

STOCHASTIC NEURAL FIELD THEORY AND THE SYSTEM-SIZE EXPANSION*

PAUL C. BRESSLOFF†

Abstract. We analyze a master equation formulation of stochastic neurodynamics for a network of synaptically coupled homogeneous neuronal populations each consisting of N identical neurons. The state of the network is specified by the fraction of active or spiking neurons in each population, and transition rates are chosen so that in the thermodynamic or deterministic limit ($N \rightarrow \infty$) we recover standard activity-based or voltage-based rate models. We derive the lowest order corrections to these rate equations for large but finite N using two different approximation schemes, one based on the Van Kampen system-size expansion and the other based on path integral methods. Both methods yield the same series expansion of the moment equations, which at $\mathcal{O}(1/N)$ can be truncated to form a closed system of equations for the first- and second-order moments. Taking a continuum limit of the moment equations while keeping the system size N fixed generates a system of integrodifferential equations for the mean and covariance of the corresponding stochastic neural field model. We also show how the path integral approach can be used to study large deviation or rare event statistics underlying escape from the basin of attraction of a stable fixed point of the mean-field dynamics; such an analysis is not possible using the system-size expansion since the latter cannot accurately determine exponentially small transitions.

Key words. neural field theory, master equations, stochastic processes, system-size expansion, path integrals

AMS subject classifications. Primary, 92; Secondary, 60

DOI. 10.1137/090756971

1. Introduction. Continuum models of neural tissue have been very popular in analyzing the spatiotemporal dynamics of large-scale cortical activity (reviewed in [24, 35, 7, 16]). Such models take the form of integrodifferential equations in which the integration kernel represents the spatial distribution of synaptic connections between different neural populations, and the macroscopic state variables represent population-averaged firing rates. Following seminal work in the 1970s by Wilson, Cowan, and Amari [72, 73, 2], many analytical and numerical results have been obtained regarding the existence and stability of spatially structured solutions to continuum neural field equations. These include electroencephalogram (EEG) rhythms [55, 46, 67], geometric visual hallucinations [25, 26, 70, 8], stationary pulses or bumps [2, 62, 50, 51, 30], traveling fronts and pulses [27, 61, 6, 74, 17, 32], spatially localized oscillations or breathers [31, 33, 59], and spiral waves [44, 49, 71, 48].

One potential limitation of neural field models is that they only take into account mean firing rates and thus do not include any information about higher-order statistics such as correlations between firing activity at different cortical locations. It is well known that the spike trains of individual cortical neurons in vivo tend to be very noisy, having interspike interval distributions that are close to Poisson [68]. However, it is far less clear how noise at the single cell level translates into noise at the population level. A number of studies of fully or sparsely connected

*Received by the editors April 24, 2009; accepted for publication (in revised form) September 11, 2009; published electronically December 11, 2009. The work of the author was partially supported by award KUK-C1-013-4 made by King Abdullah University of Science and Technology (KAUST).

<http://www.siam.org/journals/siap/70-5/75697.html>

†Department of Mathematics, University of Utah, Salt Lake City, UT 84112. Current address: Mathematical Institute, University of Oxford, 24–29 St. Giles’, Oxford OX1 3LB, UK (bressloff@maths.ox.ac.uk).

integrate-and-fire networks have shown that under certain conditions, even though individual neurons exhibit Poisson-like statistics, the neurons fire asynchronously so that the total population activity evolves according to a mean-field rate equation with a characteristic activation or gain function [1, 36, 3, 10, 9, 66]. Formally speaking, the asynchronous state only exists in the thermodynamic limit $N \rightarrow \infty$ where N is the number of neurons in the population. Moreover, even if the asynchronous state exists, it might not be stable. This has motivated a number of authors to analyze statistical correlations and finite-size effects in spiking networks [37, 54, 53, 69, 5]. A complementary approach has been developed by Cai and collaborators based on a Boltzmann-like kinetic theory of integrate-and-fire networks [14, 65], which is itself a dimension reduction by moment closure of the so-called population density method [56, 58]. The population density approach is a numerical scheme for tracking the probability density of a population of spiking neurons based on solutions of an underlying partial differential equation. In the case of simple neuron models, this can be considerably more efficient than classical Monte Carlo simulations that follow the states of each neuron in the network. On the other hand, as the complexity of the individual neuron model increases, the gain in efficiency of the population density method decreases, and this has motivated the development of moment closure schemes. However, as recently shown by Ly and Tranchina [52], considerable care must be taken when carrying out the dimension reduction, since it can lead to an ill-posed problem over a wide range of physiological parameters. That is, the truncated moment equations may not support a steady-state solution even though a steady-state probability density exists for the full system.

Although there has been some progress in analyzing the stochastic dynamics of fully connected or sparsely connected networks, there has been relatively little work in developing stochastic versions of continuum neural field models. One notable exception is the so-called path integral approach of Buice and Cowan [12] (see also [13]). This can be viewed as a minimal extension of a deterministic neural field model, in which the dynamics is described by a master equation whose state transition rates are chosen such that the original model is recovered in the mean-field limit. The latter is obtained by setting all second- and higher-order cumulants to zero. Adapting methods from nonequilibrium statistical physics previously applied to reaction-diffusion systems [19, 20, 60], Buice and Cowan [12] have shown how to construct a path integral representation of the associated moment generating functional. This path integral representation has two particularly useful features. First, diagrammatic perturbation methods (Feynman graphs) can be used to generate a truncated moment hierarchy, and thus determine corrections to mean-field theory involving coupling to two-point and higher-order cumulants [12, 13]. Second, renormalization group methods can be used to derive scaling laws for statistical correlations close to criticality, that is, close to a bifurcation point of the underlying deterministic neural field model [12].

In this paper we revisit the path integral approach of Buice and Cowan [12] in order to determine how it is related to a well-known perturbative approach to analyzing master equations, namely, the Van Kampen system-size expansion [47]. We first briefly describe the construction of population-based rate models, which reduce to neural field equations in an appropriate continuum limit, and introduce the master equation formulation of stochastic neurodynamics due to Buice and Cowan [12] (see section 2). However, we consider a rescaled version of the Buice and Cowan [12] master equation in which population activity is taken to be the fraction rather

than mean number of active neurons in each population. This allows us to treat the size N of each population as a system-size parameter and, hence, to carry out an expansion of the master equation under the assumption that fluctuations about the mean-field solutions are $\mathcal{O}(N^{-1/2})$ (see section 3). The associated moment equations can then be expressed as series expansions in powers of N^{-1} such that the coupling between the n th moment and the $(n+p)$ th moment is $\mathcal{O}(N^{-p})$. This provides a systematic procedure for truncating the moment hierarchy provided that the underlying mean-field system is not operating close to criticality. Taking an appropriate continuum limit that preserves the system-size expansion, we calculate the lowest order corrections to mean-field theory, which yields a pair of coupled integrodifferential equations describing the dynamics of the mean and covariance of population activity. We then show how these equations can also be derived by carrying out a $1/N$ loop expansion of the path integral representation of the master equation based on the steepest descents or the saddle-point method (see section 4). A similar path integral expansion was previously carried out by Buice and Cowan [12] and Buice, Cowan, and Chow [13], except that for their choice of scaling there does not exist a natural system-size or loop parameter N . Hence, it is necessary to consider different moment closure conditions, and this leads to different corrections to mean-field theory. (Note that the use of path integrals to carry out an effective system-size expansion has also been developed within the context of the Kuramoto model of coupled phase oscillators [39, 40].) We conclude the paper by showing that for our choice of scaling, the path integral representation of the master equation can be used to investigate large deviations or rare event statistics underlying escape from the basin of attraction of a stable fixed point of the mean-field dynamics (see section 5), following along similar lines to previous work on large deviations in reaction-diffusion systems [38, 21, 22, 29]. Such an analysis is not possible using the standard Van Kampen system-size expansion, since the latter cannot accurately determine exponentially small transitions. Finally, note that in this paper we consider the basic analytical formalism of the master equation approach to stochastic neurodynamics. Elsewhere we will present a numerical study based on direct Monte Carlo simulations that track the state of each neuron in a population, comparing the different choices of scaling for the master equation and characterizing the resulting statistical behavior as a function of the system size and network topology.

2. Master equation for neural field models. In this section we outline the construction of rate-based models of interacting neural populations, which in the continuum limit are described by neural field equations. Such equations can be voltage-based or activity-based [63, 24, 35]. We then show how a stochastic version of the neural field equations can be introduced using a rescaled version of the master equation formulation of Buice and Cowan [12] and Buice, Cowan, and Chow [13].

2.1. Population-based rate models. Suppose that there exist M homogeneous populations each consisting of N neurons labeled by $p \in \mathcal{I}_i$, $i = 1, \dots, M$, and $|\mathcal{I}_i| = N$. Assume that all neurons of a given population are equivalent in the sense that the interaction between neuron $p \in \mathcal{I}_i$ and neuron $q \in \mathcal{I}_j$ depends only on i and j . Each neuron p in population i receives input from all neurons q in population j with strength $w_{ij} = W_{ij}/N$ such that W_{ij} is independent of N . The total synaptic

current to neuron $p \in \mathcal{I}_i$ from the various populations is then taken to be of the form

$$(2.1) \quad a_i(t) = \sum_{j=1}^M W_{ij} \int_0^\infty \epsilon(\tau) \nu_j^{(N)}(t - \tau) d\tau,$$

where the kernel ϵ is the postsynaptic response to a single action potential or spike and $\nu_j^{(N)}(t)$ is the population activity

$$(2.2) \quad \nu_j^{(N)}(t) = \frac{1}{N} \sum_{q \in \mathcal{I}_j} \sum_{m \in \mathbb{Z}} \delta(t - T_q^m).$$

The double sum runs over all firing times T_q^m of all neurons in the population. Let us now take the thermodynamic limit $N \rightarrow \infty$ with $\nu_j(t) = \lim_{N \rightarrow \infty} \nu_j^{(N)}$. The macroscopic variable $\nu_j(t)\delta t$ represents the total activity of the j th population, that is, the fraction of neurons firing in the interval $[t, t + \delta t)$. We can then define an asynchronous state as one with constant activity $\nu_j(t) = \nu_j^*$ for all $j = 1, \dots, M$ [1, 36, 10, 9, 66]. It follows that the total synaptic current is also time independent, $a_i(t) = a_i^*$, with

$$(2.3) \quad a_i^* = \sum_j W_{ij} \nu_j^*.$$

We have assumed that ϵ is normalized according to $\int_0^\infty \epsilon(\tau) d\tau = 1$. Suppose that each homogeneous population's steady-state activity is related to the synaptic input according to some effective gain function which we denote by f , $\nu_j^* = f(a_j^*)$ [1, 36, 10, 9]. Substitution into (2.3) then yields the steady-state equation

$$(2.4) \quad a_i^* = \sum_j W_{ij} f(a_j^*).$$

We shall assume throughout that f is a positive, monotonically increasing bounded function such as a sigmoid function.

