Synchrony and Periodicity in Excitable Neural Networks with Multiple Subpopulations

Massachusetts Institute of Technology

Synchrony and Periodicity in Excitable Neural Networks with Multiple Subpopulations[*]

Lee DeVille[†] and Yi Zeng[‡]

Abstract. We consider a cascading model of excitable neural dynamics and show that over a wide variety of parameter regimes, these systems admit unique attractors. For large coupling strengths, this attractor is a limit cycle, and for small coupling strengths, it is a fixed point. We also show that the cascading model considered here is a mean-field limit of an existing stochastic model.

Key words. stochastic neuronal network, contraction mapping, mean-field limit, critical parameters

AMS subject classifications. 05C80, 37H20, 60B20, 60F05, 60J20, 82C27, 92C20

DOI. 10.1137/130943261

1. Introduction. The study of oscillator synchronization has made a significant contribution to the understanding of the dynamics of real biological systems [5, 7, 8, 11, 14, 15, 16, 17, 22, 24, 26, 27, 36, 37] and has also inspired many ideas in modern dynamical systems theory. See [28, 33, 40] for reviews. The prototypical model in mathematical neuroscience is a system of "pulse-coupled" oscillators, that is, oscillators that couple only when one of them "fires." More concretely, each oscillator has a prescribed region of its phase space where it is active, and only then does it interact with its neighbors. There is a large body of work on deterministic pulse-coupled networks [4, 6, 13, 18, 19, 24, 27, 31, 34, 35, 38, 39], mostly studying the phenomenon of synchronization on such networks.

In [9, 10], the first author and collaborators considered a specific example of a network containing both refractoriness and noise; the particular model was chosen to study the effect of *synaptic failure* on the dynamics of a neuronal network. What was observed in this class of models is that when the probability of synaptic success was taken to be small, the network acted as a stationary process with a low degree of correlation in time; when the probability of synaptic success was taken to be large, the system exhibited synchronous behavior that was close to periodic. Both of these behaviors are, of course, expected: strong coupling tends to lead to synchrony, and weak coupling tends not to do so. The most interesting observation was that for intermediate values of the coupling, the network could support both synchronized and desynchronized behaviors and would dynamically switch between the two. This was explained in [10] by showing that the large network limit was, for certain parameters, multistable. Then,

large but finite networks switch stochastically between the attractors of the limiting system.

One unusual aspect of the large network limit, or mean-field system, of [9, 10] is that it is a *hybrid* system: a system of a continuous flow coupled to a map of the phase space. This system has piecewise continuous trajectories that jump at prescribed times. This is due to the fact that the interneuronal connections in the model undergo cascades, where the firing of one neuron can cause other neurons to fire, which causes other neurons to fire, and so on, causing an avalanche of activity throughout the network. This sort of neural activity has been observed experimentally in [1, 2, 32], and a model of the type considered in this paper was matched to experimental data in [12]. Since these cascading events are on the order of the size of the network, yet happen quickly, they correspond to discontinuities in the dynamics, leading to the hybrid character of the dynamics. Moreover, as we argue below, the model we consider here is a prototypical model of cascading neuronal dynamics and is in some sense the simplest model possible of this type. The model we analyze here is a cascading version of the three-state excitable network model analyzed in [25, 29, 30] (although one notes that the details of the analysis differ significantly).

In this paper, we consider a generalization of the mean-field model that allows for several independent subpopulations with different intrinsic firing rates. We show that this model has the property that for sufficiently small interneuronal coupling, the system has a globally attracting fixed point, and for sufficiently large interneuronal coupling, the system has a globally attracting periodic orbit. We also give bounds on the parameter ranges of validity for each of the two behaviors. Moreover, we make the surprising observation that all of these attractors exist no matter how many subpopulations exist and how much their firing rates differ; in particular, we show that the critical coupling parameter for the existence of a globally attracting limit cycle does not depend on the firing rates, or relative sizes, of the subpopulations in the network.

We also connect the model studied in this paper to the stochastic cascading neural system considered in [10]. Since this result follows with only minor changes from the theorems in [10], we present a short argument on the connection between the stochastic model and its mean-field equation, but only for completeness.

2. Model definition.

2.1. Overview of model.
We consider a network of neurons which is coupled all-to-all and in which all coupling is excitatory. We also assume that the interneuronal coupling is much faster than the other timescales in the system, so that the interaction between different neurons happens in zero time. Each neuron can be in one of three states: "refractory," "excitable," or "firing." Every refractory neuron will need an input to become excitable and then takes one more input to fire. We also assume that neurons have variable firing rates.

We will assume that there is a finite number M of subpopulations of neurons and that different subpopulations have different firing rates; we will denote the fraction of neurons in subpopulation m by α_m, and the firing rate of these neurons will be denoted ρ_m. For shorthand, we will say that the refractory neurons are at level 0, and the excitable are at level 1. We use the index $k = 0, 1$ to denote the state of a given neuron and $m = 1, \ldots, M$ to index certain subpopulations. Thus we will denote the proportion of neurons of type m that are refractory by $x_{0,m}$ and the proportion that are excitable by $x_{1,m}$.

The interneuronal coupling will be determined by the parameter β. The interpretation of β is that whenever there is a proportion of neurons that are firing, they will promote a fraction of neurons in the network through synaptic connections, and β represents the ratio of neurons being promoted to those currently firing. We also assume that whenever neurons fire, we compute the entire cascade of firing until there are no longer any firing neurons. Moreover, we assume that all neurons that fire are set to the refractory state at the end of the cascade. Thus, the entire state of the network will be determined by the vector $x_{k,m}$ with $k = 0, 1$ and $m = 0, \ldots, M$, as all of the firing neurons will be processed as soon as they fire.

2.2. Mathematical definition of model. Choose a natural number M. Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_M)$ be any vector with $0 < \alpha_m < 1$, and $\sum_m \alpha_m = 1$, and let $\boldsymbol{\rho} \in (\mathbb{R}^+)^M$. The domain of our dynamical system will be

$$D^{\boldsymbol{\alpha}} := \left\{ x = \{x_{k,m}\} \in \mathbb{R}^{2M} \,|\, x_{0,m} + x_{1,m} = \alpha_m \right\}.$$

We write $y_k = y_k(x) := \sum_m x_{k,m}$ and write $D^{\boldsymbol{\alpha}}$ as the disjoint union $D^{\boldsymbol{\alpha}} = D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta} \dot{\cup} D_G^{\boldsymbol{\alpha},\beta}$, where

$$D_G^{\boldsymbol{\alpha},\beta} := \{x \in D^{\boldsymbol{\alpha}} \,|\, \beta y_1 \geq 1\}, \quad D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta} = D^{\boldsymbol{\alpha}} \setminus D_G^{\boldsymbol{\alpha},\beta}.$$

We will also write

$$\partial D_G^{\boldsymbol{\alpha},\beta} = \{x \colon y_1 = \beta^{-1}\}.$$

We now define a deterministic hybrid dynamical system $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(t)$ with state space $D^{\boldsymbol{\alpha}}$. The system will be hybrid since it will have two different rules on complementary parts of the phase space.

Definition 2.1 (definition of $\mathcal{L}$). *Consider the flow defined by*

$$(2.1) \qquad \frac{d}{dt}\xi_{k,m}(t) = \rho_m \mu(\xi)(\xi_{k-1,m}(t) - \xi_{k,m}(t)),$$

where $\mu(\xi)$ is the scalar *function*

$$\mu(\xi) = \frac{1}{1 - \beta y_1} = \frac{1}{1 - \beta \sum_m x_{1,m}},$$

and we interpret indices modulo 2. More compactly, define the matrix $\mathcal{L}$ by

$$(2.2) \qquad \mathcal{L}_{(k,m),(k',m')} = \delta_{m,m'}(-1)^{1+k+k'}\rho_m,$$

and (2.1) can be written $\dot{\xi} = \mu(\xi)\mathcal{L}\xi$.

Definition 2.2 (definition of G). *Let us now index $\mathbb{R}^{2M+1}$ by (k, m) with $k = 0, 1$, $m = 1, \ldots, M$, and a state that we denote as Q. Define the matrix $\mathcal{M}$ whose components are given by*

$$\mathcal{M}_{z,z'} = \begin{cases} -1, & z = (0, m), z' = (0, m), \\ 1, & z = (0, m), z' = (1, m), \\ 1, & z = (1, m), z' = Q, \\ 0 & else. \end{cases}$$

Define P_z as projection onto the zth coordinate, and

$$s_\star^\beta(\xi) = \inf_{s>0}\left\{s\Big|P_Q\left(e^{s\beta\mathcal{M}}\xi\right) = s\right\},$$

and then define $G^{\boldsymbol{\alpha},\beta}(\xi)$ componentwise by

$$
\begin{aligned}
(2.3) \qquad P_{(1,m)}(G^{\boldsymbol{\alpha},\beta}(\xi)) &= P_{(1,m)}(e^{\beta s_\star^\beta(\xi)\mathcal{M}}\xi), \\
P_{(0,m)}(G^{\boldsymbol{\alpha},\beta}(\xi)) &= \alpha_m - P_{(1,m)}(G^{\boldsymbol{\alpha},\beta}(\xi)).
\end{aligned}
$$

Definition 2.3 (definition of full system). *We combine the above to define a hybrid system for all $t > 0$. In short, the system uses flow given by $\mathcal{L}$ on the domain $D_\mathcal{L}^{\boldsymbol{\alpha},\beta}$, and if the system ever enters the domain $D_G^{\boldsymbol{\alpha},\beta}$, it immediately applies the map G.*

More specifically, fix $\boldsymbol{\rho}, \beta$. Define the flow map $\varphi(x,t)$ by

$$\frac{d}{dt}\varphi(\xi,t) = \mu(\xi)\mathcal{L}\xi, \quad \varphi(\xi,0) = \xi.$$

Assume $\xi(0) \in D_\mathcal{L}^{\boldsymbol{\alpha},\beta}$, and let

$$(2.4) \qquad \tau_1 = \inf_{t>0}\{\varphi(\xi(0),t) \in D_G^{\boldsymbol{\alpha},\beta}\}.$$

We then define

$$\xi(t) = \varphi(\xi(0),t) \text{ for } t \in [0,\tau_1), \quad \xi(\tau_1) = G(\varphi(\xi(0),\tau_1)).$$

(Of course, it is possible that $\tau_1 = \infty$, in which case we have defined the system for all positive time; otherwise we proceed recursively.) Now, given $\tau_n < \infty$ and $\xi(\tau_n) \in D_\mathcal{L}^{\boldsymbol{\alpha},\beta}$, define

$$(2.5) \qquad \tau_{n+1} = \inf_{t>\tau_n}\{\varphi(\xi(\tau_n),t-\tau_n) \in D_G^{\boldsymbol{\alpha},\beta}\}$$

and

$$\xi(t) = \varphi(\xi(\tau_n),t-\tau_n) \text{ for } t \in [\tau_n,\tau_{n+1}), \quad \xi(\tau_{n+1}) = G(\varphi(\xi(\tau_n),\tau_{n+1}-\tau_n)).$$

If $\tau_n = \infty$, then we define $\tau_{n+1} = \infty$ as well. We call the times $\tau_1, \tau_2, \ldots$ the big burst times, *and we call $s_\star^\beta(\xi(\tau_n))$ the* size of the big burst.

