute Trapped Gases. *Phys. Rev. Lett.*, 88(17):170409, April 2002.
[3] L P Pitaevskii and S Stringari. *Bose-Einstein Condensation*. Clarendon Press, Oxford, 2003.
[4] L F Abbott and C van Vreeswijk. Asynchronous states in networks of pulse-coupled oscillators. *Physical Review E*, 48(2):1483–1490, 1993.
[5] S Coombes. Waves, bumps, and patterns in neural field theories. *Biological Cybernetics*, 93(2):91–108, 2005.
[6] S Grossberg. Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, 1(1):17–61, 1988.
[7] F C Hoppensteadt and E M Izhikevich. *Weakly connected neural networks*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1997.
[8] F Hoppensteadt. Intermittent Chaos, Self-Organization, And Learning From Synchronous Synaptic Activity In Model Neuron Networks. *Proceedings of the national academy of sciences of the united states of america*, 86(9):2991–2995, May 1989.
[9] T P Vogels, K Rajan, and L F Abbott. Neural network dynamics. *Annual Review Of Neuroscience*, 28:357–376, 2005.
[10] J P Keener and J Sneyd. *Mathematical Physiology*. PERSEUS-BOOKS, 2000.

[11] E M Izhikevich. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting (Computational Neuroscience)*. The MIT Press, 1 edition, November 2006.

[12] E M Izhikevich and G M Edelman. Large-scale model of mammalian thalamocortical systems. *Proc. Natl. Acad. of Sciences*, 105:3593–3598, 2008.

[13] N Kopell and G B Ermentrout. Phase transitions and other phenomena in chains of coupled oscillators. *SIAM J. Appl. Math.*, 50(4):1014–1052, 1990.

[14] R E Mirollo and S H Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM J. Appl. Math.*, 50(6):1645–1662, 1990.

[15] G B Ermentrout and D Terman. *Mathematical Foundations of Neuroscience*. Springer-Verlag, New York, 2010.

[16] Y Kuramoto. Phase- and center-manifold reductions for large populations of coupled oscillators with application to non-locally coupled systems. *International Journal of Bifurcation and Chaos*, 7(4):789–805, 1997.

[17] A T Winfree. Biological rhythms and behavior of populations of coupled oscillators. *Journal of Theoretical Biology*, 16(1):15–&, 1967.

[18] H R Wilson and J D Cowan. Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophys. J.*, 12(1), 1973.

[19] P C Bressloff and S Coombes. Physics of the extended neuron. *International Journal of modern physics b*, 11(20):2343–2392, August 1997.

[20] W Gerstner and W Kistler. *Spiking Neuron Models. Single Neurons, Populations, Plasticity.* Cambridge University Press New York, NY, USA, 2002.

[21] K Kishimoto and S Amari. Existence And Stability Of Local Excitations In Homogeneous Neural Fields. *Journal Of Mathematical Biology*, 7(4):303–318, 1979.

[22] P C Bressloff. Traveling fronts and wave propagation failure in an inhomogeneous neural network. *Phys. D*, 155(1-2):83–100, 2001.

[23] C R Laing and C C Chow. Stationary bumps in networks of spiking neurons. *Neural Computation*, 13(7):1473–1494, July 2001.

[24] S Coombes and M R Owen. Bumps, breathers, and waves in a neural network with spike frequency adaptation. *Physical Review Letters*, 94(14), April 2005.

[25] C R Laing. Spiral waves in nonlocal equations. *Siam Journal On Applied Dynamical Systems*, 4(3):588–606, 2005.

[26] P C Bressloff, J D Cowan, M Golubitsky, P J Thomas, and M C Wiener. Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. *Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences*, 356(1407):299–330, March 2001.

[27] G B Ermentrout and J B Mcleod. Existence And Uniqueness Of Traveling Waves For A Neural-Network. *Proceedings Of The Royal Society Of Edinburgh Section A-Mathematics*, 123(Part 3):461–478, 1993.

[28] N Brunel and V Hakim. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural computation*, 11(7):1621–71, October 1999.

[29] C van Vreeswijk and H Sompolinsky. Chaotic Balanced State in a Model of Cortical Circuits. *Neural Computation*, 10:1321–1371, 1998.

[30] J Guckenheimer and P Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields (Applied Mathematical Sciences Vol. 42)*. Springer, 1983.

[31] R FitzHugh. Impulses and Physiological States in Theoretical Models of Nerve Membrane. *Biophys. J.*, 1(6):445–466, 1961.

[32] C Morris and H Lecar. Voltage Oscillations in the Barnacle Giant Muscle-Fiber. *Biophysical Journal*, 35(1):193–213, 1981.

[33] E M Izhikevich. Synchronization of Elliptic Bursters. *SIAM Review*, 43(2):315–344, 2001.

[34] P Dayan and L F Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems.* The MIT Press, 2005.

[35] K F Wong and X J Wang. A recurrent network mechanism of time integration in perceptual decisions. *Journal Of Neuroscience*, 26(4):1314–1328, January 2006.

[36] P Eckhoff, K F Wong-Lin, and P Holmes. Optimality and Robustness of a Biophysical Decision-Making Model under Norepinephrine Modulation. *Journal of neuroscience*, 29(13):4301–4311, April 2009.