The above analysis can be extended to the case of time-varying activities provided that each population remains in a state of incoherent firing so that the synaptic input is slowly varying in time. That is, under the approximation $\nu_j(t) = f(a_j(t))$ we obtain the closed integral equation

$$(2.5) \quad a_i(t) = \sum_j W_{ij} \int_0^\infty \epsilon(\tau) f(a_j(t - \tau)) d\tau.$$

In the case of exponential synapses, $\epsilon(\tau) = \alpha e^{-\alpha\tau}$ with $\alpha^{-1} = \tau_m$ identified as a membrane or synaptic time constant, (2.5) reduces to the form

$$(2.6) \quad \tau_m \frac{da_i}{dt} = -a_i(t) + \sum_j W_{ij} f(a_j(t)).$$

The rate equation (2.6) is expressed in terms of the postsynaptic current $a_i(t)$ and the sum over weights is outside the nonlinear gain function f . Sometimes $a_i(t)$ is identified with the postsynaptic potential rather than current, and (2.6) is referred to as a voltage-based rate model [15, 43, 24]. Note that the reduction breaks down

when the population activity changes rapidly relative to τ_m due to fast transients or the existence of a coherent state exhibiting collective oscillations, such as when the asynchronous state becomes unstable. An alternative version of the rate equation is obtained in cases where the postsynaptic response ϵ is relatively fast compared to the response of population activity to changes in synaptic input. If we model the latter as a low-pass filter and make the approximation $\epsilon(\tau) = \delta(\tau)$, then we obtain the so-called activity-based rate model [72, 73, 63]

$$(2.7) \quad \tau_s \frac{d\nu_i}{dt} = -\nu_i(t) + f \left(\sum_j W_{ij} \nu_j(t) \right).$$

2.2. Continuum limit. So far we have formulated the network dynamics in terms of a set of interacting homogeneous populations. This is motivated by the observation that neurons in cortex with similar response properties tend to be arranged in vertical columns. A classical example is the columnar-like arrangement of neurons in primary visual cortex that prefer stimuli of similar orientation [45]. Neurons within a cortical column share many inputs and are tightly interconnected, so that it is reasonable to consider the mean activity $\nu(t)$ of the given neural population, rather than keeping track of the spiking of individual cells. At the next spatial scale beyond cortical columns one observes interactions between columns that have a distinctive spatial arrangement. Roughly speaking, intercolumnar interactions depend on the distance between the interacting columns. This motivates us to consider a spatially distributed network of neural populations. For concreteness, we consider a one-dimensional network, although the construction is easily extended to higher spatial dimensions.

Suppose that there is a uniform density ρ of neural populations (columns) distributed along the x axis. We partition the x axis into discrete intervals of length Δx within which there are $\rho\Delta x$ populations. Let us denote the set of populations in the interval $[m\Delta x, (m+1)\Delta x)$ by $\mathcal{N}(m\Delta x)$. As a further approximation, suppose that the weights W_{ij} are slowly varying on the length scale Δx so that we can write

$$W_{ij} = \omega(m\Delta x, n\Delta x)$$

for all $i \in \mathcal{N}(m\Delta x)$ and $j \in \mathcal{N}(n\Delta x)$. The sum over all populations j in the rate equation (2.7) can now be decomposed as

$$\sum_j W_{ij} \nu_j = \sum_n \omega(m\Delta x, n\Delta x) \sum_{k \in \mathcal{N}(n\Delta x)} \nu_k.$$

Define the local spatially averaged activity variable $\nu(n\Delta x)$ according to

$$(2.8) \quad \nu(n\Delta x) = \frac{1}{\rho\Delta x} \sum_{k \in \mathcal{N}(n\Delta x)} \nu_k,$$

where $\rho\Delta x$ is the number of populations in each set $\mathcal{N}(n\Delta x)$. Spatially averaging the rate equation (2.7) with respect to $i \in \mathcal{N}(m\Delta x)$ then gives

$$(2.9) \quad \tau_s \frac{d\nu(m\Delta x, t)}{dt} = -\nu(m\Delta x, t) + f \left(\rho\Delta x \sum_n \omega(m\Delta x, n\Delta x) \nu(n\Delta x, t) \right).$$

Finally, setting $x = m\Delta x$ and $y = n\Delta x$ and taking the continuum limit $\Delta x \rightarrow 0$ we obtain the Wilson–Cowan equation

$$(2.10) \quad \tau_s \frac{\partial}{\partial t} \nu(x, t) = -\nu(x, t) + f \left(\rho \int \omega(x, y) \nu(y, t) dy \right).$$

Note that our choice of continuum limit involves grouping together local populations in an infinitesimal interval Δx rather than distributing neurons within a local population across an interval Δx [35].

A similar analysis can be applied to the voltage-based rate equation (2.6). In this case we make the approximation

$$\begin{aligned} \sum_j W_{ij} f(a_j) &= \sum_n \omega(m\Delta x, n\Delta x) \sum_{k \in \mathcal{N}(n\Delta x)} f(a_k) \\ &\approx \rho \Delta x \sum_n \omega(m\Delta x, n\Delta x) f \left(\sum_{k \in \mathcal{N}(n\Delta x)} a_k \right) \\ &= \rho \Delta x \sum_n \omega(m\Delta x, n\Delta x) f(a(n\Delta x, t)). \end{aligned}$$

Spatially averaging (2.6) with respect to $i \in \mathcal{N}(m\Delta x)$ and taking the continuum limit gives

$$(2.11) \quad \tau_m \frac{\partial}{\partial t} a(x, t) = -a(x, t) + \rho \int \omega(x, y) f(a(y, t)) dy.$$

(2.10) and (2.11) are examples of neural field equations, which have been widely used to study large-scale spatiotemporal dynamics of cortical tissue [72, 73, 2, 25, 24, 8, 7, 16]. They can easily be extended to higher spatial dimensions as well as to separate populations of excitatory and inhibitory neurons.

2.3. Rate models as Markov processes. The construction of the rate models (2.6) and (2.7) relies on various forms of spatial and temporal averaging. In particular, each interacting population is assumed to be homogeneous, and the thermodynamic limit is taken so that an asynchronous state of constant population activity can be defined. The resulting neural field equation is then deterministic even though the spike trains of individual neurons within each homogeneous population is stochastic [1, 36, 10, 9]. A more realistic network model would have a finite number of neurons in each population, would allow for heterogeneities within and between populations, and would incorporate various sources of intrinsic and extrinsic noise.

In this paper we follow Buice and Cowan [12] by considering a phenomenological master equation approach to incorporating noise into neural field models. However, in contrast to Buice and Cowan [12], we keep N finite rather than taking the thermodynamic limit $N \rightarrow \infty$ and keep track of how various quantities such as the synaptic weights scale with N . This will allow us to carry out a system-size expansion of the master equation along the lines of Van Kampen [47] (section 3). It also provides a natural small parameter, $1/N$, that can be used to carry out a systematic loop expansion in the path integral approach developed previously by Buice and Cowan [12] (section 4). In order to illustrate the basic idea, let us write down a stochastic version of the rate equation (2.7). The first step is to reinterpret the population activity $\nu_i^{(N)}(t)$. We now assume that neurons in each homogeneous population can be

in one of two states, quiescent or active, i.e., in the process of generating an action potential. Suppose that in the interval $[t, t + \Delta t)$ there are $m_i(t)$ active neurons in the i th population. It follows that the activity of this population is given by

$$(2.12) \quad \nu_i^{(N)}(t) = \frac{m_i(t)}{N\Delta t}.$$

We will take Δt to be equal to the width of an action potential (approximately 1msec) and fix time units such that $\Delta t = 1$. The state or configuration of the network is now specified by the vector $\mathbf{m}(t) = (m_1(t), m_2(t), \dots, m_M(t))$, where each $m_i(t)$ is treated as a stochastic variable that evolves according to a one-step jump Markov process. The rates for the state transitions $m_i \rightarrow m_i \pm 1$ are then chosen so that in the thermodynamic limit $N \rightarrow \infty$ one recovers the deterministic rate equation (2.7). Let $P(\mathbf{n}, t) = \text{Prob}[\mathbf{m}(t) = \mathbf{n}]$ denote the probability that the network of interacting populations has configuration $\mathbf{n} = (n_1, n_2, \dots, n_M)$ at time $t, t > 0$, given some initial distribution $P(\mathbf{n}, 0)$.¹ The probability distribution evolves according to the master equation [12]

$$(2.13) \quad \frac{dP(\mathbf{n}, t)}{dt} = \alpha \sum_i \left[(n_i + 1)P(\mathbf{n}_{i+}) - n_i P(\mathbf{n}, t) + Nf \left(W_{ii}(n_i - 1)/N \right. \right. \\ \left. \left. + \sum_{j \neq i} W_{ij} n_j / N \right) P(\mathbf{n}_{i-}, t) - Nf \left(\sum_j W_{ij} n_j / N \right) P(\mathbf{n}, t) \right]$$

for $0 \leq n_i \leq N$. Here $\mathbf{n}_{i\pm}$ denotes the configuration with n_i replaced by $n_i \pm 1$ and $\alpha = 1/\tau_s$. Equation (2.14) is supplemented by the boundary conditions $P(\mathbf{n}, t) \equiv 0$ if $n_i = N + 1$ or $n_i = -1$ for some i . The master equation preserves the normalization condition $\sum_{i=1}^M \sum_{n_i=0}^N P(\mathbf{n}, t) = 1$ for all $t \geq 0$.

The one major difference between (2.13) and the corresponding master equation of Buice and Cowan [12] is the relationship between the activity variable ν_i in the rate equation (2.7) and the mean $\langle m_i(t) \rangle \equiv \sum_{\mathbf{n}} P(\mathbf{n}, t) n_i$. Buice and Cowan make the identification $\nu_i = \langle m_i \rangle$ and allow $m_i(t)$ to be unbounded; that is, they first take the thermodynamic limit $N \rightarrow \infty$ with Nf replaced by f . The rate equation (2.7) is then recovered by taking a mean-field limit in which all second- and higher-order correlations are set to zero. This particular choice of scaling effectively assumes that the neurons in each local population are strongly correlated such that the population activity is Poisson-like [13, 11]. On the other hand, we take $\nu_i = \langle m_i \rangle / N$ and $w_{ij} = W_{ij} / N$ with W_{ij} independent of N , where N is the size of each local population. We then naturally recover the rate equation (2.7) in the thermodynamic limit $N \rightarrow \infty$ provided that f is scaled by the factor N in the master equation. (In general one might expect the gain function f itself to depend on the size N , but we ignore this additional complication here.) Our choice of scaling is motivated by the analysis of asynchronous states in homogeneous networks [1, 36, 10, 9, 66]. That is, we consider a regime in which the asynchronous state of each local population is stable so that the firing patterns of the neurons within a local population are approximately

¹In probability theory it is conventional to distinguish between a random variable and the particular value it takes by writing the former as uppercase and the latter as lowercase. In this paper we use lowercase for both. Thus $\langle m(t) \rangle = n$ states that the expectation value of the random variable $m(t)$ at time t is n .

uncorrelated for large N and the law of large numbers holds. From a mathematical perspective, there is a major difference between dealing with finite N and taking the thermodynamic limit ab initio. In the former case, the master equation describes a jump process on a large but finite-dimensional state space consisting of N^M states. Assuming that the weight matrix \mathbf{W} is irreducible, it follows that the corresponding transition matrix of the master equation (2.13) is also irreducible and, hence, there exists a unique stationary solution [47]. However, the presence of the nonlinear gain function f means that it is not possible to obtain exact solutions of the master equation so that some approximation scheme is required. In this paper, we will consider two such approximation schemes, one based on the Van Kampen system-size expansion (section 3) and the other based on path integrals and steepest descents (section 4). Both approaches assume that $m_i(t)$ is unbounded, while treating N as a fixed parameter of the transition rates, and involve perturbation expansions in the small parameter N^{-1} about deterministic (mean-field) solutions. In general, these perturbative approximations will not generate the correct stationary behavior of the finite system in the limit $t \rightarrow \infty$. However, they can still capture quasi-stationary (metastable) states in which the finite network spends a long but finite time.