Remark 2.4. We note that the definition given above is well defined and gives a unique trajectory for $t \in [0,\infty)$ iff we know that $G(\xi) \in D_\mathcal{L}$ for any $\xi \in D_G$. We will show below that this is the case. We will also see below that some trajectories have infinitely many big bursts, and some have finitely many; this depends both on parameters and initial conditions.

2.3. Intuition behind definition. This is no doubt a complicated description, but all of the pieces of this definition can be well motivated. We give an intuitive description of this justification now and make a connection to a stochastic model in section 2.4 below.

First consider an infinite network where each parent neuron firing gives rise to an independent random number of children neurons firing, and the expected number of children per parent is β. Then assume that when a neuron fires, we recursively compute all descendants

of this initial neuron until the activity dies away. The expected number firing in the first generation is β, and the expected number firing in the second generation is β^2, etc. It is clear that the expected number of neurons that fire in the entire cascade is $\sum_{\ell=0}^{\infty} \beta^{\ell} = (1 - \beta)^{-1}$ if this sum converges, i.e., if $\beta < 1$. Let us call this $\beta < 1$ case *subcritical*. If $\beta > 1$, then the expected size of the cascade is infinite, and let us call this case *supercritical*.

Now consider the network defined above. Notice that a neuron is only primed to fire if it is excitatory, and the total proportion of excitatory neurons is y_1. Thus, when a neuron fires, the proportion of neurons that are "available" is given by y_1, and the average number of children per parent is βy_1, and so we should consider the case $\beta y_1 < 1$ as subcritical, and the cascade size is $(1 - \beta y_1)^{-1}$. This means that the "force multiplier" of each neuron is $\mu(\xi)$ as defined above, by which we mean if an infinitesimal fraction of neurons enters the firing state, then the total size of the burst that arises should be $\mu(\xi)$ times this fraction. With β positive but subcritical, this just "speeds up" the ODE by the multiplicative factor $\mu(\xi)$.

If the state ξ is supercritical ($\beta y_1 > 1$), then the above argument implies an infinite cascade. However, notice that y_1 will evolve during the cascading process as neurons are drawn from the excitatory state into the firing state. To model this, we should consider a system where neurons in the queue are being processed and thrown away at rate 1, and this induces neurons to move from refractory to excitatory at rate β times the proportion that are refractory, and from excitatory to firing at rate β times the proportion that are excitatory. But notice the definition of $\mathcal{M}$: this is exactly what happens as the system evolves, and we stop the system when the proportion of neurons in the queue is equal to the time that we have evolved—which is of course equivalent to saying that if we are removing neurons from the queue at constant rate 1, then it is the first time the queue is empty. Then, all of the neurons that have fired are then reset to be refractory, which is the same as saying that they are reinjected at level zero.

2.4. Connection to stochastic model. We now present a stochastic neuronal network model that generalizes the one considered in [9, 10]. The model has N neurons, each of which has an intrinsic firing rate ρ_n. Each neuron can be in one of three states: "quiescent," "excitable," and "firing," which we denote as levels 0,1,2.

If there are no neurons firing, we promote the nth neuron in the network with rate ρ_n; i.e., we choose N independent random times T_n, where T_n is exponentially distributed with rate ρ_n, and define

$$T = \min_n T_n, \quad n^{\star} = \arg\min_n T_n,$$

and then we promote neuron $n^{\star}$ by one level and increment the time variable by T.

If there are neurons firing, we compute the effect of a cascade as follows: for each neuron in the firing queue, we promote each other neuron in the network, independently, with probability p. If any neurons are raised to the level of firing, we add them to the queue, and we continue this process until the firing queue is empty. Note that the probability of any neuron promoting any other neuron is the same, so it will not matter how we process the neurons in the queue (First In First Out (FIFO), Last In First Out (LIFO), etc.). However, if a neuron fires in a given burst, we temporarily remove it from the population until the burst is completed and then reinsert all of the neurons that have fired back to the quiescent state. This is a type of refractoriness in that no neuron can fire more than once in a burst.
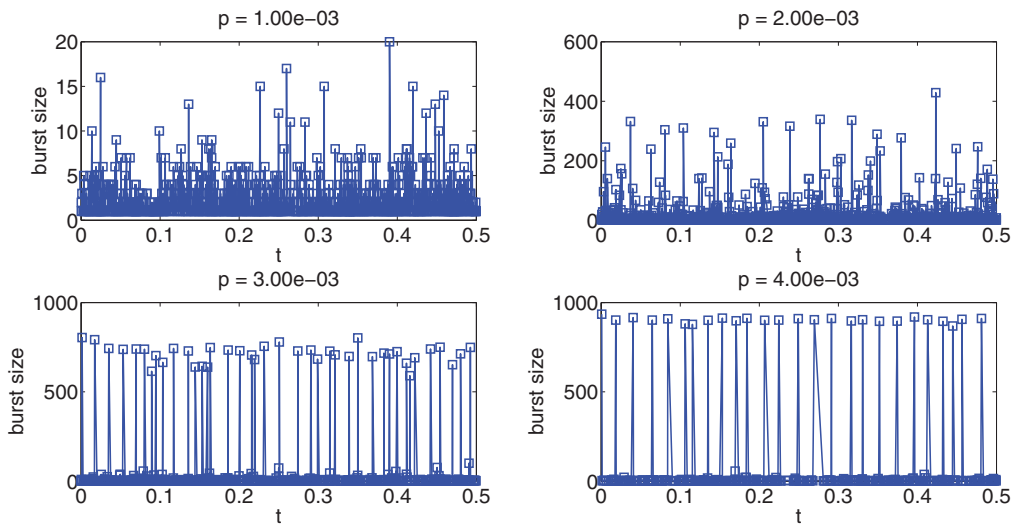
Figure 1. *Different behaviors of the model. We fix $M = 10$ and $N = 1000$ and plot different dynamics of the model that correspond to different p. As we increase p, we see the change from asynchronous and irregular behavior to synchronous and periodic behavior.*

Clearly, all of the interneuronal coupling in this model is through the parameter p. The larger the value of p, the more tightly coupled the system is. What has been observed for models of this type [9, 10, 12] is that when p is small, the typical event size in the system is small, and the system is incoherent; conversely, when p is large, the system is synchronous and periodic (see Figure 1 for an example, but see other references for more detail).

We can now consider a limit as $N \to \infty$ for this system. Choose M a natural number and $\boldsymbol{\alpha}, \boldsymbol{\rho}$ as defined in the system above, i.e., $\rho_m > 0$ for all m, $0 < \alpha_m < 1$ for all m, and $\sum_m \alpha_m = 1$. For each N, define a partition of N into M disjoint sets, denoted by $A_m^{(N)}$, and require that the firing rate of every neuron in $A_m^{(N)}$ be ρ_m. As $N \to \infty$, assume that $||A_m^{(N)}| - \alpha_m N| < 1$ for all m. (Note that $\alpha_m N$ is not in general an integer, but we require that $|A_m^{(N)}|$ be as close to this number as possible.)

It is not hard to see that the description defines a stochastic process with parameters $N, \boldsymbol{\alpha}, \boldsymbol{\rho}, p$, which we will denote as $X_t^{N,\boldsymbol{\alpha},\boldsymbol{\rho},p}$ below. In the limit $N \to \infty$, we will state the convergence theorem of the stochastic neuronal network to a mean-field limit; the proof given there will work with some technical changes.

Theorem 2.5. *Consider any $x \in D^{\boldsymbol{\alpha}} \cap \mathbb{Q}^{2M}$. For N sufficiently large, Nx has integral components, and we can define the neuronal network process $X_t^{N,\boldsymbol{\alpha},\boldsymbol{\rho},p}$ as above, with initial condition $X_0^{N,\boldsymbol{\alpha},\boldsymbol{\rho},p} = Nx$.*

Choose and fix $\epsilon, h, T > 0$. Let $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(t)$ be the solution to the mean field defined in Definition 2.3 with initial condition $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(0) = x$. Define the times $\tau_1, \tau_2, \dots$ at which the mean field jumps, and define $b_{min}(T) = \min\{s_\star^\beta(\xi(\tau_k)) : \tau_k < T\}$, i.e., b_{min} is the size of the smallest big burst which occurs before time T, and let $m(T) = \arg\max_k \tau_k < T$, i.e., $m(T)$ is the number of big bursts in $[0, T]$.

Pick any $\gamma < b_{min}(T)$. For the stochastic process $X_t^{N,\boldsymbol{\alpha},\boldsymbol{\rho},p}$, denote by $T_k^{(N)}$ the (random) times at which the $X_t^{N,\boldsymbol{\alpha},\boldsymbol{\rho},p}$ has a burst of size larger than γN. Then there exist $C_{0,1}(\epsilon) \in [0,\infty)$ and $\omega(M) \geq 1/(5M)$ such that for N sufficiently large,

$$(2.6) \qquad \mathbb{P}\left(\sup_{j=1}^{m(T)} \left|T_j^{(N)} - \tau_j\right| > \epsilon\right) \leq C_0(\epsilon)Ne^{-C_1(\epsilon)N^{\omega(M)}}.$$

Moreover, if we define $\mathcal{T} := ([0,T] \setminus \cup_{j=1}^{m(T)}(T_j^{(N)} - \epsilon, T_j^{(N)} + \epsilon))$, and

$$\varphi(t) = t - (T_j^{(N)} - \tau_j), \;\; where \; j = \max\{k \colon \tau_k < t\},$$

then

$$(2.7) \qquad \mathbb{P}\left(\sup_{t \in \mathcal{T}} \left|N^{-1}X_t^{K,\boldsymbol{\alpha},\boldsymbol{\rho},p} - \xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(\varphi(t))\right| > \epsilon\right) \leq C_0(\epsilon)Ne^{-C_1(\epsilon)N^{\omega(M)}}.$$

In summary, the theorem has two main conclusions about what happens if we consider a stochastic neuronal network with N large. The first is that (up to some technical details) the stochastic system is a fluctuation around the mean-field system when N is sufficiently large. Recalling Figure 1 again, we will show below that the mean-field system has an attracting fixed point for β sufficiently small, and the incoherent dynamics for small p correspond to fluctuations around this fixed point. Conversely, we show that for β sufficiently large, the mean-field system has a limit cycle, and the periodic dynamics for large p correspond to fluctuations around this limit cycle.