[37] A Treves. Mean-Field Analysis of Neuronal Spike Dynamics. *Network-Computation in Neural Systems*, 4(3):259–284, 1993.

[38] C VanVreeswijk and H Sompolinsky. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293):1724–1726, 1996.

[39] S H Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, March 2001.

[40] D A Wiley, S H Strogatz, and M Girvan. The size of the sync basin. *CHAOS*, 16(1), March

2006.

[41] V N Belykh, I V Belykh, and M Hasler. Connection graph stability method for synchronized coupled chaotic systems. *Physica D*, 195(1-2):159–187, 2004.

[42] S Coombes and M Zachariou. Gap junctions and emergent rhythms. In *Coherent Behaviour in Neuronal Networks*, pages 77–94. Springer, 2007.

[43] S Coombes. Neuronal networks with gap junctions: A study of piece-wise linear planar neuron models. *SIAM Journal on Applied Dynamical Systems*, 7(3):1101–1129, 2008.

[44] E Steur, I Tyukin, and H Nijmeijer. Semi-Passivity And Synchronization Of Diffusively Coupled Neuronal Oscillators. *Physica D-Nonlinear Phenomena*, 238(21):2119–2128, 2009.

[45] M Golubitsky and D G Schaeffer. *Singularities and Groups in Bifurcation Theory I*. Springer-Verlag, New York, 1985.

[46] M Golubitsky and I Stewart. Nonlinear dynamics of networks: the groupoid formalism. *Bulletin of the American Mathematical Society*, 43(03):305–365, 2006.

[47] M Golubitsky and I Stewart. Synchrony versus symmetry in coupled cells. *Network*, (x):1–12, 2003.

[48] M B Feller, D A Butts, H L Aaron, D S Rokhsar, and C J Shatz. Dynamic processes shape spatiotemporal properties of retinal waves. *NEURON*, 19(2):293–306, 1997.

[49] M B Feller. Spontaneous correlated activity in developing neural circuits. *NEURON*, 22(4):653–656, 1999.

[50] A Sarti, G Citti, and M Manfredini. From neural oscillations to variational problems in the visual cortex. *Journal of Physiology-Paris*, 97(2-3):379–385, 2003.

[51] D Hansel and H Sompolinsky. Chaos and synchrony in a model of a hypercolumn in visual cortex. *Journal of Computational Neuroscience*, 3(1):7–34, 1996.

[52] J A Riffell, H Lei, and J G Hildebrand. Neural correlates of behavior in the moth Manduca sexta in response to complex odors. *Proceedings of the national academy of sciences of the united states of america*, 106(46):19219–19226, 2009.

[53] S. Ahn, B. H. Smith, A. Borisyuk, and D. Terman. Analyzing neuronal networks using discrete-time dynamics. *Physica D*, 239(9):515–528, 2010.

[54] A Kumar, S Rotter, and A Aertsen. Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nature Reviews Neuroscience*, 11(9):615–627, 2010.

[55] R Ben-Yishai, D Hansel, and J Rinzel. Traveling waves and the processing of weakly tuned inputs in a cortical network module. *Journal of Computational Neuroscience*, 4:57–77, 1997.

[56] P Holmes, J Lumley, and G Berkooz. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press, Cambridge, 1996.

[57] L Sirovich. Modeling the functional organization of the visual cortex. *Physica D: Nonlinear Phenomena*, 96(1-4):355–366, September 1996.

[58] L Sirovich and R Everson. Management and Analysis of Large Scientific Datasets. *The International Journal Of Supercomputer Applications*, 6(1):50–68, 1992.

[59] N Kopell and G B Ermentrout. Mechanisms of Phase-Locking and Frequency Control in pairs of coupled Neural oscillators. *Handbook of Dynamical Systems edited by B. Fiedler (Elsevier, Amsterdam, 2002),*, 2:3–54, 2002.

[60] I. Belykh and A. Shilnikov. When Weak Inhibition Synchronizes Strongly Desynchronizing Networks of Bursting Neurons. *Physical Review Letters*, 101(7):1–4, August 2008.

[61] H. Sompolinsky, A. Crisanti, and HJ Sommers. Chaos in Random Neural Networks. *Physical Review Letters*, 61(3):259–262, 1988.

[62] A Destexhe, Z F Mainen, and Terrence J Sejnowski. Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *Journal of Computational Neuroscience*, 1(3):195–230, 1994.

[63] J Rubin and D Terman. Geometric analysis of population rhythms in synaptically coupled neuronal networks. *Neural computation*, 12(3):597–645, March 2000.

[64] M Leite and M Golubitsky. Homogeneous three-cell networks. *Nonlinearity*, 19(10):2313–2363, 2006.

[65] G . B . Ermentrout and N Kopell. Parabolic Bursting in an Excitable System Coupled with a Slow Oscillation. *SIAM Journal on Applied Mathematics*, 46(2):233–253, 2010.

[66] E M Izhikevich. Hybrid spiking models. *Phil. Trans. Roy. Soc. Lond. A*, 368:5061–5070, 2010.

[67] E Shlizerman and P Holmes. Neural dynamics, bifurcations and firing rates in a quadratic integrate-and-fire model with a recovery variable. I: deterministic behavior. *to appear in Neural Computation*, 2012.