Another important observation is that the construction of the master equation is nonunique. For example, any master equation of the form

$$(2.14) \quad \frac{dP(\mathbf{n}, t)}{dt} = \alpha \sum_i [(n_i + 1)G_i(\mathbf{n}_{i+}/N)P(\mathbf{n}_{i+}, t) - n_i G_i(\mathbf{n}/N)P(\mathbf{n}, t) + N(H_i(\mathbf{n}_{i-}/N)P(\mathbf{n}_{i-}, t) - H_i(\mathbf{n}/N))P(\mathbf{n}, t)]$$

will reduce to the rate equation (2.7) in the large- N limit, provided that

$$(2.15) \quad G_i(\boldsymbol{\nu}) = 1 + \Gamma_i(\boldsymbol{\nu}), \quad H_i(\boldsymbol{\nu}) = f\left(\sum_j W_{ij}\nu_j\right) + \nu_i \Gamma_i(\boldsymbol{\nu}),$$

where Γ_i is an arbitrary positive function. This raises two issues. How do corrections to mean-field theory depend on Γ_i ? Are there additional constraints that can determine Γ_i ? We will attempt to address the first question in this paper. A master equation of the general form (2.14) can also be constructed for the voltage-based model given by (2.6). However, we first need to write down a rate equation for the activity $\nu_i(t)$. This can be achieved by an application of the chain rule:

$$(2.16) \quad \frac{d\nu_i}{dt} = f'(a_i) \frac{da_i}{dt}.$$

Assuming that f is invertible such that $a_i = g(\nu_i)$ and using (2.6), we find that

$$(2.17) \quad \frac{d\nu_i}{dt} = \frac{\alpha}{g'(\nu_i)} \left[-g(\nu_i) + \sum_j W_{ij}\nu_j \right],$$

where we have used the relation $f'(a) = 1/g'(\nu)$. We are assuming that $g'(\nu) \neq 0$ for all ν . By analogy with the previous case, the master equation (2.14) reduces to the above rate equation in the thermodynamic limit provided that

$$(2.18) \quad G_i(\boldsymbol{\nu}) = \frac{g(\nu_i)}{\nu_i g'(\nu_i)} + \Gamma_i(\boldsymbol{\nu}), \quad H_i(\boldsymbol{\nu}) = \frac{1}{g'(\nu_i)} \sum_j W_{ij}\nu_j + \nu_i \Gamma_i(\boldsymbol{\nu}).$$

In this case, Γ_i should be chosen such that both G_i and H_i are positive transition rates. That is, in general the weight matrix W_{ij} includes both excitatory (positive) and inhibitory (negative) synapses. One possibility is to choose Γ_i such that all inhibitory inputs contribute to the transition rate G_i .

In this paper we will develop both the Van Kampen system-size expansion and the path integral approach for the general master equation (2.14). Our results can then be applied to activity-based and voltage-based neural field models using (2.15) and (2.18), respectively, and taking the appropriate continuum limit. However, we will illustrate our results using the simpler activity-based model in order to compare directly with Buice and Cowan [12].

3. System-size expansion. In this section we carry out a system-size expansion of the general master equation (2.14), following the method originally developed by Van Kampen [47] in the context of chemical reactions and birth-death processes. For simplicity, we will assume that the initial state of the network is known so that $m_i(0) = \bar{n}_i$ with probability one

$$(3.1) \quad P(\mathbf{n}, 0) = \prod_i \delta_{n_i, \bar{n}_i}.$$

Note that a system-size expansion for a single homogeneous network was previously carried out by Ohira and Cowan [57]. Let us rewrite (2.14) in the more compact form

$$(3.2) \quad \frac{dP(\mathbf{n}, t)}{dt} = \alpha \sum_i [(\mathbb{E}_i - 1) [n_i G_i(\mathbf{n}/N) P(\mathbf{n}, t)] + N (\mathbb{E}_i^{-1} - 1) [H_i(\mathbf{n}/N) P(\mathbf{n}, t)]],$$

where \mathbb{E}_i is a translation operator: $\mathbb{E}_i^{\pm 1} F(\mathbf{n}) = F(\mathbf{n}_{i\pm})$ for any function F . Perform the change of variables

$$(3.3) \quad (n_i, t) \rightarrow (N\bar{\nu}_i(t) + \sqrt{N}\xi_i, t),$$

where $\bar{\nu}_i(t)$ is some time-dependent function to be determined below, and set

$$(3.4) \quad \Pi(\boldsymbol{\xi}, t) = P(N\bar{\boldsymbol{\nu}}(t) + \sqrt{N}\boldsymbol{\xi}, t).$$

Note that since n_i is integer valued, ξ_i is also a discrete variable but takes on values separated by integer multiples of $1/\sqrt{N}$. Thus, in the large- N limit we can treat ξ_i as continuous valued. Differentiation with respect to time gives

$$(3.5) \quad \begin{aligned} \frac{\partial P(\mathbf{n}, t)}{\partial t} &= \frac{\partial \Pi(\boldsymbol{\xi}, t)}{\partial t} + \sum_{j=1}^M \frac{\partial \xi_j}{\partial t} \frac{\partial \Pi(\boldsymbol{\xi}, t)}{\partial \xi_j} \\ &= \frac{\partial \Pi(\boldsymbol{\xi}, t)}{\partial t} - \sqrt{N} \sum_{j=1}^M \frac{\partial \bar{\nu}_j(t)}{\partial t} \frac{\partial \Pi(\boldsymbol{\xi}, t)}{\partial \xi_j}, \end{aligned}$$

where we have used the fact that the variable n_i is kept fixed when taking the partial derivative with respect to t . The next step is to set $n_i/N = \bar{\nu}_i(t) + N^{-1/2}\xi_i$ in the coefficients of the master equation (3.2) and to Taylor expand with respect to $\boldsymbol{\xi}/\sqrt{N}$.

For example, denoting by $\mathbf{1}_i$ the vector with components $[\mathbf{1}_i]_j = \delta_{i,j}$ we have

$$(3.6) \quad \mathbb{E}_i^{-1}[H_i(\mathbf{n}/N)P(\mathbf{n}, t)] = H_i\left(\bar{\mathbf{v}}(t) + N^{-1/2}(\boldsymbol{\xi} - N^{-1/2}\mathbf{1}_i)\right)\Pi(\boldsymbol{\xi} - N^{-1/2}\mathbf{1}_i, t) \\ = \left(H_i(\bar{\mathbf{v}}(t)) - N^{-1/2} \sum_j \frac{\partial H_i(\bar{\mathbf{v}}(t))}{\partial \phi_j} (\xi_j - N^{-1/2}\delta_{i,j}) + \dots \right) \\ \times \left(\Pi(\boldsymbol{\xi}, t) - N^{-1/2} \frac{\partial \Pi(\boldsymbol{\xi}, t)}{\partial \xi_i} + \frac{N^{-1}}{2} \frac{\partial^2 \Pi(\boldsymbol{\xi}, t)}{\partial \xi_i^2} + \dots \right).$$

The final result is a Taylor expansion of the master equation in powers of $N^{-1/2}$. Identifying the leading order $N^{1/2}$ terms gives

$$(3.7) \quad \sum_i \frac{d\bar{v}_i(t)}{dt} \frac{\partial \Pi(\boldsymbol{\xi}, t)}{\partial \xi_i} = \alpha \sum_{i=1}^M [-\bar{v}_i(t)G_i(\bar{\mathbf{v}}(t)) + H_i(\bar{\mathbf{v}}(t))] \frac{\partial \Pi(\boldsymbol{\xi}, t)}{\partial \xi_i}.$$

These terms cancel provided that we take $\bar{\mathbf{v}}(t)$ to be a solution of the rate equation

$$(3.8) \quad \tau_s \frac{d\bar{\mathbf{v}}_i(t)}{dt} = F_i(\bar{\mathbf{v}}(t)) \equiv -\bar{v}_i(t)G_i(\bar{\mathbf{v}}(t)) + H_i(\bar{\mathbf{v}}(t)).$$

Note that substitution of (2.15) or (2.18) into (3.8) recovers the corresponding rate equation (2.7) or (2.6). Assuming that $\bar{\mathbf{v}}(t)$ satisfies (3.8), the master equation expansion to order $N^{-1/2}$ is then

$$(3.9) \quad \tau_s \frac{\partial \Pi}{\partial t} = - \sum_i \frac{\partial}{\partial \xi_i} \left[F_i^{(1)}(\boldsymbol{\xi}, \bar{\mathbf{v}})\Pi(\boldsymbol{\xi}, t) \right] + \frac{1}{2} \sum_i \frac{\partial^2}{\partial \xi_i^2} \left[F_i^{(2)}(\boldsymbol{\xi}, \bar{\mathbf{v}})\Pi(\boldsymbol{\xi}, t) \right] \\ - \frac{N^{-1/2}}{6} \sum_i \frac{\partial^3}{\partial \xi_i^3} [F_i(\bar{\mathbf{v}})\Pi(\boldsymbol{\xi}, t)] + \mathcal{O}(N^{-1}),$$

where

$$(3.10) \quad F_i^{(1)}(\boldsymbol{\xi}, \bar{\mathbf{v}}) = \sum_k A_{ik}(\bar{\mathbf{v}})\xi_k + \frac{N^{-1/2}}{2} \sum_{k,l} A_{ikl}(\bar{\mathbf{v}})\xi_k \xi_l + \mathcal{O}(N^{-1}),$$

$$(3.11) \quad F_i^{(2)}(\boldsymbol{\xi}, \bar{\mathbf{v}}) = B_i(\bar{\mathbf{v}}) + N^{-1/2} \sum_k B_{ik}(\bar{\mathbf{v}})\xi_k + \mathcal{O}(N^{-1}),$$

and

$$(3.12) \quad A_{ik}(\bar{\mathbf{v}}) = \frac{\partial F_i(\bar{\mathbf{v}})}{\partial \bar{v}_k}, \quad A_{ikl}(\bar{\mathbf{v}}) = \frac{\partial^2 F_i(\bar{\mathbf{v}})}{\partial \bar{v}_k \partial \bar{v}_l},$$

$$(3.13) \quad B_i(\bar{\mathbf{v}}) = \bar{v}_i G_i(\bar{\mathbf{v}}) + H_i(\bar{\mathbf{v}}), \quad B_{ik}(\bar{\mathbf{v}}) = \frac{\partial B_i(\bar{\mathbf{v}})}{\partial \bar{v}_k}.$$

3.1. Linear noise approximation. Keeping only the $\mathcal{O}(1)$ terms in (3.9) leads to the linear Fokker–Planck equation [47]

$$(3.14) \quad \tau_s \frac{\partial \Pi}{\partial t} = - \sum_{i,k} A_{ik}(\bar{\mathbf{v}}) \frac{\partial}{\partial \xi_i} [\xi_k \Pi(\boldsymbol{\xi}, t)] + \frac{1}{2} \sum_i B_i(\bar{\mathbf{v}}) \frac{\partial^2}{\partial \xi_i^2} \Pi(\boldsymbol{\xi}, t)$$

with $\bar{\mathbf{v}}(t)$ evolving according to (3.8). Suppose, for the moment, that we treat $\Pi(\boldsymbol{\xi}, t)$ as a probability density on \mathbb{R}^M such that

$$\prod_{i=1}^M \int_{-\infty}^{\infty} d\xi_i \Pi(\boldsymbol{\xi}, t) = 1.$$

In order to represent $\Pi(\boldsymbol{\xi}, t)$ as a probability density in the large- N limit, we have to introduce a scale factor in (3.4) of the form $P = N^{-M/2}\Pi$ and make the approximation $N^{-1/2}\sum_{\xi_i} \approx \int d\xi_i$ for each $i = 1, \dots, M$. We are also assuming that there is a negligible probability of the stochastic process approaching the boundaries at $m_i = 0, N$ so that we can take $-\infty < \xi_i < \infty$. Under these simplifying assumptions, the linear noise approximation decomposes the discrete random variables $m_i(t)$ into a deterministic component and a stochastic component such that

$$(3.15) \quad N^{-1}m_i(t) = \bar{\nu}_i(t) + N^{-1/2}\xi_i(t)$$

with $\xi_i(t)$ approximated by a multivariate Gaussian process corresponding to solutions of the Fokker–Planck equation (3.14). In other words, we can represent the stochastic dynamics in terms of $\mathcal{O}(N^{-1/2})$ fluctuations around a background deterministic trajectory, with the latter corresponding to a solution of the rate equation (3.8).