In Figure 2, we numerically show the convergence result in another way: in dark blue, we plot the mean and standard deviation of the sizes of burst in the stochastic model, and in red we plot the corresponding quantity in the mean-field model, the function $s_\star(\beta)$ defined in Lemma 3.4 below. We see that they match well even for $N = 1000$.

The guaranteed rate of convergence is subexponential due to the presence of the $\omega(M)$ power in the exponent, but note that the convergence is asymptotically faster than any polynomial. Numerical simulations done for the case of $M = 1$ were reported in [9] and show that $\omega(1)$ seemed to be close to 1, and this closeness was uniform in K. This suggests that the lower bound is pessimistic and that the convergence may in fact be exponential. However, the lower bound given in the theorem above seems to be the best that can be achieved by the authors' method of proof. For the details comprising a complete proof of Theorem 2.5, see [10].

3. Main theorem and analysis. The main result of this paper is to prove that for any M, $\boldsymbol{\alpha}$, and $\boldsymbol{\rho}$, then for β sufficiently small, the system has a globally attractive fixed point, and for β sufficiently large, the system has a globally attracting periodic orbit.

It should be noted that there is no clear a priori method of analyzing the stability of the model considered here. As is well known, the analysis of hybrid systems can be exceedingly complicated [3, 21]; questions just about the stability of fixed points are much more complicated than in the nonhybrid (flow or map) case, and stability of periodic orbits is more complicated still. As we see below, the state-of-the-art technique for this kind of problem is very problem-specific; in general, one contrives to construct some sort of Lyapunov function for the system, and this is what we are able to do here.
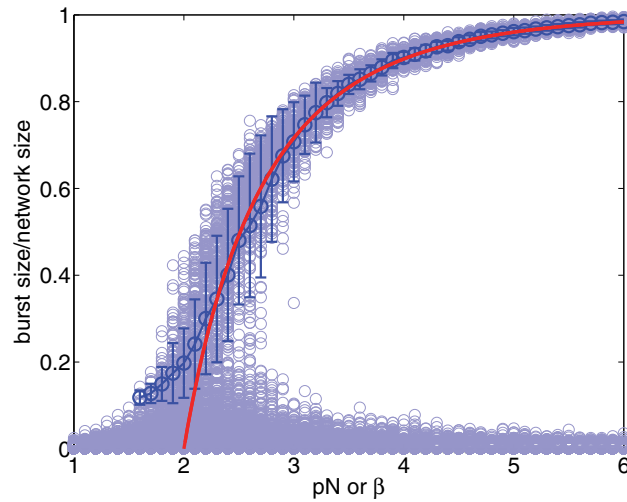
Figure 2. *The meaning of the blue data: We fix a choice of $\boldsymbol{\alpha}$, and $N = 1000$; then we run the stochastic neuronal network described in this section. We plot the burst sizes in light blue. For p large enough, we also plot the mean and standard deviations of the burst sizes for all of the bursts larger than one-tenth the size of the network. In red, we plot the deterministic burst size (as a proportion of network size) in the deterministic limit defined in section 3.2 (in fact, we are plotting the function $s_\star(\beta)$ defined in Lemma 3.4). The result of Theorem 2.5 is that the dark blue circles lie on the red curve, and the error bars get small, as $N \to \infty$. The numerics seem to verify this.*

3.1. Main result.
We now state the main result of the paper.

Theorem 3.1. *Choose and fix $M, \boldsymbol{\alpha}, \boldsymbol{\rho}$, and consider the hybrid system $\xi^{\boldsymbol{\alpha}, \boldsymbol{\rho}, \beta}(t)$ defined in Definition 2.3. Then the following hold:*

- *For $\beta < 2$ and all M, the system has a globally attracting fixed point $\xi_{\mathsf{FP}}^{\boldsymbol{\alpha}, \boldsymbol{\rho}, \beta}$.*
- *For any $M \geq 1$, there exists $\beta_M \geq 2$ such that, for $\beta > \beta_M$, the hybrid system has a globally attracting limit cycle $\xi_{\mathsf{LC}}^{\boldsymbol{\alpha}, \boldsymbol{\rho}, \beta}(t)$. This orbit $\xi_{\mathsf{LC}}^{\boldsymbol{\alpha}, \boldsymbol{\rho}, \beta}(t)$ undergoes infinitely many big bursts. Moreover, $\limsup_{M \to \infty} \beta_M / \log(\sqrt{M}) \leq 1$.*

We delay the formal proof of the main theorem until after we have stated and proved all of the auxiliary results below, but we give a sketch here.

The main analytic technique we use is a contraction mapping theorem, and we prove this in two parts. We first show that for any two initial conditions, the flow part of the system stretches the distance between them by no more than $1 + \sqrt{M}/2$ (Theorem 3.11). We then show that the map G_β is a contraction, and, moreover, its modulus of contraction can be made as small as desired by choosing β large enough (Theorem 3.14). The stretching modulus of one "flow, map" step of the hybrid system is the product of these two numbers, and as long as this is less than one, we have a contraction. Finally, we also show that for $\beta > 2$, there exists an orbit with infinitely many big bursts (Lemma 3.7); in fact, we show the stronger result that all initial conditions give an orbit with infinitely many big bursts. All of this together, plus the compactness of the phase space, implies that this orbit is globally attracting.

We point out that several parts of the argument that seem straightforward at first glance are actually nontrivial for a few reasons.

First, consider the task of computing the growth rate for the flow part of the hybrid system. Clearly $e^{t\mathcal{L}}$ is a contraction, since its eigenvalues are

$$\{0^M, -2\rho_1, -2\rho_2, \ldots, -2\rho_M\},$$

and the vector in the null space is unique once $\boldsymbol{\alpha}$ is chosen. (We use the notation 0^M to denote M repeated eigenvalues at 0.) However, even though the linear flow $e^{t\mathcal{L}}$ is contracting, and clearly $\left| e^{t\mathcal{L}}x - e^{t\mathcal{L}}x' \right| < |x - x'|$ for any fixed $t > 0$, the difficulty is that two different initial conditions can flow for a different interval of time until the first big burst, and clearly we cannot guarantee that $e^{t\mathcal{L}}x$ and $e^{t'\mathcal{L}}x'$ are close at all. For example, consider the extreme case where the flow $e^{t\mathcal{L}}x$ hits the set $D_G^{\boldsymbol{\alpha},\beta}$ at some finite time, and the flow $e^{t\mathcal{L}}x'$ never does; then these trajectories can end up arbitrarily far apart, regardless of the spectrum of $\mathcal{L}$. For both of these reasons, we cannot simply use the spectral analysis of $\mathcal{L}$ for anything useful and have to work harder at establishing a uniform contraction bound.

Moreover, we point out another subtlety of hybrid systems, which is that the composition of two stable systems is not stable in general. In fact, establishing stability properties for hybrid systems, even when all components are stable and linear, is generally a very nontrivial problem (see, for example, [20]). We get around this by showing the subsystems are each contractions (i.e., we show that $\|\cdot\|_2$ is a strict Lyapunov function for the system), but needing to control every potential direction of stretching adds complexity to the analysis.

3.2. Intermediate results. This section lists several intermediate results that we now quickly summarize. In Lemma 3.2 we show that we can in practice ignore the scalar function $\mu(\xi)$. In Lemma 3.3, we show that the size of the big burst can be written as the root of a certain analytic function. In Lemma 3.4, we show that for fixed parameters, the size of the big burst is independent of when we enter $D_G^{\boldsymbol{\alpha},\beta}$, and we derive some of the properties of the size of the big burst when parameterized by β.

Lemma 3.2. *Recall* (2.1), *written as*

$$\frac{d\xi}{dt} = \mu(\xi)\mathcal{L}\xi.$$

If we replace the scalar function $\mu(\xi)$ with any constant, this does not affect the trajectories of the hybrid system whatsoever (although it does affect the speed at which they are traced).

Proof. Since $\mu(\xi(t))$ is a scalar function of time, we can remove it by the time change $\tau = \mu(\xi)t$, and then we have

$$\frac{d}{d\tau} = \mathcal{L}\xi.$$

Clearly this does not affect the trajectories of the hybrid system and thus will not affect any of the conclusions of Theorem 3.1. Thus w.l.o.g. we will drop μ below. ∎

Since the flow has the form

$$(3.1) \qquad \frac{d}{dt}\begin{pmatrix} x_{0,m} \\ x_{1,m} \end{pmatrix} = \rho_m \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_{0,m} \\ x_{1,m} \end{pmatrix},$$

the solution is

$$(3.2) \qquad x_{1,m}(t) = \frac{\alpha_m}{2} + \frac{x_{1,m}(0) - x_{0,m}(0)}{2}e^{-2\rho_m t} = \frac{\alpha_m}{2} - \left(\frac{\alpha_m}{2} - x_{1,m}(0)\right)e^{-2\rho_m t},$$

and of course $x_{0,m}(t) = \alpha_m - x_{1,m}(t)$.

Lemma 3.3. *If we define*

$$
(3.3) \qquad
\begin{aligned}
\psi^\beta(x,s) &= -s + \sum_{m=1}^M x_{1,m}\left(1 - e^{-s\beta}\right) + \sum_{m=1}^M x_{0,m}\left(1 - e^{-s\beta} - s\beta e^{-s\beta}\right) \\
&= -s + y_1\left(1 - e^{-s\beta}\right) + y_0\left(1 - e^{-s\beta} - s\beta e^{-s\beta}\right),
\end{aligned}
$$

then

$$
s_\star^\beta(x) = \inf_{s>0} \psi^\beta(x,s).
$$

Proof. If $\dot z = \beta \mathcal{M} z$, then writing this in coordinates gives

$$
\dot z_Q = \beta \sum_{m=1}^M z_{1,m}, \quad \dot z_{1,m} = \beta(z_{0,m} - z_{1,m}), \quad \dot z_{0,m} = -\beta z_{0,m}.
$$

One can compute directly that

$$
z_Q(s) = \sum_{m=1}^M z_{1,m}(0)\left(1 - e^{-s\beta}\right) + \sum_{m=1}^M z_{0,m}(0)\left(1 - e^{-s\beta} - s\beta e^{-s\beta}\right),
$$

and thus $z_Q(s) = s$ iff $\psi^\beta(z,s) = 0$. The remainder follows from the definition of $s_\star^\beta$. ∎

Lemma 3.4. $s_\star^\beta(x)$ *is constant on* $\partial D_G^{\alpha,\beta}$, *and its value depends only on* β. *We write* $s_\star(\beta)$ *for its value on this set.* $s_\star(\beta)$ *is an increasing function of* β, *and*

$$
\lim_{\beta \to \infty} s_\star(\beta) = 1.
$$

Proof. We see from (3.3) that $\psi^\beta(x,s)$, and thus $s_\star^\beta(x)$, depend on x only through the sums y_0 and y_1. By definition y_0 and y_1 are constant on $\partial D_G^{\alpha,\beta}$, and therefore $s_\star^\beta(\cdot)$ is as well. On $\partial D_G^{\alpha,\beta}$, $y_0 = (\beta-1)/\beta$ and $y_1 = 1/\beta$, so on this set we can ignore x and simplify ψ to

$$
(3.4) \qquad \psi^\beta(s) = 1 - s - e^{-s\beta} - \frac{\beta-1}{\beta} s\beta e^{-s\beta} = 1 - s - ((\beta-1)s+1)e^{-s\beta}.
$$

It follows from this formula that

$$
\psi^\beta(0) = 0, \quad \psi^\beta(1) = -\beta e^{-\beta} < 0, \quad \frac{d\psi^\beta}{ds}(0) = 0, \quad \frac{d^2\psi^\beta}{ds^2}(0) = \beta(\beta-2).
$$

If $\beta < 2$, then $\psi^\beta(s)$ is negative for some interval of s around zero, and thus $s_\star(\beta) = 0$. If $\beta > 2$, then the graph $\psi^\beta(s)$ is tangent to the x-axis at $(0,0)$ but is concave up, and thus positive for some interval of s around zero, and therefore $s_\star(\beta) > 0$. Since $\psi^\beta(1) < 0$, it is clear that $s_\star(\beta) < 1$. Taking β large, we see that $\psi^\beta(s) \approx 1 - s$, so that $s_\star(\beta) \approx 1$ for β large.

Finally, thinking of $\psi^\beta(s)$ as a function of both s and β, we have

$$
\frac{\partial}{\partial s}\psi^\beta(s) = e^{-s\beta}\left(1 - e^{s\beta} + \beta(\beta-1)s\right), \quad \frac{\partial}{\partial \beta}\psi^\beta(s) = e^{-\beta s}(\beta-1)s^2.
$$

Since the second derivative of $e^{s\beta}\partial\psi^\beta/\partial s$ is always negative, this means that $\partial\psi^\beta/\partial s$ can have at most two roots, and one of them is at $s = 0$. From the fact that $\psi^\beta(s)$ is concave up at zero, this means that the single positive root of $\partial\psi^\beta/\partial s$ is strictly less than $s_\star(\beta)$. From this it follows that $\partial\psi^\beta/\partial s|_{s=s_\star(\beta)} > 0$. It is clear from inspection that $\partial\psi^\beta/\partial\beta|_{s=s_\star(\beta)} < 0$, and from this and the implicit function theorem, we have $\partial s_\star/\partial\beta > 0$. ∎

Remark 3.5. By definition, a big burst occurs on the set $D_G^{\boldsymbol{\alpha},\beta}$, where $y_1 \geq \beta^{-1}$. Since the flow has continuous trajectories, it must enter $D_G^{\boldsymbol{\alpha},\beta}$ on the boundary $\partial D_G^{\boldsymbol{\alpha},\beta}$, and note that on this set, formula (3.4) is valid.

We can further simplify the formula for $G^{\boldsymbol{\alpha},\beta}$ as follows:

(3.5)
$$G_{0,m}^{\boldsymbol{\alpha},\beta}(x) = \alpha_m - e^{-\beta s_\star^\beta(x)}(\beta s_\star^\beta(x)x_{0,m} + x_{1,m}),$$
$$G_{1,m}^{\boldsymbol{\alpha},\beta}(x) = e^{-\beta s_\star^\beta(x)}(\beta s_\star^\beta(x)x_{0,m} + x_{1,m}).$$

Note that different subpopulations are coupled only through $s_\star^\beta(x)$.

3.3. Infinitely many big bursts. In this section, we show that for $\beta > 2$, all orbits of $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(t)$ have infinitely many big bursts.

Lemma 3.6.
$$G^{\boldsymbol{\alpha},\beta}: D_G^{\boldsymbol{\alpha},\beta} \to D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta}.$$

Proof. Let $x \in D_G^{\boldsymbol{\alpha},\beta}$, and consider the flow $\dot{z} = \beta\mathcal{M}z$, with $z(0) = x$. Since $dz_Q(s)/ds > 0$ and $z_Q(0) = 0$, we have that $z_Q(s) > 0$ for $s \in [0, s_\star^\beta(x))$. This means that

$$\frac{d}{ds}(z_Q(s) - s) < 0,$$

or

$$1 > \frac{d}{ds}z_Q(s) = \beta\sum_{m=1}^{M} z_{1,m}(s).$$

From this, it follows that

$$\sum_{m=1}^{M} G_{1,m}^{\boldsymbol{\alpha},\beta}(x) < \frac{1}{\beta},$$

and $G^{\boldsymbol{\alpha},\beta}(x) \in D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta}$. ∎

It is apparent that the flow (3.1) has a family of attracting fixed points given by $x_{0,m} = x_{1,m}$ and, moreover, that $x_{0,m} + x_{1,m}$ is a conserved quantity under this flow. Therefore, if we assume that $x_{0,m}(t) + x_{1,m}(t) = \alpha_m$ for some t, then this is true for all t. Under this restriction, there is a unique attracting fixed point $\xi_{\mathsf{FP}}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}$ given by

$$\left(\xi_{\mathsf{FP}}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}\right)_{0,m} = \left(\xi_{\mathsf{FP}}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}\right)_{1,m} = \frac{\alpha_m}{2}.$$

Lemma 3.7. *If $\beta > 2$, then $\xi_{\mathsf{FP}}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta} \in D_G^{\boldsymbol{\alpha},\beta}$, and every initial condition gives rise to a solution with infinitely many big bursts. Moreover, the time it takes any initial condition to enter $D_G^{\boldsymbol{\alpha},\beta}$ is uniformly bounded above.*

Proof. Notice that

$$\sum_{m=1}^{M} \left(\xi_{\mathsf{FP}}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta} \right)_{1,m} = \sum_{m=1}^{M} \frac{\alpha_m}{2} = \frac{1}{2}.$$

If $\beta > 2$, this is greater than β^{-1}; every initial condition will enter $D_G^{\boldsymbol{\alpha},\beta}$ under the flow. A stronger result is true: for any fixed $\beta > 2$, and any initial condition $x \in D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta}$, there is a global upper bound on the amount of time the system will flow until it hits $D_G^{\boldsymbol{\alpha},\beta}$. Let $\rho_{\mathsf{min}} = \min_{m=1}^{M} \rho_m$, and note that the initial condition $x_{0,m}(0) \leq \alpha_m$ for all m. Then $x_{0,m}(t) = \alpha_m e^{-\rho_m t}$, and we have

$$\sum_{m=1}^{M} x_{0,m}(t) \leq \sum_{m=1}^{M} \alpha_m e^{-\rho_m t} \leq \sum_{m=1}^{M} \alpha_m e^{-\rho_{\mathsf{min}} t} = e^{-\rho_{\mathsf{min}} t},$$

so that at some time less than $t = \rho_{\mathsf{min}}^{-1} \log(\beta/(\beta-1))$, we have $y_0 = 1 - \beta^{-1}$ and thus $y_1 = \beta^{-1}$. By existence-uniqueness and using the fact that different m modes are decoupled in the flow, any other initial condition must reach this threshold at least as quickly.

Since the only way for the hybrid system to have finitely many big bursts is for it to stay in the flow mode for an infinite time, we are done. ∎

3.4. Growth properties of stopped flow. The main result of this subsection is Theorem 3.11, from which we obtain an upper bound on the maximal stretching given by the stopped flow.

Definition 3.8.

$$F^{\boldsymbol{\alpha}} := \left\{ x \in D^{\boldsymbol{\alpha}} : x_{1,m} < \frac{\alpha_m}{2} \text{ for all } m \right\}.$$

Lemma 3.9. *For any $\beta > 2$, there exists $n_{\star}(\beta)$ such that for any $\boldsymbol{\rho} > 0$, and any solution of the hybrid system $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(t)$ with initial condition $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(0) \in D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta}$, we have $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(t) \in F^{\boldsymbol{\alpha}}$ for all $t > \tau_{n_{\star}(\beta)}$.*

Remark 3.10. In short, this lemma says that any initial condition will remain in $F^{\boldsymbol{\alpha}}$ after a finite number of big bursts, and this number depends only on β.

Proof. We will break this proof up into two steps: first, we will show that $F^{\boldsymbol{\alpha}}$ is absorbing; second, we will show that every initial condition will enter it after $n_{\star}(\beta)$ big bursts. Together, this will prove the lemma.