A Gaussian process is completely specified by its first- and second-order statistics. Let us define the mean

$$(3.16) \quad \langle \xi_i(t) \rangle \equiv \int \xi_i \Pi(\boldsymbol{\xi}, t) d\boldsymbol{\xi} = N^{-1/2} \langle m_i(t) \rangle$$

and covariance

$$(3.17) \quad C_{kl}(t) \equiv \langle \xi_k(t)\xi_l(t) \rangle - \langle \xi_k(t) \rangle \langle \xi_l(t) \rangle = N^{-1} [\langle m_k(t)m_l(t) \rangle - \langle m_k(t) \rangle \langle m_l(t) \rangle].$$

Taking moments of the Fokker–Planck equation (3.14), it can be shown that (see also section 3.2)

$$(3.18) \quad \tau_s \frac{\partial \langle \xi_p \rangle}{\partial t} = \sum_k \frac{\partial F_p(\bar{\boldsymbol{\nu}})}{\partial \bar{\nu}_k} \langle \xi_k \rangle$$

and

$$(3.19) \quad \frac{d\mathbf{C}}{dt} = \mathbf{A}(\bar{\boldsymbol{\nu}})\mathbf{C} + \mathbf{C}\mathbf{A}(\bar{\boldsymbol{\nu}})^T + \mathbf{B}(\bar{\boldsymbol{\nu}}),$$

where the matrix components \mathbf{A}_{kl} are given by (3.12) and $\mathbf{B}_{kl} = B_k\delta_{k,l}$ with B_k given by (3.13). The deterministic initial conditions (3.1) imply that we can set $\langle \xi_i(0) \rangle = 0$. (3.18) then implies that $\langle \xi(t) \rangle = 0$ for all $t > 0$ and, hence, $\langle m_i(t) \rangle = N\bar{\nu}_i(t)$. Suppose that the rate equation (3.8) has a unique stable fixed point solution $\boldsymbol{\nu}^*$. It follows that the Fokker–Planck equation has a stationary solution given by the multidimensional Gaussian

$$(3.20) \quad \Pi(\boldsymbol{\xi}) = \frac{1}{\sqrt{(2\pi)^M \det \mathbf{C}}} \exp \left(- \sum_{k,l} \xi_k C_{kl} \xi_l \right)$$

with the stationary covariance matrix satisfying the equation

$$(3.21) \quad \mathbf{A}(\boldsymbol{\nu}^*)\mathbf{C} + \mathbf{C}\mathbf{A}(\boldsymbol{\nu}^*)^T + \mathbf{B}(\boldsymbol{\nu}^*) = 0.$$

The autocorrelation matrix of the stationary process, $\Xi_{ij}(t) = \langle \xi_i(t)\xi_j(s) \rangle$, is given by

$$(3.22) \quad \Xi(t) = \mathbf{C}e^{\mathbf{A}(\boldsymbol{\nu}^*)|t-s|}.$$

This can be derived by noting that [47], for $t > s$,

$$\langle \xi_i(t)\xi_j(s) \rangle = \langle \xi_j(s)\langle \xi_i(t) \rangle_{\xi(s)} \rangle = \langle \xi_j(s)\xi_i(s) \rangle e^{\mathbf{A}(\nu^*)|t-s|}.$$

That is, we solve the first moment equation (3.18) given the initial condition $\xi(s)$ and then set $\langle \xi_i(s)\xi_j(s) \rangle = C_{ij}$ after averaging with respect to the stationary distribution. A similar result holds for $t < s$. Noting from (3.12) that the matrix $\mathbf{A}(\nu^*)$ is the Jacobian of the fixed point ν^* , it follows that $\mathbf{A}(\nu^*)$ has negative definite eigenvalues and hence $\Xi(t) \rightarrow 0$ as $t \rightarrow \infty$. However, (3.21) implies that \mathbf{C} becomes singular as $\det \mathbf{A}(\nu^*) \rightarrow 0$. This reflects the well-known result that the standard system-size expansion breaks down at critical points [47]. That is, fluctuations near the critical point are enhanced so that as the critical point is approached more and more terms in the system-size expansion are required to maintain a given accuracy. At the critical point itself, it is usually necessary to carry out the expansion in powers of $N^{-\mu}$ for some $\mu \neq 1/2$ [47]. One of the interesting aspects of the path integral approach developed by Buiuc and Cowan [12] is that renormalization group methods can be used to extract the power law behavior of statistical correlations in the vicinity of a critical point.

For a general weight matrix W_{ij} we expect the rate equations (2.6) and (2.7) to have multiple fixed points [43]. In this case, the system-size expansion can still be used to analyze the effects of fluctuations well within the basin of attraction of a locally stable fixed point. However, there is now a small probability that there is a noise-induced transition to the basin of attraction of another fixed point or to an absorbing boundary. Since the probability of such a transition is usually of order e^{-N} except close to the boundary of the basin of attraction, such a contribution cannot be analyzed accurately by carrying out a perturbation expansion in $N^{-1/2}$ and using standard Fokker–Planck methods [47]. These exponentially small transitions play a crucial role in allowing the network to approach the unique stationary state (if it exists) in the asymptotic limit $t \rightarrow \infty$. In section 5 we will show how the path integral representation of the master equation presented in section 4 can be used to estimate the lifetime of the active metastable state, following recent work on the analysis of rare event statistics in reaction-diffusion systems [38, 21, 22, 29].

3.2. Moment equations. Returning to the more general Taylor expanded master equation (3.9), we can generate a hierarchy of moment equations by multiplying both sides with products of components of ξ , integrating with respect to ξ , and then performing successive integrations by parts. Our choice of scaling with respect to the system size N means that we can systematically truncate this moment hierarchy by specifying the order of N . For simplicity, we will truncate at order $N^{-1/2}$. (The accuracy of the truncation will depend on how close the system is to criticality; see section 3.1.) As in the $\mathcal{O}(1)$ linear noise approximation, we treat $\Pi(\xi, t)$ as a probability density with the appropriate normalization and ignore boundary effects.

In the case of the first moment,

$$\begin{aligned} \tau_s \frac{\partial \langle \xi_i \rangle}{\partial t} &= - \sum_j \int d\xi \xi_i \frac{\partial}{\partial \xi_j} \left[F_j^{(1)}(\xi, \bar{\nu}) \Pi(\xi, t) \right] \\ (3.23) \qquad &= \int d\xi F_i^{(1)}(\xi, \bar{\nu}) \Pi(\xi, t). \end{aligned}$$

Substituting for $F_i^{(1)}$ using (3.10) and (3.12) then gives

$$\begin{aligned}
 \tau_s \frac{\partial \langle \xi_i \rangle}{\partial t} &= \sum_k A_{ik}(\bar{\nu}) \langle \xi_k \rangle + \frac{N^{-1/2}}{2} \sum_{k,l} A_{ikl}(\bar{\nu}) \langle \xi_k \xi_l \rangle + \mathcal{O}(N^{-1}) \\
 (3.24) \qquad &= \sum_k \frac{\partial F_i(\bar{\nu})}{\partial \bar{\nu}_k} \langle \xi_k \rangle + \frac{N^{-1/2}}{2} \sum_{k,l} \frac{\partial^2 F_i(\bar{\nu})}{\partial \bar{\nu}_k \partial \bar{\nu}_l} \langle \xi_k \xi_l \rangle + \mathcal{O}(N^{-1}).
 \end{aligned}$$

Similarly, the second moment equation is

$$\begin{aligned}
 \tau_s \frac{\partial \langle \xi_i \xi_j \rangle}{\partial t} &= - \sum_k \int d\xi \xi_i \xi_j \frac{\partial}{\partial \xi_k} \left[F_k^{(1)}(\xi, \bar{\nu}) \Pi(\xi, t) \right] \\
 &\quad + \frac{1}{2} \sum_k \int d\xi \xi_i \xi_j \frac{\partial^2}{\partial \xi_k^2} \left[F_k^{(2)}(\xi, \bar{\nu}) \Pi(\xi, t) \right] \\
 &= \int d\xi \left[\xi_j F_i^{(1)}(\xi, \bar{\nu}) + \xi_i F_j^{(1)}(\xi, \bar{\nu}) \right] \Pi(\xi, t) \\
 (3.25) \qquad &\quad + \delta_{i,j} \int d\xi F_i^{(2)}(\xi, \bar{\nu}) \Pi(\xi, t).
 \end{aligned}$$

Substituting for $F_i^{(2)}$ using (3.11) then gives

$$\begin{aligned}
 \tau_s \frac{\partial \langle \xi_i \xi_j \rangle}{\partial t} &= \sum_k A_{ik}(\bar{\nu}) \langle \xi_k \xi_j \rangle + \sum_k A_{jk}(\bar{\nu}) \langle \xi_k \xi_i \rangle + B_i(\bar{\nu}) \delta_{i,j} \\
 &\quad + \frac{N^{-1/2}}{2} \sum_{k,l} [A_{ikl}(\bar{\nu}) \langle \xi_j \xi_k \xi_l \rangle + A_{jkl}(\bar{\nu}) \langle \xi_i \xi_k \xi_l \rangle] \\
 (3.26) \qquad &\quad + N^{-1/2} \delta_{i,j} \sum_k B_{ik}(\bar{\nu}) \langle \xi_k \rangle + \mathcal{O}(N^{-1}).
 \end{aligned}$$

At first sight it would appear that, since the second moments couple to third moments at $\mathcal{O}(N^{-1/2})$, we need to determine the equation for third (and possibly higher) moments by going to order $\mathcal{O}(N^{-1})$. Note, however, that the first moment already couples to the second moment at order $\mathcal{O}(N^{-1/2})$; see (3.24). Therefore, if we are only interested in the lowest order corrections to the rate equation (3.8) that couple the first and second moments together, we can drop the $\mathcal{O}(N^{-1/2})$ terms from the second moment equation. Thus,

$$\begin{aligned}
 \tau_s \frac{\partial \langle \xi_i \xi_j \rangle}{\partial t} &= B_i(\bar{\nu}) \delta_{i,j} + \sum_k A_{ik}(\bar{\nu}) \langle \xi_k \xi_j \rangle + \sum_k A_{jk}(\bar{\nu}) \langle \xi_k \xi_i \rangle + \mathcal{O}(N^{-1/2}) \\
 &= [\bar{\nu}_i G_i(\bar{\nu}) + H_i(\bar{\nu})] \delta_{i,j} + \sum_k \left[\frac{\partial F_i(\bar{\nu})}{\partial \bar{\nu}_k} \langle \xi_k \xi_j \rangle + \frac{\partial F_j(\bar{\nu})}{\partial \bar{\nu}_k} \langle \xi_k \xi_i \rangle \right] \\
 (3.27) \qquad &\quad + \mathcal{O}(N^{-1/2}).
 \end{aligned}$$

The various coefficients in the moment equations (3.24) and (3.27) depend on the time-dependent solution $\bar{\nu}(t)$ of the deterministic rate equation (3.8). It is useful to combine (3.8) and (3.24) by setting

$$(3.28) \qquad \nu_i = \bar{\nu}_i + N^{-1/2} \langle \xi_i \rangle$$

such that

$$\begin{aligned}
 (3.29) \quad \tau_s \frac{\partial \nu_i}{\partial t} &= \tau_s \frac{\partial \nu_i}{\partial t} + N^{-1/2} \tau_s \frac{\partial \langle \xi_i \rangle}{\partial t} \\
 &= F_i(\boldsymbol{\nu}) + \frac{1}{2N} \sum_{k,l} \frac{\partial^2 F_i(\boldsymbol{\nu})}{\partial \nu_k \partial \nu_l} C_{kl} + \mathcal{O}(N^{-2}),
 \end{aligned}$$

where C_{kl} is the covariance matrix (3.17). The corresponding initial conditions are $\nu_i(0) = \nu_i^{(0)}$ and $C_{kl}(0) = 0$ (deterministic initial conditions). The covariance matrix \mathbf{C} satisfies the same equation to lowest order as the second moments,

$$\begin{aligned}
 (3.30) \quad \tau_s \frac{\partial C_{ij}}{\partial t} &= [\nu_i G_i(\boldsymbol{\nu}) + H_i(\boldsymbol{\nu})] \delta_{i,j} + \sum_k \left[\frac{\partial F_i(\boldsymbol{\nu})}{\partial \nu_k} C_{kj} + \frac{\partial F_j(\boldsymbol{\nu})}{\partial \nu_k} C_{ki} \right] \\
 &+ \mathcal{O}(N^{-1}),
 \end{aligned}$$

which recovers the $\mathcal{O}(1)$ result (3.19). (3.29) and (3.30) generate the lowest order correction to the rate equation (3.8) that couples the mean activity to the covariance matrix. In the thermodynamic limit $N \rightarrow \infty$ this coupling vanishes, and we recover the deterministic rate equation. An important observation is that the right-hand side of (3.30) depends on the sum $B_i(\boldsymbol{\nu}) = \nu_i G_i(\boldsymbol{\nu}) + H_i(\boldsymbol{\nu})$. It follows from (2.18) and (2.15) that the covariance depends on the arbitrary function $\Gamma_i(\boldsymbol{\nu})$; the latter cancels in the case of $F_i(\boldsymbol{\nu})$. The resulting moment equations for the activity-based rate model (2.7) can be obtained by substituting (2.15) into (3.29) and (3.30):

$$(3.31) \quad \tau_s \frac{\partial \nu_i}{\partial t} = -\nu_i + f \left(\sum_j W_{ij} \nu_j \right) + \frac{1}{2N} f'' \left(\sum_j W_{ij} \nu_j \right) \sum_{k,l} W_{ik} W_{il} C_{kl} + \mathcal{O}(N^{-2}),$$

and

$$\begin{aligned}
 (3.32) \quad \tau_s \frac{\partial C_{ij}}{\partial t} &= \left[\nu_i + f \left(\sum_p W_{ip} \nu_p \right) \right] \delta_{i,j} - 2C_{ij} \\
 &+ \sum_k \left[f' \left(\sum_p W_{ip} \nu_p \right) W_{ik} C_{kj} + f' \left(\sum_p W_{jp} \nu_j \right) W_{jk} C_{ki} \right] \\
 &+ \mathcal{O}(N^{-1}),
 \end{aligned}$$

where we have set $\Gamma_i = 0$ for simplicity.

3.3. Continuum limit of moment equations. Following section 2, we can obtain a continuum limit of the moment equations (3.31) and (3.32) of the activity-based model by considering a continuum of populations (cortical columns) and performing local spatial averaging. This preserves the system size N of each homogeneous population so that one can still apply the system-size expansion. Introducing the spatially averaged activity variable $\nu(m\Delta x, t)$ according to (2.8), we spatially average (3.31) over an infinitesimal region to give ($\Gamma_i \equiv 0$)

$$\begin{aligned}
 \tau_s \frac{\partial \nu(m\Delta x, t)}{\partial t} &= -\nu(m\Delta x, t) + f(u(m\Delta x, t)) + \frac{N^{-1}}{2} f''(u(m\Delta x, t)) \\
 &\times [\rho\Delta x]^2 \sum_{k,l} \omega(m\Delta x, k\Delta x) \omega(m\Delta x, l\Delta x) C(k\Delta x, l\Delta x, t),
 \end{aligned}$$

where

$$(3.33) \quad u(m\Delta x, t) = \rho\Delta x \sum_n \omega(m\Delta x, n\Delta x)\nu(n\Delta x, t)$$

and

$$(3.34) \quad C(k\Delta x, l\Delta x, t) = \frac{1}{[\rho\Delta x]^2} \sum_{i \in \mathcal{N}(k\Delta x)} \sum_{j \in \mathcal{N}(l\Delta x)} C_{ij}(t).$$

Similarly, spatially averaging equation (3.32)

$$\begin{aligned} \tau_s \frac{\partial C(k\Delta x, l\Delta x, t)}{\partial t} &= [\nu(k\Delta x) + f(u(k\Delta x, t))] \frac{\delta_{k,l}}{\rho\Delta x} - 2C(k\Delta x, l\Delta x, t) \\ &\quad + f'(u(k\Delta x, t)) \rho\Delta x \sum_p \omega(k\Delta x, p\Delta x) C(p\Delta x, l\Delta x, t) \\ &\quad + f'(u(l\Delta x, t)) \rho\Delta x \sum_p \omega(l\Delta x, p\Delta x) C(p\Delta x, k\Delta x, t). \end{aligned}$$

Finally, taking the continuum limit leads to the lowest order corrections to the Wilson–Cowan equations:

$$(3.35) \quad \begin{aligned} \tau_s \frac{\partial \nu(x, t)}{\partial t} &= -\nu(x, t) + f(u(x, t)) \\ &\quad + \frac{1}{2N} f''(u(x, t)) \rho^2 \iint \omega(x, y)\omega(x, y') C(y, y', t) dy dy', \end{aligned}$$

and

$$(3.36) \quad \begin{aligned} \tau_s \frac{\partial C(x, x', t)}{\partial t} &= [\nu(x) + f(u(x, t))] \frac{\delta(x-x')}{\rho} - 2C(x, x', t) \\ &\quad + f'(u(x, t)) \rho \int \omega(x, z) C(z, x', t) dz \\ &\quad + f'(u(x', t)) \rho \int \omega(x', z) C(z, x, t) dz \end{aligned}$$

with

$$(3.37) \quad u(x, t) = \rho \int \omega(x, y)\nu(y, t) dy.$$

(3.35) and (3.36) provide the minimal extension of the Wilson–Cowan neural field equations that take into account the coupling between mean activity and the second-order cumulant. The extended neural field equations can be used to study how intrinsic noise due to the finite size of local populations affects spatially structured solutions of neural field equations such as traveling pulses and fronts, spiral waves and bumps. However, as we discussed in section 3.1, they are only valid provided that the network does not operate close to criticality. Moreover, they do not take into account large deviations (rare events) of the full stochastic dynamics, which can induce escape from the basin of attraction of a stable solution of the neural field equations (see section 5). (3.35) and (3.36) are almost identical to the corresponding neural field equations originally derived by Buice and Cowan [12] using path integral methods,

except that in our case the mean activity $\nu(x)$ couples to the covariance $C(x, x')$ rather than the so-called normal-ordered covariance $\widehat{C}(x, x') = C(x, x') - \rho\nu(x)\delta(x - x')$; the latter is obtained by taking the continuum limit of $\widehat{C}_{ij} = C_{ij} - \nu_i\delta_{i,j}$. The significance of the normal-ordered covariance in the neural field equations of Buice and Cowan [12] was recently made explicit by Buice, Cowan, and Chow [13], who used a more direct approach to deriving these equations based on closure of the moment hierarchy. Indeed, one way to interpret the difference between our neural field equations (3.35) and (3.36) and those of Buice and Cowan [12] is that they involve different moment closure conditions (see section 4).

4. Path integrals and the $1/N$ expansion. In this section we show how the system-size expansion of the master equation (2.14) is equivalent to a $1/N$ (loop) expansion of a corresponding path integral representation of the master equation. The latter method was recently introduced by Buice and Cowan [12] and further refined by Buice, Cowan, and Chow [13], based on a modification of previous work on reaction-diffusion systems [19, 20, 60]. It should be noted that, as in the case of the Van Kampen system-size expansion, the path integral representation assumes that there is no upper bound on the number of active neurons m_i in each local population, while N is treated as a fixed parameter of the transition rates.

4.1. Path integral representation of the master equation. The technical details regarding the derivation of the path integral representation are presented in Appendix A (see also [12]). Here we follow a more heuristic approach along the lines of Buice, Cowan, and Chow [13]. Let us first consider the representation of the joint probability density for the fields $\Phi_i = \{\Phi_i(s), 0 \leq s \leq t\}$, with $\Phi_i = N\nu_i$ and ν_i satisfying the deterministic rate equation (3.8). Assuming an initial condition $\Phi_i(0) = \Phi_i^{(0)}$, the joint probability density can be written formally as an infinite product of Dirac delta functions that enforce the solution of the rate equation at each point in time:

$$(4.1) \quad P[\Phi] = \mathcal{N} \prod_{s \leq t} \prod_i \delta \left(\partial_t \Phi_i + \alpha \Phi_i G_i(\Phi/N) - \alpha N H_i(\Phi/N) - \delta(t) \Phi_i^{(0)} \right),$$

where \mathcal{N} is a normalization factor. Introducing the standard Fourier representation of the Dirac delta function, we can rewrite this equation as

$$(4.2) \quad P[\Phi] = \int \prod_i D\tilde{\Phi}_i e^{-S[\Phi, \tilde{\Phi}]},$$

where we have absorbed the normalization factor into the integration measure

$$D\tilde{\Phi}_i \sim \prod_{s \leq t} d\tilde{\Phi}_i(s),$$

each $\tilde{\Phi}_i(s)$ is integrated along the imaginary axis, and S is the so-called action

$$S[\Phi, \tilde{\Phi}] = \int dt \sum_i \tilde{\Phi}_i(t) [\partial_t \Phi_i + \alpha \Phi_i G_i(\Phi/N) - \alpha N H_i(\Phi/N)] - \sum_i \tilde{\Phi}_i(0) \Phi_i^{(0)}.$$

Following along similar lines to Buice and Cowan [12], it can be shown that when fluctuations described by the master equation (2.14) are taken into account, the probability density $P[\Phi]$ is still given by (4.2) except that the action S now takes the form

(see Appendix A)

$$(4.3) \quad S[\Phi, \tilde{\Phi}] = \int dt \sum_i \tilde{\Phi}_i \left[\partial_t \Phi_i + \alpha \hat{G}_i(\Psi/N) \Phi_i - \alpha N \hat{H}_i(\Psi/N) \right] - \sum_i \tilde{\Phi}_i(0) \Phi_i^{(0)},$$

where $\Psi_j = \tilde{\Phi}_j \Phi_j + \Phi_j$ and, for simplicity, the initial distribution is given by a product of independent Poisson processes with means $\bar{n}_i = \Phi_i^{(0)}$,

$$(4.4) \quad P(\mathbf{n}, 0) = \frac{[\bar{n}_i]^{n_i} e^{-\bar{n}_i}}{n_i!}.$$

Note that the functions \hat{G} and \hat{H} are obtained from the functions G and H after “normal-ordering” the action. That is, after moving all fields $\tilde{\Phi}$ to the right of all fields Φ using repeated application of the commutation rule

$$(4.5) \quad \Phi_i \tilde{\Phi}_j = \tilde{\Phi}_j \Phi_i + \delta_{i,j}.$$

Thus, if $g(\Psi) = (\tilde{\Phi}_i + 1) \Phi_i (\tilde{\Phi}_j + 1) \Phi_j$, then

$$\hat{g}(\Psi) = (\tilde{\Phi}_i + 1) (\tilde{\Phi}_j + 1) \Phi_i \Phi_j + (\tilde{\Phi}_i + 1) \Phi_i \delta_{i,j}.$$

(Of course, $\Phi_i(t)$ and $\tilde{\Phi}_i(t)$ are just numbers rather than operators: the commutation rule is simply a prescription for how to construct the functions \hat{G} and \hat{H} ; see Appendix A.)