First assume that $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(t) \in F^{\boldsymbol{\alpha}}$, and let τ_n be the time of the next big burst after t. From (3.2), the $(1,m)$ coordinate cannot cross $\alpha_m/2$ under the flow, so $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(\tau_n-) \in F^{\boldsymbol{\alpha}}$. Let us denote $x = \xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(\tau_n-)$, and, recalling (3.5), we have

$$(3.6) \qquad G_{1,m}^{\boldsymbol{\alpha},\beta}(x) = e^{-\beta s_{\star}^{\beta}(x)} (\beta s_{\star}^{\beta}(x) x_{0,m} + x_{1,m}).$$

This is a linear combination of $x_{0,m} \in [\alpha_m/2, \alpha_m]$ and $x_{1,m} \in [0, \alpha_m/2]$, so we need only check the extremes. If we take $x_{0,m} = \alpha_m$ and $x_{1,m} = 0$, then we have $G_{1,m}^{\boldsymbol{\alpha},\beta}(x) = z e^{-z} \alpha_m$ for some $z > 0$, and $\sup_{z>0} z e^{-z} = 1/e$. Considering the other extreme gives $G_{1,m}^{\boldsymbol{\alpha},\beta}(x) = (z+1) e^{-z} \alpha_m/2$, and $\sup_{z>0} (z+1) e^{-z} = 1$. In either case, we have $G_{1,m}^{\boldsymbol{\alpha},\beta}(x) < \alpha_m/2$, and we see that $F^{\boldsymbol{\alpha}}$ is absorbing.

Now assume that $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(0) \notin F^{\boldsymbol{\alpha}}$. Since $\beta > 2$, it follows from Lemma 3.7 that $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(t)$ has infinitely many big bursts. Let $x = \xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(\tau_1-)$, noting by definition that $x \in \partial D_G^{\boldsymbol{\alpha},\beta}$. Using (3.6) and $x_{1,m} > \alpha_m/2$, $x_{0,m} < x_{1,m}$,

$$G_{1,m}^{\boldsymbol{\alpha},\beta}(x) < e^{-\beta s_\star^\beta(x)}(\beta s_\star^\beta(x) + 1)x_{1,m}.$$

By Lemma 3.4 and again recalling that $(z+1)e^{-z} < 1$ for all $z > 0$, this means that there is an $h(\beta) \in (0,1)$ with

$$\xi_{1,m}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(\tau_1) < h(\beta) \cdot x_{1,m}.$$

If $h(\beta)x_{1,m} < \alpha_m/2$, then we are done. If not, notice that the flow generated by $\mathcal{L}$ will make the $(1,m)$ coordinate decrease, so it is clear that if $\xi_{1,m}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(t) \notin F^{\boldsymbol{\alpha}}$ for all $t \in [0, \tau_n)$, then by induction $\xi_{1,m}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(\tau_n) < (h(\beta))^n \alpha_m$. Choose $n_\star(\beta)$ so that $(h(\beta))^{n_\star(\beta)} < 1/2$, and we have that $\xi_{1,m}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(\tau_{n_\star(\beta)}) < \alpha_m/2$ and thus $\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(\tau_{n_\star(\beta)}) \in F^{\boldsymbol{\alpha}}$. ∎

Theorem 3.11. *Choose any two initial conditions $x(0), \widetilde{x}(0) \in F^{\boldsymbol{\alpha}} \cap D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta}$, and define $\tau, \widetilde{\tau}$ as in (2.4). Then*

$$\left\| e^{\tau\mathcal{L}}x(0) - e^{\widetilde{\tau}\mathcal{L}}\widetilde{x}(0) \right\| \leq \left(1 + \frac{\sqrt{M}}{2}\right) \|x(0) - \widetilde{x}(0)\| ;$$

i.e., for any two initial conditions, the distance at the time of the first big burst has grown by no more than a factor of $1 + \sqrt{M}/2$.

Proof. Before we start, recall that the map $e^{\tau\mathcal{L}}x$ is nonlinear in x, because τ itself depends nonlinearly on x. Let $\mathbf{1}_M$ be the all-ones column vector in $\mathbb{R}^M$. Let $x(0) \in D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta}$, consider a perturbation $\boldsymbol{\epsilon} = \{\epsilon_m\}$ with $\sum_m \epsilon_m = 0$, i.e., $\boldsymbol{\epsilon} \in \mathbf{1}_M^\perp$, and define $\widetilde{x}(0)$ by

$$\widetilde{x}_{m,1}(0) = x_{m,1}(0) + \epsilon_m, \quad \widetilde{x}_{m,0}(0) = x_{m,0}(0) - \epsilon_m.$$

Define $\tau, \widetilde{\tau}$ as the burst times associated with these initial conditions as in (2.4), and by definition, we have

$$\sum_{m=1}^M x(\tau-)_{1,m} = \sum_{m=1}^M \widetilde{x}(\widetilde{\tau}-)_{1,m} = \frac{1}{\beta}.$$

Writing $\widetilde{\tau} = \tau + \delta$ and using (3.2), we have

$$\sum_{m=1}^M \left(\frac{\alpha_m}{2} - \left(\frac{\alpha_m}{2} - x_{1,m}(0)\right)e^{-2\rho_m\tau}\right) = \sum_{m=1}^M \left(\frac{\alpha_m}{2} - \left(\frac{\alpha_m}{2} - \widetilde{x}_{1,m}(0)\right)e^{-2\rho_m(\tau+\delta)}\right).$$

Since $\widetilde{x} - x = O(\epsilon)$ and $e^{-2\rho_m\delta} = (1 + O(\delta))$, we can see from this expression that the leading order terms in both ϵ and δ are of the same order. Thus, Taylor expanding to first order in ϵ and δ and canceling gives a solution for δ:

$$(3.7) \qquad \delta = -\frac{\sum_\ell \epsilon_\ell e^{-2\rho_\ell\tau}}{2\sum_\ell \rho_\ell \left(\frac{\alpha_\ell}{2} - x_{1,\ell}(0)\right)e^{-2\rho_\ell\tau}} + O(\epsilon^2).$$

We then have

$$\widetilde{x}_{1,m}(\tau + \delta) - x_{1,m}(\tau) = \epsilon_m e^{-2\rho_m\tau} - 2\rho_m \left(\frac{\alpha_m}{2} - x_{1,m}(0)\right)\delta e^{-2\rho_m\tau}$$

$$= \epsilon_m e^{-2\rho_m\tau} - c_m \sum_\ell \epsilon_\ell e^{-2\rho_\ell\tau} + O(\epsilon^2),$$

where

$$(3.8) \qquad c_m = \frac{\rho_m \left(\frac{\alpha_m}{2} - x_{1,m}(0)\right) e^{-2\rho_m\tau}}{\sum_\ell \rho_\ell \left(\frac{\alpha_\ell}{2} - x_{1,\ell}(0)\right) e^{-2\rho_\ell\tau}}.$$

Since $x(0) \in F^{\boldsymbol{\alpha}}$, $c_m > 0$. It is then clear from the definition that $c_m < 1$. Writing this in matrix form in terms of ϵ gives

$$(3.9) \qquad \begin{pmatrix} \widetilde{x}_{1,m}(\tau + \delta) - x_{1,m}(\tau) \\ \widetilde{x}_{2,m}(\tau + \delta) - x_{2,m}(\tau) \\ \vdots \\ \widetilde{x}_{2,M}(\tau + \delta) - x_{2,M}(\tau) \end{pmatrix} = \mathbf{M}_M \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{pmatrix} + O(\epsilon^2),$$

where the matrix $\mathbf{M}_M$ is defined as

$$(3.10) \qquad \mathbf{M}_M = \begin{pmatrix} e^{-2\rho_1\tau} - c_1 e^{-2\rho_1\tau} & -c_1 e^{-2\rho_2\tau} & \cdots & -c_1 e^{-2\rho_M\tau} \\ -c_2 e^{-2\rho_1\tau} & e^{-2\rho_2\tau} - c_2 e^{-2\rho_2\tau} & \cdots & -c_2 e^{-2\rho_M\tau} \\ \vdots & \vdots & \ddots & \vdots \\ -c_M e^{-2\rho_1\tau} & -c_N e^{-2\rho_2\tau} & \cdots & e^{-2\rho_M\tau} - c_N e^{-2\rho_M\tau} \end{pmatrix},$$

or, more compactly,

$$(\mathbf{M}_M)_{ij} = -c_i e^{-2\rho_j\tau} + \delta_{ij} e^{-2\rho_i\tau}.$$

Thus, the map $e^{\tau\mathcal{L}}x$ has Jacobian $\mathbf{M}_M$. Since $\mathbf{M}_M$ has zero column sums, it is apparent that $\mathbf{1}^{\mathsf{T}}\mathbf{M}_M = \mathbf{0}$ and thus $0 \in \mathrm{Spec}(\mathbf{M}_M)$. Since all of the nondiagonal entries of $\mathbf{M}_M$ are bounded above by one, the standard Gershgorin estimate implies that all of the eigenvalues of $\sqrt{\mathbf{M}_M^{\mathsf{T}}\mathbf{M}_M}$ lie in a disk of radius $O(M)$ around the origin, but this is not good enough to establish our result.

We can work out a more delicate bound: by the definition of $D^{\boldsymbol{\alpha}}$, we need only consider zero sum perturbations, and so in fact we are concerned with $\mathbf{M}_M$ restricted to $\mathbf{1}_M^\perp$. From this and the fundamental theorem of calculus, it follows that

$$\left\| e^{\tau\mathcal{L}}x(0) - e^{\widetilde{\tau}\mathcal{L}}\widetilde{x}(0) \right\| \le \left\| \mathbf{M}_M|_{\mathbf{1}_m^\perp} \right\|_2 \left\| x(0) - \widetilde{x}(0) \right\|,$$

where $\|\cdot\|_2$ is the spectral norm of a matrix (see Definition 3.12 below). Using the bound in Lemma 3.13 proves the theorem. ∎

Definition 3.12. *We define the* spectral norm *of a square matrix A by*

$$\|A\|_2 = \sup_{x \ne \mathbf{0}} \frac{\|Ax\|_2}{\|x\|_2},$$

where $\|\cdot\|_2$ is the Euclidean (L^2) norm of a vector.

The spectral norm of a matrix is equal to its largest singular value, and if the matrix is symmetric, this is the same as the largest eigenvalue. In particular, it follows from the definition that

$$\|Ax\|_2 \le \|A\|_2 \|x\|_2.$$

Theorem 3.13. *Let $\mathbf{1}_M^\perp \subseteq \mathbb{R}^M$ denote the subspace of zero-sum vectors. Then $\mathbf{M}_M \colon \mathbf{1}_M^\perp \to \mathbf{1}_M^\perp$ since it is a zero column sum matrix, and thus the restriction is well defined. Then*

$$(3.11) \qquad \left\| \mathbf{M}_M \big|_{\mathbf{1}_M^\perp} \right\|_2 < 1 + \frac{\sqrt{M}}{2}.$$

Proof. Let us denote $\mathbf{I}_M$ to be the M-by-M identity matrix and $\mathbf{1}_M$ the all-ones column vector in $\mathbb{R}^M$. We will also define the matrix $\mathbf{D}_M$ and vector $\mathbf{d}_M$ by

$$\mathbf{d}_M = [e^{-2\rho_1 s}, e^{-2\rho_2 s}, \ldots, e^{-2\rho_N s}]^\intercal,$$

and $\mathbf{D}_M$ is the matrix with $\mathbf{d}_M$ on the diagonal, i.e., $(\mathbf{D}_M)_{ij} = \delta_{ij} e^{-2\rho_i \tau}$.

Any vector $\mathbf{v} \in \mathbf{1}_M^\perp$ is in the null space of the matrix $\mathbf{1}\mathbf{1}^\intercal$, and thus $(\mathbf{I}_M - M^{-1}\mathbf{1}\mathbf{1}^\intercal)\mathbf{v} = \mathbf{v}$, and $\mathbf{M}_M = \mathbf{M}_M(\mathbf{I}_M - M^{-1}\mathbf{1}\mathbf{1}^\intercal)$ on $\mathbf{1}^\perp$, so it suffices for our result to bound the norm of $\mathbf{M}_M(\mathbf{I}_M - M^{-1}\mathbf{1}\mathbf{1}^\intercal)$.

We can factorize

$$(3.12) \qquad \mathbf{M}_M = (\mathbf{I} - \mathbf{c}\mathbf{1}^\intercal)\mathbf{D}_M,$$

where the components of $\mathbf{c}$ are given in (3.8). To see this, we compute

$$((\mathbf{I} - \mathbf{c}\mathbf{1}^\intercal)\mathbf{D}_M)_{ij} = (\mathbf{D}_M)_{ij} - (\mathbf{c}\mathbf{1}^\intercal \mathbf{D}_M)_{ij} = (\mathbf{D}_M)_{ij} - \sum_k c_i \cdot 1 \cdot \delta_{k,j} e^{-2\rho_j \tau}$$

$$= \delta_{ij} e^{-2\rho_i \tau} - c_i e^{-2\rho_j \tau}.$$

Let us first write

$$\mathbf{M}_M = (\mathbf{I} - \mathbf{c}\mathbf{1}^\intercal)\mathbf{D}_M = (\mathbf{D}_M - \mathbf{D}_M \mathbf{c}\mathbf{1}^\intercal + \mathbf{D}_M \mathbf{c}\mathbf{1}^\intercal - \mathbf{c}\mathbf{1}^\intercal \mathbf{D}_M)$$
$$= \mathbf{D}_M(\mathbf{I} - \mathbf{c}\mathbf{1}^\intercal) + (\mathbf{D}_M \mathbf{c}\mathbf{1}^\intercal - \mathbf{c}\mathbf{d}_M^\intercal),$$

where we use the relation $\mathbf{1}^\intercal \mathbf{D}_M = \mathbf{d}_M^\intercal$, and then

$$(3.13) \quad \mathbf{M}_M(\mathbf{I} - M^{-1}\mathbf{1}\mathbf{1}^\intercal) = \mathbf{D}_M(\mathbf{I} - \mathbf{c}\mathbf{1}^\intercal)(\mathbf{I} - M^{-1}\mathbf{1}\mathbf{1}^\intercal) + (\mathbf{D}_M \mathbf{c}\mathbf{1}^\intercal - \mathbf{c}\mathbf{d}_M^\intercal)(\mathbf{I} - M^{-1}\mathbf{1}\mathbf{1}^\intercal).$$

We break this into two parts. Using the fact that $\mathbf{1}^\intercal \mathbf{1} = M$, we have

$$\mathbf{c}\mathbf{1}^\intercal(\mathbf{I}_M - M^{-1}\mathbf{1}\mathbf{1}^\intercal) = \mathbf{I}_M \mathbf{c}\mathbf{1}^\intercal - M^{-1}\mathbf{c}\mathbf{1}^\intercal \mathbf{1}\mathbf{1}^\intercal = \mathbf{c}\mathbf{1}^\intercal - \mathbf{c}\mathbf{1}^\intercal = 0,$$

and thus the first term can be simplified to

$$(3.14) \qquad \begin{aligned} &\mathbf{D}_M(\mathbf{I} - \mathbf{c}\mathbf{1}^\intercal)(\mathbf{I} - M^{-1}\mathbf{1}\mathbf{1}^\intercal) \\ &= \mathbf{D}_M(\mathbf{I} - M^{-1}\mathbf{1}\mathbf{1}^\intercal) - \mathbf{D}_M(\mathbf{c}\mathbf{1}^\intercal)(\mathbf{I}_M - M^{-1}\mathbf{1}\mathbf{1}^\intercal) = \mathbf{D}_M(\mathbf{I} - M^{-1}\mathbf{1}\mathbf{1}^\intercal). \end{aligned}$$

Since the matrix $M^{-1}\mathbf{1}\mathbf{1}^\intercal$ is an orthogonal projection matrix with norm 1 and rank 1, it follows that $\mathbf{I}_M - M^{-1}\mathbf{1}\mathbf{1}^\intercal$ is also a projection matrix with norm 1 and rank $M - 1$. By Cauchy–Schwarz, the norm can be bounded by

$$(3.15) \qquad \left\|\mathbf{D}_M(\mathbf{I} - M^{-1}\mathbf{1}\mathbf{1}^\intercal)\right\|_2 \leq \|\mathbf{D}_M\|_2 \left\|\mathbf{I}_M - M^{-1}\mathbf{1}\mathbf{1}^\intercal\right\|_2 = \|\mathbf{D}_M\|_2 < 1.$$

(The last inequality follows from the fact that $\mathbf{D}_M$ is diagonal and all entries are less than one in magnitude.)

For the second term in (3.13) and noting that $\mathbf{d}^\intercal\mathbf{1} = \sum_m d_m$, we obtain

$$(\mathbf{D}_M\mathbf{c}\mathbf{1}^\intercal - \mathbf{c}\mathbf{d}^\intercal)(\mathbf{I} - M^{-1}\mathbf{1}\mathbf{1}^\intercal)$$

$$= \mathbf{D}_M\mathbf{c}\mathbf{1}^\intercal - \mathbf{c}\mathbf{d}^\intercal - \mathbf{D}_M\mathbf{c}\mathbf{1}^\intercal + \frac{\mathbf{d}^\intercal\mathbf{1}}{M}\mathbf{c}\mathbf{1}^\intercal = \mathbf{c}\left(\frac{\sum_m d_m}{M}\mathbf{1} - \mathbf{d}\right)^\intercal.$$

This outer product is of rank 1, and thus it has exactly one nonzero singular value; this singular value is the product of the L^2 norms of the two vectors, and therefore

$$\left\|(\mathbf{D}_M\mathbf{c}\mathbf{1}^\intercal - \mathbf{c}\mathbf{d}^\intercal)(\mathbf{I}_M - M^{-1}\mathbf{1}\mathbf{1}^\intercal)\right\|_2 = \|\mathbf{c}\|_2\left\|\mathbf{d} - \frac{\sum_m d_m}{M}\mathbf{1}\right\|_2 < 1 \cdot \frac{\sqrt{M}}{2}.$$

Using (3.13) and the triangle inequality gives the result. ∎

3.5. Contraction of the big burst map. In this section, we demonstrate that $G^{\boldsymbol{\alpha},\beta}$ is a contraction for β large enough and, moreover, that one can make the contraction modulus as small as desired by choosing β sufficiently large.

Theorem 3.14. For any $M \geq 1$ and $\delta > 0$, there is a $\beta_1(M,\delta)$ such that for all $\beta > \beta_1(M,\delta)$ and $x, \widetilde{x} \in \partial D_G^{\boldsymbol{\alpha},\beta}$,

$$\left\|G^{\boldsymbol{\alpha},\beta}(x) - G^{\boldsymbol{\alpha},\beta}(\widetilde{x})\right\| \leq \delta\|x - \widetilde{x}\|.$$

In particular, by choosing β sufficiently large, we can make this map have as small a modulus of contraction as required.

Proof. Let us define the vector $\boldsymbol{\epsilon}$ by

$$\epsilon_m = \widetilde{x}_m - x_m.$$

Since $x, \widetilde{x}$ are both in $\partial D_G^{\boldsymbol{\alpha},\beta}$, $\boldsymbol{\epsilon} \perp \mathbf{1}$. It follows from (3.3) that $\nabla_{\boldsymbol{\epsilon}}\psi^\beta(s,x) = 0$. Recall from (3.5) that

$$G_{1,m}^{\boldsymbol{\alpha},\beta}(x) = e^{-\beta s_\star^\beta(x)}(\beta s_\star^\beta(x)x_{0,m} - x_{1,m}),$$

and thus

$$\nabla_{\boldsymbol{\epsilon}}G_{1,m}^{\boldsymbol{\alpha},\beta}(x) = e^{-\beta s^*(x)}\left(-\beta\nabla_{\boldsymbol{\epsilon}}s_\star^\beta(x)\right)(\beta s_\star^\beta(x)x_{0,m} - x_{1,m})$$

$$+ e^{-\beta s_\star^\beta(x)}\left(\beta\nabla_{\boldsymbol{\epsilon}}x_{0,m} - \nabla_{\boldsymbol{\epsilon}}x_{1,m}\right)$$

$$= e^{-\beta s_\star^\beta(x)}(\beta s_\star^\beta(x)(-1) - 1),$$

so

$$\nabla_{\boldsymbol{\epsilon}}G^{\boldsymbol{\alpha},\beta}(x) = -(e^{-\beta s_\star^\beta(x)}(\beta s_\star^\beta(x) + 1))\mathbf{1}.$$

Note that Lemma 3.4 implies that $\beta s_\star^\beta(x) \to \infty$ as $\beta \to \infty$ for any x. If we define the function $g(z) = e^{-z}(1+z)$, then it is easy to see that

$$0 < g(z) < 1 \text{ for } z \in (0, \infty), \quad \lim_{z \to \infty} g(z).$$

From this and the fundamental theorem of calculus, the result follows. ∎

3.6. Proof of main theorem. Finally, to prove the theorem, we will show that, under sufficient conditions on β, the composition of the map and the flow is eventually a strict contraction for any initial condition.

Definition 3.15. *We define*

$$\mathcal{H}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta} \colon D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta} \to D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta},$$
$$x \mapsto G^{\boldsymbol{\alpha},\beta}(e^{\tau\mathcal{L}}x),$$

where τ is the first hitting time defined in (2.4).