Given the probability distribution $P[\Phi]$, we can calculate mean fields according to²

$$(4.6) \quad \begin{aligned} \langle\langle \Phi_k(t_1) \rangle\rangle &= \int \prod_i D\Phi_i \Phi_k(t_1) P[\Phi] \\ &= \int \prod_i D\Phi_i \int \prod_i D\tilde{\Phi}_i \Phi_k(t_1) e^{-S[\Phi, \tilde{\Phi}]} \end{aligned}$$

Similarly two-point correlations are given by

$$(4.7) \quad \langle\langle \Phi_k(t_1) \Phi_l(t_2) \rangle\rangle = \int \prod_i D\Phi_i \int \prod_i D\tilde{\Phi}_i \Phi_k(t_1) \Phi_l(t_2) e^{-S[\Phi, \tilde{\Phi}]}.$$

In terms of the statistics of the physical activity variables $m_i(t)$ one finds that

$$(4.8) \quad \langle m_k(t) \rangle \equiv \sum_{\mathbf{n}} n_k P(\mathbf{n}, t) = \langle\langle \Phi_k(t) \rangle\rangle,$$

whereas the covariance is given by [12, 13]

$$(4.9) \quad \langle m_k(t) m_l(t) \rangle - \langle m_k(t) \rangle \langle m_l(t) \rangle = \langle\langle \Phi_k(t) \Phi_l(t) \rangle\rangle - \langle\langle \Phi_k(t) \rangle\rangle \langle\langle \Phi_l(t) \rangle\rangle + \langle\langle \Phi_k(t) \rangle\rangle \delta_{k,l}.$$

Another important characterization of the system is how the mean activity responds to small external inputs. Suppose that we add a small external source

²We use double angular brackets to distinguish between expectation values of fields within the path integral formalism and expectation values of random variables with respect to some probability distribution.

term $h_i(t)$ onto the right-hand side of the deterministic rate equation (3.8). Linearizing about the time-dependent solution of the unperturbed equation ($\mathbf{h} \equiv 0$) leads to the following (nonautonomous) linear equation for the perturbed solution $u_i(t) = \Phi_i^h(t) - \Phi_i(t)$ where $\Phi_i(t) = N\nu_i(t)$:

$$(4.10) \quad \tau_s \frac{du_i}{dt} = \sum_k \frac{\partial F_i(\Phi/N)}{\partial \Phi_k} u_k(t) + Nh_i(t).$$

Introducing the Green’s function or propagator $\mathcal{G}_{ij}^0(t, t')$ according to the adjoint equation

$$(4.11) \quad -\tau_s \frac{d\mathcal{G}_{ij}^0(t, t')}{dt'} = \sum_k \frac{\partial F_k(\Phi/N)}{\partial \Phi_j} \mathcal{G}_{ik}^0(t, t') + \delta_{i,j} \delta(t - t'),$$

we can express the linear response as

$$(4.12) \quad u_i(t) = N \int^t \sum_j \mathcal{G}_{ij}^0(t, t') h_j(t') dt'.$$

In other words, in terms of functional derivatives

$$(4.13) \quad \frac{\delta \Phi_i(t)}{N \delta h_j(t')} = \mathcal{G}_{ij}^0(t, t').$$

Now suppose that we add a source term in the path integral representation. This corresponds to adding a term $\int \sum_i h_i(t) \tilde{\Phi}_i(t) dt$ to the action (4.3). It follows that the associated Green’s function for the full stochastic model is given by

$$(4.14) \quad \mathcal{G}_{ij}(t, t') \equiv \frac{\delta \langle \Phi_i(t) \rangle}{N \delta h_j(t')} = \langle \langle \Phi_i(t) \tilde{\Phi}_j(t') \rangle \rangle.$$

The above analysis motivates the introduction of the generating functional

$$(4.15) \quad Z[\mathbf{J}, \tilde{\mathbf{J}}] = \int \prod_i D\Phi_i \int \prod_i D\tilde{\Phi}_i e^{-S[\Phi, \tilde{\Phi}]} e^{\int dt \sum_j [\tilde{\Phi}_j(t) J_j(t) + \tilde{J}_j(t) \Phi_j(t)]}.$$

Various moments of physical interest can then be obtained by taking functional derivatives with respect to the “current sources” $\mathbf{J}, \tilde{\mathbf{J}}$. For example,

$$(4.16) \quad \langle \langle \Phi_i(t) \rangle \rangle = \left. \frac{\delta}{\delta \tilde{J}_i(t)} Z[\mathbf{J}, \tilde{\mathbf{J}}] \right|_{\mathbf{J}=\tilde{\mathbf{J}}=0},$$

$$(4.17) \quad \langle \langle \Phi_i(t) \Phi_j(t') \rangle \rangle = \left. \frac{\delta}{\delta \tilde{J}_i(t)} \frac{\delta}{\delta \tilde{J}_j(t')} Z[\mathbf{J}, \tilde{\mathbf{J}}] \right|_{\mathbf{J}=\tilde{\mathbf{J}}=0},$$

$$(4.18) \quad \langle \langle \Phi_i(t) \tilde{\Phi}_j(t') \rangle \rangle = \left. \frac{\delta}{\delta \tilde{J}_i(t)} \frac{\delta}{\delta J_j(t')} Z[\mathbf{J}, \tilde{\mathbf{J}}] \right|_{\mathbf{J}=\tilde{\mathbf{J}}=0}.$$

4.2. Loop (1/N) expansion. Suppose that we perform the rescaling $\Phi_i \rightarrow \phi_i = \Phi_i/N$ so that the generating functional can be rewritten as

$$(4.19) \quad Z[\mathbf{J}, \tilde{\mathbf{J}}] = \int \prod_i D\phi_i \int \prod_i D\tilde{\phi}_i e^{-NS[\phi, \tilde{\phi}]} e^{N \int dt \sum_j [\tilde{\phi}_j(t) J_j(t) + \tilde{J}_j(t) \phi_j(t)]}$$

with the action

$$(4.20) \quad S[\phi, \tilde{\phi}] = \int dt \left[\sum_i \tilde{\phi}_i \partial_t \phi_i + \mathcal{H}(\phi, \tilde{\phi}) \right] - \sum_i \tilde{\phi}_i(0) \phi_i^{(0)},$$

where

$$(4.21) \quad \mathcal{H}(\phi, \tilde{\phi}) = \alpha \sum_i \tilde{\phi}_i \left[\widehat{G}_i(\psi) \phi_i - \widehat{H}_i(\psi) \right]$$

and $\psi_i = \tilde{\phi}_i \phi_i + \phi_i$. For convenience of notation we have set $\tilde{\Psi} = \tilde{\phi}$ and rescaled the current \mathbf{J} . Under this rescaling, the normal ordering of a function g is generated using the commutation rule

$$(4.22) \quad \phi_i \tilde{\phi}_j = \tilde{\phi}_j \phi_i + N^{-1} \delta_{i,j}.$$

(3.17), (4.8), and (4.9) imply that

$$(4.23) \quad \nu_k(t) \equiv N^{-1} \langle m_k(t) \rangle = \langle \langle \phi_k(t) \rangle \rangle$$

and

$$(4.24) \quad \begin{aligned} N^{-1} C_{kl} &= N^{-2} [\langle m_k(t) m_l(t) \rangle - \langle m_k(t) \rangle \langle m_l(t) \rangle] \\ &= \langle \langle \phi_k(t) \phi_l(t) \rangle \rangle - \langle \langle \phi_k(t) \rangle \rangle \langle \langle \phi_l(t) \rangle \rangle + N^{-1} \langle \langle \phi_k(t) \rangle \rangle \delta_{k,l}. \end{aligned}$$

Moreover, (4.14) can be rewritten as

$$(4.25) \quad N^{-1} \mathcal{G}_{ij}(t, t') = \langle \langle \phi_i(t) \tilde{\phi}_j(t') \rangle \rangle.$$

The loop expansion of the path integral (4.19) is essentially a diagrammatic method for carrying out a $1/N$ expansion based on steepest descents or the saddle-point method. In the limit $N \rightarrow \infty$, the path integral is dominated by the “classical” solutions $\mathbf{a}(t)$ and $\tilde{\mathbf{a}}(t)$, which extremize the exponent of the generating functional:

$$(4.26) \quad \left. \frac{\delta S[\phi, \tilde{\phi}]}{\delta \phi_i(t)} \right|_{\tilde{\phi}=\tilde{\mathbf{a}}, \phi=\mathbf{a}} = -\tilde{J}_i(t), \quad \left. \frac{\delta S[\phi, \tilde{\phi}]}{\delta \tilde{\phi}_i(t)} \right|_{\tilde{\phi}=\tilde{\mathbf{a}}, \phi=\mathbf{a}} = -J_i(t).$$

In the case of zero currents $\mathbf{J} = \tilde{\mathbf{J}} = 0$, these equations reduce to

$$(4.27) \quad \frac{\partial a_i}{\partial t} = -\frac{\partial \mathcal{H}(\mathbf{a}, \tilde{\mathbf{a}})}{\partial \tilde{a}_i}, \quad \frac{\partial \tilde{a}_i}{\partial t} = \frac{\partial \mathcal{H}(\mathbf{a}, \tilde{\mathbf{a}})}{\partial a_i}.$$

(4.27) take the form of a Hamiltonian dynamical system in which a_i is a “coordinate” variable, \tilde{a}_i is its “conjugate momentum,” and \mathcal{H} is the Hamiltonian. It immediately follows from the form of \mathcal{H} , (4.21), that one type of classical solution is the mean-field classical solution $\tilde{a}_i(t) \equiv 0$ for all i , which implies that $\mathbf{a}(t)$ satisfies the rate equation (3.8), since to lowest order we can ignore any normal ordering of F_i . Interestingly, there are also non-mean-field classical solutions, $\tilde{\mathbf{a}}(t) \neq 0$, which play an important role in determining large deviations or rare event statistics [22]; see section 5. We now perform a perturbation expansion of the path integral (4.19) about the classical solution by setting

$$\phi = \mathbf{a} + \varphi, \quad \tilde{\phi} = \tilde{\mathbf{a}} + \tilde{\varphi}$$

and expanding to second order in $\varphi, \tilde{\varphi}$:

$$(4.28) \quad Z[\mathbf{J}, \tilde{\mathbf{J}}] \approx e^{-NS[\mathbf{a}, \tilde{\mathbf{a}}]} e^{N \int dt \sum_j [\tilde{a}_j(t) J_j(t) + \tilde{J}_j(t) a_j(t)]} \int \prod_i D\varphi_i \int \prod_i D\tilde{\varphi}_i \\ \times \exp \left[-N \int dt dt' \sum_{i,r} \sum_{j,s} \varphi_i^r(t) \frac{\delta^2 S}{\delta \phi_i^r(t) \delta \phi_j^s(t')} \Big|_{\tilde{\phi}=\tilde{\mathbf{a}}, \phi=\mathbf{a}} \varphi_i^s(t') \right].$$

We have introduced the extra index $r = 1, 2$ such that $\varphi_i^1 = \varphi_i, \varphi_i^2 = \tilde{\varphi}_i$. Evaluating the infinite-dimensional Gaussian integral gives

$$Z[\mathbf{J}, \tilde{\mathbf{J}}] \approx e^{-NS[\mathbf{a}, \tilde{\mathbf{a}}]} e^{N \int dt \sum_j [\tilde{a}_j(t) J_j(t) + \tilde{J}_j(t) a_j(t)]} \text{Det} [\mathcal{D}[\mathbf{a}, \tilde{\mathbf{a}}]]^{-1/2},$$

where $\mathcal{D}[\mathbf{a}, \tilde{\mathbf{a}}]$ is the matrix with components

$$(4.29) \quad \mathcal{D}[\mathbf{a}, \tilde{\mathbf{a}}]^{rs}(i, t; j, t') = \frac{\delta^2 S}{\delta \phi_i^r(t) \delta \phi_j^s(t')} \Big|_{\tilde{\phi}=\tilde{\mathbf{a}}, \phi=\mathbf{a}}.$$