Proof of Theorem 3.1. If we consider any solution of the hybrid system $\xi^{2,\boldsymbol{\alpha},\boldsymbol{\rho}}\beta(t)$ that has infinitely many big bursts, then it is clear from chasing definitions that

$$\xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(\tau_n) = \left(\mathcal{H}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}\right)^n \xi^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}(0).$$

$\mathcal{H}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}$ is the composition of two maps—one coming from a stopped flow and the other coming from the map G. It follows from Theorem 3.11 that the modulus of contraction of the stopped flow is no more than $1 + \sqrt{M}/2$ on the set $F^{\boldsymbol{\alpha}}$ whenever $\beta > 2$. It follows from Theorem 3.14 that we can make the modulus of the second flow less than δ by choosing $\beta > \beta_1(M, \delta)$. Let us define

$$\beta_M := \beta_1\left(M, \frac{1}{1 + \sqrt{M}/2}\right),$$

and then by composition it follows that $\mathcal{H}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}$ is a strict contraction on $F^{\boldsymbol{\alpha}}$. From Lemma 3.9, it follows that $D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta}$ is mapped into $F^{\boldsymbol{\alpha}}$ in a finite number of iterations, so that $\mathcal{H}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}$ is eventually strictly contracting on $D_{\mathcal{L}}^{\boldsymbol{\alpha},\beta}$, and therefore $\mathcal{H}^{\boldsymbol{\alpha},\boldsymbol{\rho},\beta}$ has a globally attracting fixed point, which means that the hybrid system has a globally attracting limit cycle.

Finally, we want to understand the asymptotics as $M \to \infty$. Choose any $0 < \gamma_1, \gamma_2 < 1$. By Lemma 3.4, $\beta s_\star(\beta) > \gamma_1\beta$ for β sufficiently large, and it is clear that $e^{-z}(z+1) < e^{-\gamma_2 z}$ for z sufficiently large. From these it follows that for β sufficiently large,

$$e^{-\beta s_\star(\beta)}(\beta s_\star(\beta) + 1) < e^{-\gamma_1\gamma_2\beta}.$$

From this we have that $\beta_M < \ln(1 + \sqrt{M}/2)/\gamma_1\gamma_2$, and the result follows. ∎

We have shown that β_M is finite and have determined its asymptotic scaling as $M \to \infty$. It was shown in [10] that $\beta_1 = 2$, and we can now show, as follows, that this is the case as well for $M = 2$.

Proposition 3.16. *The computations in the proof of Theorem* 3.1 *imply that β_2 is, at most, the largest solution of the equation $e^{-\beta s_\star(\beta)}(\beta s_\star(\beta) + 1) < 2/3$. Numerical approximation of this root gives $\beta \approx 2.48$. However, in fact, $\beta_2 = 2$.*

Proof. In $\mathbb{R}^2$, $\mathbf{1}^\perp$ is a one-dimensional space spanned by $(1, -1)^\mathsf{T}$, and thus we need only compute the eigenvalue associated with this vector. If we define $\mathbf{v} = \mathbf{M}_2 \cdot (1, -1)^\mathsf{T}$ and show $|v_1 - v_2| < 2$, then we have established the result. When $M = 2$, we can write (3.10) as

$$(3.16) \qquad \mathbf{M}_2 = \begin{pmatrix} e^{-2\rho_1\tau} - c_1 e^{-2\rho_1\tau} & -c_1 e^{-2\rho_2\tau} \\ -c_2 e^{-2\rho_1\tau} & e^{-2\rho_2\tau} - c_2 e^{-2\rho_2\tau} \end{pmatrix},$$

and thus

$$\mathbf{v} = \begin{pmatrix} e^{-2\rho_1\tau} - c_1 e^{-2\rho_1\tau} + c_1 e^{-2\rho_2\tau} \\ -c_2 e^{-2\rho_1\tau} - e^{-2\rho_2\tau} + c_2 e^{-2\rho_2\tau} \end{pmatrix}.$$

Thus

$$v_1 - v_2 = e^{-2\rho_1\tau}(1 - c_1 + c_2) + e^{-2\rho_2\tau}(1 + c_1 - c_2).$$

Using $c_1 + c_2 = 1$, this simplifies to

$$v_1 - v_2 = 2c_2 e^{-2\rho_1\tau} + 2c_1 e^{-2\rho_2\tau}.$$

Since it is clear that $v_1 - v_2 > 0$, we need to show that $v_1 - v_2 < 2$, or

$$2c_1 e^{2\rho_1\tau} + 2c_2 e^{2\rho_2\tau} < 2e^{(\rho_1 + \rho_2)\tau}.$$

Writing $A = \rho_1(\alpha_1/2 - x_{1,1}(0))$, $B = \rho_2(\alpha_2/2 - x_{1,2}(0))$, this becomes

$$(3.17) \qquad \frac{A + B}{Ae^{2\rho_2\tau} + Be^{2\rho_1\tau}} < 1,$$

but this must be satisfied, since $e^{2\rho_1\tau}, e^{2\rho_2\tau} > 1$. ∎

Remark 3.17. We conjecture from numerical evidence (cf. Figure 5) that, in fact, $\beta_M = 2$ for all M.

4. Numerical simulations. In this section, we present some numerical simulations; we verify the existence of the unique attractor whose existence is proven above and give evidence for the conjecture that $\beta_M = 2$ for all M.

We first numerically solve the hybrid ODE-mapping system, with $M = 3$ and random α_i, ρ_i. The ODE portion of the hybrid system can be solved explicitly, and we use MATLAB's `fsolve` to determine the hitting times τ_i. We plot the trace of the system for $\beta = 2.1$, $\beta = 2.5$ for a single initial condition in Figure 3. We observe that each neuron population is attracted to a periodic orbit after several bursts.

To further demonstrate convergence, we also plot trajectories for the same parameters for various initial conditions in Figure 4. We see that after three to four bursts, the trajectories converge to the same periodic orbit.

We also present some numerics verifying the conjecture in Remark 3.17; i.e., the numerical evidence in Figure 5 suggests that $\beta_M = 2$ in general, or, at least, it is much less than the upper bound given in the main theorem. To check this, we choose $10,000$ initial conditions uniformly random in the simplex and verify that all initial conditions converge to the attracting limit cycle for all $\beta > 2$. We also see that the eigenvalues of the map $\mathcal{H}^{\alpha,\rho,\beta}$ have a complicated dependence on β: there seem to be regions where this map has negative eigenvalues and some where it does not, which can be detected by whether we converge monotonically to the limit cycle or not. But it seems to always converge for any $\beta > 2$.
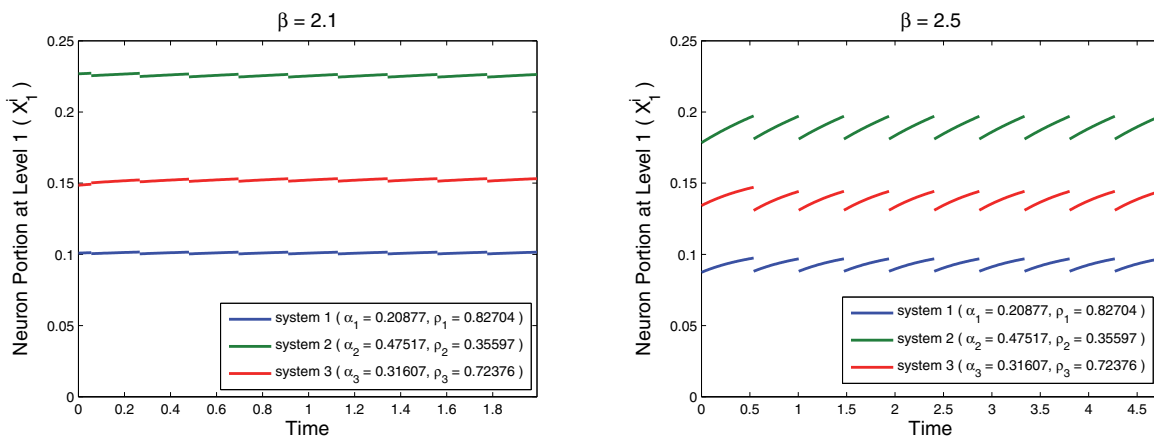
Figure 3. *Plots of the hybrid ODE-mapping system numerical simulation results with $\beta = 2.1$ (left) and $\beta = 2.5$ (right). Both of them are with three neuron populations. The neuron portions at energy level 1 over simulation time are shown in the plots.*
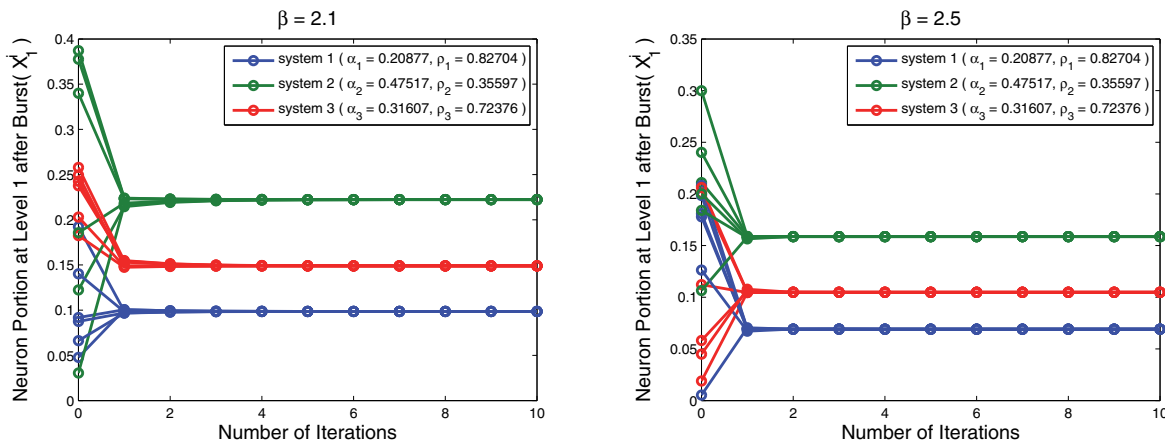

Figure 4. *Plots of neuron proportions after each burst iteration with $\beta = 2.1$ (left) and $\beta = 2.5$ (right). Both subfigures are for $M = 3$. For all initial conditions, the population seems to converge after about four bursts.*

5. Conclusion. We generalized the mean-field model derived in of [9, 10] to the case of multiple subpopulations with different intrinsic firing rates. We analyzed the limiting mean field in the case where each neuron has at most two inactive states and proved that for sufficiently large coupling parameters, the mean-field limit has a globally attracting limit cycle.