Using the following identity for a matrix \mathbf{M} :

$$\text{Det} \mathbf{M} = e^{\text{Tr} \log \mathbf{M}},$$

we obtain the $\mathcal{O}(N^{-1})$ approximation

$$(4.30) \quad Z[\mathbf{J}, \tilde{\mathbf{J}}] \approx e^{-NS_{\text{eff}}[\mathbf{a}, \tilde{\mathbf{a}}]} e^{N \int dt \sum_j [\tilde{a}_j(t) J_j(t) + \tilde{J}_j(t) a_j(t)],}$$

where

$$(4.31) \quad S_{\text{eff}}[\mathbf{a}, \tilde{\mathbf{a}}] = S[\mathbf{a}, \tilde{\mathbf{a}}] + \frac{1}{2} N^{-1} \text{Tr} \log [\mathcal{D}[\mathbf{a}, \tilde{\mathbf{a}}]].$$

In order to use the above expansion to determine corrections to the dynamical mean-field equation (3.8), it is first necessary to introduce a little more formalism. Given the mean fields $\nu_k = \langle\langle \phi_k \rangle\rangle$ and $\tilde{\nu}_k = \langle\langle \tilde{\phi}_k \rangle\rangle$, consider the Legendre transformation

$$(4.32) \quad \Gamma[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}] = W[\mathbf{J}, \tilde{\mathbf{J}}] + \int dt \sum_j [\tilde{\nu}_j(t) J_j(t) + \tilde{J}_j(t) \nu_j(t)],$$

where $W[\mathbf{J}, \tilde{\mathbf{J}}] = -N^{-1} \log Z[\mathbf{J}, \tilde{\mathbf{J}}]$ and Γ is known as the effective action. Since

$$(4.33) \quad \nu_k(t) = -\frac{\delta W}{\delta \tilde{J}_k(t)}, \quad \tilde{\nu}_k(t) = -\frac{\delta W}{\delta J_k(t)},$$

it follows from functionally differentiating equation (4.32) that

$$(4.34) \quad J_k(t) = \frac{\delta \Gamma}{\delta \tilde{\nu}_k(t)}, \quad \tilde{J}_k(t) = \frac{\delta \Gamma}{\delta \nu_k(t)}.$$

Dynamical equations for the physical mean fields $\nu_k(t)$ are then generated by setting $\mathbf{J} = 0 = \tilde{\mathbf{J}}$ in (4.34). Another useful result is obtained by functionally differentiating equations (4.33) with respect to the mean fields $\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}$:

$$(4.35) \quad \delta(t-t') \delta_{r,s} \delta_{k,l} = \frac{\delta \nu_k^r(t)}{\delta \nu_l^s(t')} = -\frac{\delta^2 W}{\delta J_k^r(t) \delta \nu_l^s(t')} = -\sum_{u=1,2} \sum_{p=1}^M \int \frac{\delta^2 W}{\delta J_k^r(t) \delta J_p^u(y)} \frac{\delta J_p^u(y)}{\delta \nu_l^s(t')} dy,$$

where $\nu_i^1 = \nu_i, \nu_i^2 = \tilde{\nu}_i$, etc. Differentiating (4.34) with respect to $\mathbf{J}, \tilde{\mathbf{J}}$ then shows that

$$(4.36) \quad \sum_{u=1,2} \sum_{p=1}^M \int \frac{\delta^2 W}{\delta J_k^r(t) \delta J_p^u(y)} \frac{\delta^2 \Gamma}{\delta \nu_p^u(y) \delta \nu_l^s(t')} dy = -\delta_{r,s} \delta_{k,l} \delta(t-t').$$

In other words, defining the infinite-dimensional matrix $\widehat{\mathcal{D}}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]$ according to

$$(4.37) \quad \widehat{\mathcal{D}}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]^{r,s}(i, t; j, t') = \frac{\delta^2 \Gamma}{\delta \nu_i^r(t) \delta \nu_j^s(t')},$$

we see that $\widehat{\mathcal{D}}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]$ is the inverse of the two-point covariance matrix with components

$$(4.38) \quad -\frac{\delta^2 W}{\delta J_i^r(t) \delta J_j^s(t')} = N [\langle \langle \phi_i^r(t) \phi_j^s(t') \rangle \rangle - \langle \langle \phi_i^r(t) \rangle \rangle \langle \langle \phi_j^s(t') \rangle \rangle].$$

Returning to the expansion about the classical solution, we make the observation that the classical solution $(\mathbf{a}, \tilde{\mathbf{a}})$ and mean-field solution $(\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}})$ differ by terms of $\mathcal{O}(N^{-2})$ due to the fact that the classical solution is an extremum of the action S . Hence, keeping only terms up to $\mathcal{O}(N^{-1})$ we can equate the two solutions. It follows from (4.30) and (4.32) that $\Gamma[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}] = S_{\text{eff}}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}] + \mathcal{O}(N^{-2})$. Moreover, (4.29) and (4.37) imply that $\widehat{\mathcal{D}}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}] = \mathcal{D}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}] + \mathcal{O}(N^{-1})$; that is, we can take $\mathcal{D}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]$ to be the inverse of the two-point covariance matrix. The first-order correction to the mean-field equation is then obtained from (4.34) after setting $\mathbf{J} = \tilde{\mathbf{J}} = \tilde{\boldsymbol{\nu}} = 0$:

$$(4.39) \quad 0 = \left. \frac{\delta \Gamma[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]}{\delta \tilde{\nu}_i(t)} \right|_{\tilde{\boldsymbol{\nu}}=0} = \left. \frac{\delta S[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]}{\delta \tilde{\nu}_i(t)} \right|_{\tilde{\boldsymbol{\nu}}=0} + \frac{1}{2} N^{-1} \text{Tr} \mathcal{D}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]^{-1} \left. \frac{\delta \mathcal{D}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]}{\delta \tilde{\nu}_i(t)} \right|_{\tilde{\boldsymbol{\nu}}=0}.$$

Functionally differentiating $\mathcal{D}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]$ with respect to $\tilde{\nu}_i(t)$ forces $[\mathcal{D}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]^{-1}]^{r,s}(i, t; j, t')$ to be evaluated at the equal times $t = t'$. Since the only nonvanishing equal-time two-point correlation functions exist for $r = s = 1$ [13], it follows that

$$(4.40) \quad N^{-1} \text{Tr} \mathcal{D}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]^{-1} \left. \frac{\delta \mathcal{D}[\boldsymbol{\nu}, \tilde{\boldsymbol{\nu}}]}{\delta \tilde{\nu}_i(t)} \right|_{\tilde{\boldsymbol{\nu}}=0} = - \sum_{k,l} [\langle \langle \phi_k(t) \phi_l(t) \rangle \rangle - \langle \langle \phi_k(t) \rangle \rangle \langle \langle \phi_l(t) \rangle \rangle] \frac{\partial^2 \widehat{F}_i(\boldsymbol{\nu})}{\partial \nu_k \partial \nu_l},$$

where $\widehat{F}_i(\boldsymbol{\nu}) = -\widehat{G}_i(\boldsymbol{\nu})\nu_i + \widehat{H}_i(\boldsymbol{\nu})$. Adding this to the $\mathcal{O}(1)$ rate equation (3.8) and using (4.24), we have

$$(4.41) \quad \tau_s \frac{\partial \nu_i}{\partial t} = \widehat{F}_i(\boldsymbol{\nu}) + \frac{1}{2N} \sum_{k,l} \frac{\partial^2 \widehat{F}_i(\boldsymbol{\nu})}{\partial \nu_k \partial \nu_l} \widehat{C}_{kl} + \mathcal{O}(N^{-2}),$$

where

$$(4.42) \quad \widehat{C}_{kl} = C_{kl} - \nu_k \delta_{k,l} = N^{-1} (\langle m_k m_l \rangle - \langle m_k \rangle \langle m_l \rangle - \langle m_k \rangle \delta_{k,l})$$

is the normal-ordered covariance [13]. Finally, using the fact that

$$\widehat{F}_i(\boldsymbol{\nu}) = F_i(\boldsymbol{\nu}) + \frac{1}{2N} \sum_k \frac{\partial^2 F_i(\boldsymbol{\nu})}{\partial \nu_k^2} \nu_k + \mathcal{O}(N^{-2}),$$

we have

$$(4.43) \quad \tau_s \frac{\partial \nu_i}{\partial t} = F_i(\boldsymbol{\nu}) + \frac{1}{2N} \sum_{k,l} \frac{\partial^2 F_i(\boldsymbol{\nu})}{\partial \nu_k \partial \nu_l} C_{kl} + \mathcal{O}(N^{-2}),$$

which is identical to the first-order moment equation (3.29) derived using the Van Kampen system-size expansion. It is also possible to derive a corresponding dynamical equation for \mathbf{C} by extending the definition of the effective action along the lines of Buice, Cowan, and Chow [13], and this is consistent with (3.30). We have thus established that the Van Kampen system-size expansion and the loop expansion of the corresponding path integral representation yield the same lowest order corrections to mean-field theory; we expect this agreement to hold to arbitrary orders in N^{-1} .

The above path integral loop expansion differs in one crucial respect from the corresponding expansion of the activity-based rate model carried out by Buice and Cowan [12] and Buice, Cowan, and Chow [13]. That is, these authors interpreted a_i as the mean number rather than the fraction of active neurons in a local population and took the thermodynamic limit $N \rightarrow \infty$ ab initio with $NH_i \rightarrow H_i$ in the action (4.3) and G_i, H_i given by (2.15). Under this particular choice of scaling there does not exist a natural small parameter N^{-1} with which to carry out a system-size or loop expansion. Nevertheless, one can still formally carry out the one loop expansion of the path integral to obtain (4.41) without the factor N^{-1} . However, some other moment closure condition is now needed in order to justify truncating the loop expansion away from critical points. As shown by Buice, Cowan, and Chow [13], such a truncation can be carried out provided that both the nonlinear gain function f and the cumulants are kept in normal-ordered form. Substituting (2.15) into (4.41) and taking the continuum limit then leads to the neural field equation (3.35), with f and C replaced by their normal-ordered forms and the factor N^{-1} removed. This form of moment closure assumes that the population activity is approximately Poissonian [13].

4.3. Continuum limit of the action. In the case of the activity-based model, it is possible to take the continuum limit of the path integral prior to carrying out the loop expansion. Substituting for G_i and H_i using (2.15) and setting $\Gamma_i = 0$, the action (4.20) takes the particular form

$$(4.44) \quad S[\phi, \tilde{\phi}] = \int dt \sum_i \tilde{\phi}_i \left[\partial_t \phi_i + \alpha \phi_i - \alpha \hat{f} \left(\sum_j W_{ij} [\tilde{\phi}_j \phi_j + \phi_j] \right) \right] - \sum_i \tilde{\phi}_i(0) \phi_i^{(0)},$$

where \hat{f} is the normal-ordered gain function. In order to determine the continuum limit of the above action, we perform spatial averaging along the lines of sections 2 and 3, that is, grouping populations together over an interval of length Δx . Suppose that there is a uniform density ρ of neural populations distributed along the x axis. As before, we partition the x axis into discrete intervals of length Δx within which there are $\rho \Delta x$ populations and denote the set of populations in the interval $[n\Delta x, (n+1)\Delta x)$

by $\mathcal{N}(n\Delta x)$. The action (4.44) can then be rewritten as

$$(4.45) \quad \begin{aligned} S[\phi, \tilde{\phi}] = & \int dt \rho \Delta x \sum_n \tilde{\phi}(n\Delta x, t) \left[\partial_t \phi(n\Delta x, t) + \alpha \phi(n\Delta x, t) \right. \\ & \left. - \alpha \hat{f} \left(\rho \Delta x \sum_m \omega(n\Delta x, m\Delta x) [\tilde{\phi}(m\Delta x, t) \phi(m\Delta x, t) + \phi(m\Delta x, t)] \right) \right] \\ & - \rho \Delta x \sum_n \tilde{\phi}(n\Delta x, 0) \phi^{(0)}(n\Delta x). \end{aligned}$$

We have assumed that the auxiliary field $\tilde{\phi}$ is a slowly varying function of space, which is consistent with the classical solution $\tilde{\phi}(n\Delta x, t) = 0$. If we now consider a generating functional that generates moments of the spatially averaged fields, we obtain the following continuum path integral in space and time:

$$(4.46) \quad Z[\mathbf{J}, \tilde{\mathbf{J}}] = \int D\phi D\tilde{\phi}_i e^{-NS[\phi, \tilde{\phi}]} e^{N \int dt dx \tilde{\phi}(x, t) J(x, t) + \tilde{\mathbf{J}}(x, t) \phi(x, t)}$$

with

$$(4.47) \quad \begin{aligned} S[\phi, \tilde{\phi}] = & \rho \int dt dx \tilde{\phi}(x, t) \left[\partial_t \phi(x, t) + \alpha \phi(x, t) \right. \\ & \left. - \hat{f} \left(\rho \int dx' \omega(x, x') [\tilde{\phi}(x', t) \phi(x', t) + \phi(x', t)] \right) \right] \\ & - \rho \int dx \tilde{\phi}(x, 0) \phi^{(0)}(x). \end{aligned}$$

Applying steepest descents to the continuum path integral generates the continuum moment equations (3.35) and (3.36), after expanding the normal-ordered gain function f in powers of N^{-1} .