We point out a few similar results in the literature. A similar three-state model was considered in [25, 29, 30] where the number of neurons in the analogous firing state affected the firing rates of all of the neurons in the system. The mean-field model derived there was a delay system instead of a hybrid system, but that model also exhibited the coexistence of an attracting periodic orbit and an attracting limit cycle, similar to the case considered here. In
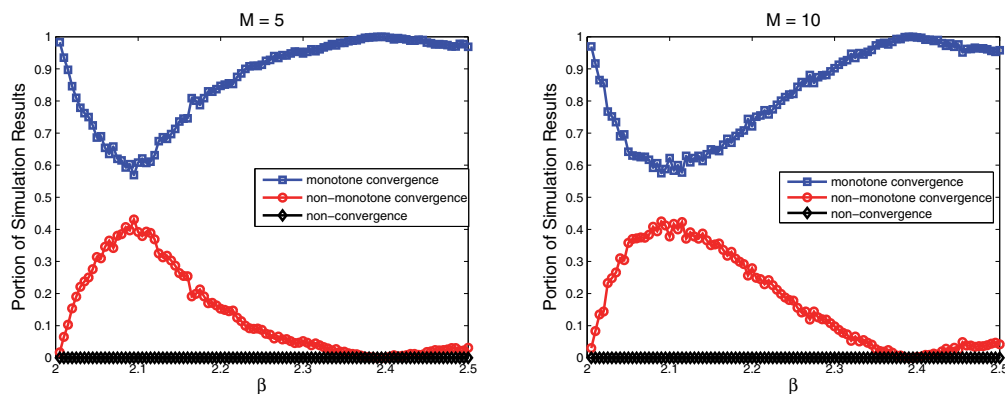
Figure 5. *Proportions of initial conditions that converge monotonically, converge nonmonotonically, or do not converge for $M = 5$ and $M = 10$ subpopulations. The parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\rho}$ are chosen at random and fixed. For each β, we choose $10,000$ initial conditions uniformly in the simplex and determine which proportion falls into each of three categories: monotone convergent, nonmonotone convergent, and nonconvergent. We vary β from 2.005 to 2.5. We see that all initial conditions converge, but the monotonicity of the convergence depends on β.*

a different direction, the complete characterization of the dynamics of a network of interacting theta neurons was studied in [23]. Again, this model exhibits the coexistence of macroscopic limit cycles and fixed points, but the bifurcation structure described in [23] has more distinct features than the one observed in the present paper (cf. Figure (8b) of [23] and Figure 2 of the current paper). The model considered here is at first glance quite different from these other two models in that it explicitly incorporates cascades directly into the dynamics; interestingly, it shows many of the same macroscopic phenomena.

REFERENCES

[1] J. M. BEGGS AND D. PLENZ, *Neuronal avalanches in neocortical circuits*, J. Neurosci., 23 (2003), pp. 11167–11177.

[2] J. M. BEGGS AND D. PLENZ, *Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures*, J. Neurosci., 24 (2004), pp. 5216–5229.

[3] M. S. BRANICKY, *Stability of hybrid systems: State of the art*, in Proceedings of the 36th IEEE Conference on Decision and Control, Vol. 1, IEEE, Washington, DC, 1997, pp. 120–125.

[4] P. C. BRESSLOFF AND S. COOMBES, *Desynchronization, mode locking, and bursting in strongly coupled integrate-and-fire oscillators*, Phys. Rev. Lett., 81 (1998), pp. 2168–2171.

[5] J. BUCK AND E. BUCK, *Mechanism of rhythmic synchronous flashing of fireflies*, Science, 159 (1968), pp. 1319–1327.

[6] S. R. CAMPBELL, D. L. L. WANG, AND C. JAYAPRAKASH, *Synchrony and desynchrony in integrate-and-fire oscillators*, Neural Comput., 11 (1999), pp. 1595–1619.

[7] J. H. E. CARTWRIGHT, V. M. EGULUZ, E. HERNNDEZ-GARCA, AND O. PIRO, *Dynamics of elastic excitable media*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 9 (1999), pp. 2197–2202.

[8] C. A. CZEISLER, E. WEITZMAN, M. C. MOORE-EDE, J. C. ZIMMERMAN, AND R. S. KNAUER, *Human sleep: Its duration and organization depend on its circadian phase*, Science, 210 (1980), pp. 1264–1267.

[9] R. E. L. DEVILLE AND C. S. PESKIN, *Synchrony and asynchrony in a fully stochastic neural network*, Bull. Math. Biol., 70 (2008), pp. 1608–1633.

[10] R. E. L. DeVille, C. S. Peskin, and J. H. Spencer, *Dynamics of stochastic neuronal networks and the connections to random graph theory*, Math. Model. Natural Phenomena, 5 (2010), pp. 26–66.

[11] G. B. Ermentrout and J. Rinzel, *Reflected waves in an inhomogeneous excitable medium*, SIAM J. Appl. Math., 56 (1996), pp. 1107–1128.

[12] N. Friedman, S. Ito, B. A. W. Brinkman, M. Shimono, R. E. L. DeVille, K. A. Dahmen, J. M. Beggs, and T. C. Butler, *Universal critical dynamics in high resolution neuronal avalanche data*, Phys. Rev. Lett., 108 (2012), 208102.

[13] W. Gerstner and J. L. van Hemmen, *Coherence and incoherence in a globally-coupled ensemble of pulse-emitting units*, Phys. Rev. Lett., 71 (1993), pp. 312–315.

[14] L. Glass, A. L. Goldberger, M. Courtemanche, and A. Shrier, *Nonlinear dynamics, chaos and complex cardiac arrhythmias*, Proc. Roy. Soc. London Ser. A, 413 (1987), pp. 9–26.

[15] M. R. Guevara and L. Glass, *Phase locking, period doubling bifurcations and chaos in a mathematical model of a periodically driven oscillator: A theory for the entrainment of biological oscillators and the generation of cardiac dysrhythmias*, J. Math. Biol., 14 (1982), pp. 1–23.

[16] D. Hansel and H. Sompolinsky, *Synchronization and computation in a chaotic neural network*, Phys. Rev. Lett., 68 (1992), pp. 718–721.

[17] R. Kapral and K. Showalter, eds., *Chemical Waves and Patterns*, Springer, New York, 1994.

[18] B. W. Knight, *Dynamics of encoding in a population of neurons*, J. General Physiology, 59 (1972), pp. 734–766.

[19] Y. Kuramoto, *Collective synchronization of pulse-coupled oscillators and excitable units*, Phys. D, 50 (1991), pp. 15–30.

[20] J. C. Lagarias and Y. Wang, *The finiteness conjecture for the generalized spectral radius of a set of matrices*, Linear Algebra Appl., 214 (1995), pp. 17–42.

[21] D. Liberzon and A. S. Morse, *Basic problems in stability and design of switched systems*, IEEE Control Systems, 19 (1999), pp. 59–70.

[22] Z.-H. Liu and P. M. Hui, *Collective signaling behavior in a networked-oscillator model*, Phys. A, 383 (2007), pp. 714–724.

[23] T. B. Luke, E. Barreto, and P. So, *Complete classification of the macroscopic behavior of a heterogeneous network of theta neurons*, Neural Comput., 25 (2013), pp. 3207–3234.

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

[25] B. Naundorf, T. Prager, and L. Schimansky-Geier, *A Novel Approach to Synchronization in Coupled Excitable Systems*, preprint, http://arxiv.org/abs/cond-mat/0211011, 2002.

[26] K. Pakdaman and D. Mestivier, *Noise induced synchronization in a neuronal oscillator*, Phys. D, 192 (2004), pp. 123–137.

[27] C. S. Peskin, *Mathematical Aspects of Heart Physiology*, Notes based on a course given at New York University during the year 1973/74, Courant Institute of Mathematical Sciences, New York University, New York, NY, 1975.

[28] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences*, Cambridge University Press, Cambridge, UK, 2003.

[29] T. Prager, B. Naundorf, and L. Schimansky-Geier, *Coupled three-state oscillators*, Phys. A, 325 (2003), pp. 176–185.

[30] T. Prager and L. Schimansky-Geier, *Stochastic Resonance in a Non Markovian Discrete State Model for Excitable Systems*, preprint, http://arxiv.org/abs/cond-mat/0310624, 2003.

[31] W. Senn and R. Urbanczik, *Similar nonleaky integrate-and-fire neurons with instantaneous couplings always synchronize*, SIAM J. Appl. Math., 61 (2000), pp. 1143–1155.

[32] W. L. Shew, H. Yang, T. Petermann, R. Roy, and D. Plenz, *Neuronal avalanches imply maximum dynamic range in cortical networks at criticality*, J. Neurosci., 29 (2009), pp. 15595–15600.

[33] S. Strogatz, *Sync: The Emerging Science of Spontaneous Order*, Hyperion, New York, 2003.

[34] D. Terman, N. Kopell, and A. Bose, *Dynamics of two mutually coupled slow inhibitory neurons*, Phys. D, 117 (1998), pp. 241–275.

[35] M. Tsodyks, I. Mitkov, and H. Sompolinsky, *Pattern of synchrony in inhomogeneous networks of oscillators with pulse interactions*, Phys. Rev. Lett., 71 (1993), pp. 1280–1283.

[36] J. J. Tyson, C. I. Hong, C. D. Thron, and B. Novak, *A simple model of circadian rhythms based on dimerization and proteolysis of PER and TIM*, Biophys. J., 77 (1999), pp. 2411–2417.

[37] J. J. Tyson and J. P. Keener, *Singular perturbation theory of traveling waves in excitable media (a review)*, Phys. D, 32 (1988), pp. 327–361.

[38] C. van Vreeswijk, L. Abbott, and G. Ermentrout, *When inhibition not excitation synchronizes neural firing*, J. Comp. Neurosci., 1 (1994), pp. 313–321.

[39] C. van Vreeswijk and H. Sompolinsky, *Chaotic balance state in a model of cortical circuits*, Neural Comput., 10 (1998), pp. 1321–1372.

[40] A. T. Winfree, *The Geometry of Biological Time*, 2nd ed., Interdiscip. Appl. Math. 12, Springer-Verlag, New York, 2001.