5. Hamiltonian dynamics and rare event statistics. Although it is interesting from a mathematical perspective to establish the relationship between the loop expansion of the path integral and Van Kampen's system-size expansion, one could argue that the latter approach is much simpler. This then raises an issue regarding the usefulness of the path integral formalism within the context of stochastic neural field theory. Buice and Cowan [12] have already demonstrated one important application, namely, that renormalization group methods can be used to study power law behavior in the vicinity of critical points. Moreover, as we have clarified in this paper, the path integral approach provides alternative perturbation schemes in the absence of a natural system-size parameter, for example, normal ordering the gain function and cumulants [13]. In this section we illustrate a further application of path integral methods, namely, to the study of large deviations or rare event statistics. More specifically, Elgart and Kamenev [22] have shown that in the case of reaction-diffusion systems, the Hamiltonian-like dynamical system obtained by extremizing the associated path integral action (cf. (4.27)) can be used to determine the most probable or optimal path that leads to escape from the basin of attraction of a stable fixed point of the mean-field dynamics. This provides an estimate of the lifetime of a metastable state. We will show how the same ideas can be used to analyze metastability in the activity-based neural master equation (2.13). For the sake of illustration we consider a single homogeneous population ($M = 1$).

Let us begin by considering fixed point solutions of the activity-based rate equation (2.7) for a single population:

$$(5.1) \quad \tau_s \frac{d\nu}{dt} = -\nu + f(W\nu).$$

The fixed points and their stability can be determined graphically by plotting the intercepts of the function $f(\nu)$ with the linear function ν/W . This is illustrated in Figure 5.1(a) for $f(\nu) = \tanh(\nu)$, $\nu \geq 0$. If $W < 1$, then there is a single stable fixed point at $\nu = 0$. The fixed point becomes critical at $W = 1$ such that for $W > 1$ it is unstable and there is an additional stable fixed point at $\nu = \nu_c > 0$. In Figure 5.1(b) we show corresponding results for the gain function $f(\nu) = e^{-r/(\nu-\kappa)^2} \Theta(\nu-\kappa)$, where Θ is the Heaviside function, κ is a positive threshold, and r^{-1} is a gain parameter. The exponential factor ensures that the function is continuously differentiable at $\nu = \kappa$ for $r > 0$ [51]; if $r = 0$, then we recover the standard Heaviside gain function. Note that this gain function is similar in form to the gain function derived from studies of asynchronous states in noisy spiking networks [10, 9, 35, 66]. Now the zero fixed point is always stable for $\kappa \geq 0$, whereas an additional stable/unstable pair of fixed points arises via a saddle-node bifurcation at the critical point $W = W_c$. Thus, the deterministic network is bistable when $W > W_c$. A similar scenario would occur on varying the threshold κ . How does the basic picture of mean-field theory change when fluctuations are taken into account? The linear noise approximation (see section 3.1) suggests that, depending on initial conditions, the dynamics is characterized by Gaussian fluctuations about one of the stable fixed points. The mean and variance evolve according to the equations

$$(5.2) \quad \tau_s \frac{\partial \langle \xi \rangle}{\partial t} = -\langle \xi \rangle + f'(W\nu)W \langle \xi \rangle$$

and

$$(5.3) \quad \tau_s \frac{dC}{dt} = \nu + f(W\nu) + 2[-1 + Wf'(W\nu)]C.$$

Criticality of a fixed point ν^* occurs when $W = W_c$ with $W_c^{-1} = f'(W_c\nu^*)$. One problem with the linear noise approximation away from critical points is that it does not take into account the fact that the number of active neurons $n = N\nu$ is bounded, $0 \leq n \leq N$. However, a more serious problem arises from the observation that if $f(0) = 0$, then the zero activity state $n = 0$ is an absorbing state of the corresponding master equation. This means that $P_n(t) \rightarrow \delta_{n,0}$ in the limit $t \rightarrow \infty$; that is, $\delta_{n,0}$ is a unique stationary probability distribution. It follows that the nonzero stable fixed point appearing in Figures 5.1(a,b) is only metastable.

Let us now consider the generating functional (4.19) for a single population, which takes the form

$$(5.4) \quad Z[J, \tilde{J}] = \int D\phi \int D\tilde{\phi} e^{-NS[\phi, \tilde{\phi}]} e^{N \int dt \tilde{\phi}(t)J(t) + \tilde{J}(t)\phi(t)},$$

where S is the single population model version of the activity-based action (4.44):

$$(5.5) \quad S[\phi, \tilde{\phi}] = \int dt \left[\tilde{\phi} \partial_t \phi + \mathcal{H}(\phi, \tilde{\phi}) \right] - \tilde{\phi}(0)\phi^{(0)}$$

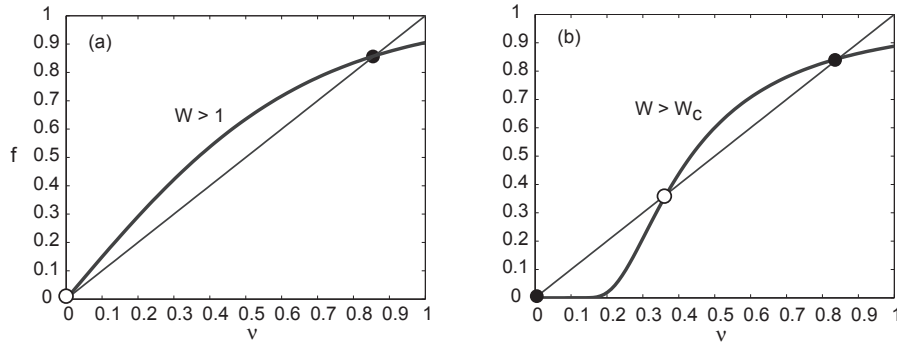


FIG. 5.1. Fixed points of the mean-field equation (5.1) for (a) $f(v) = \tanh(v)$ and (b) $f(v) = e^{-r/(v-\kappa)^2} \Theta(v-\kappa)$ with $r = 1$ and $\kappa = 0.1$.

with

$$(5.6) \quad \mathcal{H}(\phi, \tilde{\phi}) = \alpha \tilde{\phi} \phi - \alpha \tilde{\phi} f(W[\phi \tilde{\phi} + \phi]).$$

Since we will be working to leading order in N^{-1} , we have dropped the normal ordering of the gain function f . It follows from the saddle-point method that the path integral is dominated by the classical solution $\phi = q, \tilde{\phi} = p$, which is obtained by extremizing the action (5.5). (In contrast to section 4, we write the classical solution as (q, p) rather than (a, \tilde{a}) in order to emphasize the connection with Hamiltonian dynamics.) This leads to the Hamiltonian dynamical system

$$(5.7) \quad \frac{dq}{dt} = -\frac{\partial \mathcal{H}}{\partial p} = -\alpha q + \alpha f(Wq(p+1)) + pqWf'(Wq(p+1))$$

and

$$(5.8) \quad \tau_s \frac{dp}{dt} = \frac{\partial \mathcal{H}}{\partial q} = \alpha p (1 - (p+1)Wf'(Wq(p+1))).$$

(5.7) and (5.8) have an integral of motion given by the conserved “energy”

$$(5.9) \quad E = \mathcal{H}(q(t), p(t)) = \alpha p(t) [q(t) - f(Wq(t)(p(t)+1))].$$

The fact that $\mathcal{H}(q, 0) = 0$ is a direct consequence of probability conservation, whereas the fact that $\mathcal{H}(0, p) = 0$ for $f(0) = 0$ reflects the existence of an absorbing state. It follows from energy conservation that the action evaluated along a classical trajectory is given by

$$(5.10) \quad S[q, p] = Et + \int_0^t p \partial_\tau q d\tau - \frac{\bar{n}}{N} p(0).$$

The action is a functional of the classical solutions $p(\tau)$ and $q(\tau)$ with $0 \leq \tau \leq t$.

As shown by Elgart and Kamenev [22] in the case of reaction-diffusion systems, the plot of constant energy trajectories in the phase plane (p, q) corresponding to solutions of the Hamiltonian equations (5.7) and (5.8) provides important insights into the dynamics underlying destabilization of a metastable state due to the presence of an absorbing state. This is illustrated in Figure 5.2 for the gain function $f(q) = \tanh(q)$

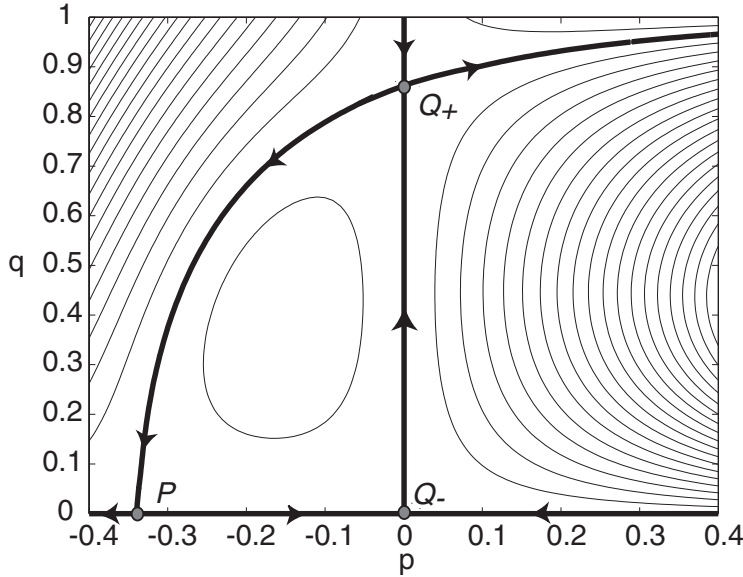


FIG. 5.2. Phase portrait of constant energy trajectories for the Hamiltonian system given by (5.7) and (5.8) with $f(q) = \tanh(q)$ and $W = 1.5$. Zero-energy trajectories are indicated by thick curves. The stable and unstable fixed points of the mean-field dynamics are denoted by Q_+ and Q_- . These are connected to a fluctuational fixed point P via zero-energy heteroclinic connections. The curve Q_+P is the optimal path from the metastable state to the absorbing state.

(cf. Figure 5.1(a)). The zero-energy trajectories are highlighted as thicker curves, since they play a crucial role in determining the long-term dynamics of the network. One zero-energy curve is the line $p = 0$ along which (5.7) reduces to the mean-field equation (3.8). It follows from (5.9) that the energy of the mean-field solution is zero and $S[q, 0] = 0$. If the dynamics were restricted to the one-dimensional manifold $p = 0$, then the nonzero fixed point of the mean-field equation would be stable. However, it becomes a saddle point of the full dynamics in the (p, q) plane, denoted by Q_+ in Figure 5.2, reflecting the fact that it is metastable when fluctuations are taken into account. A second zero-energy curve is the absorbing line $q = 0$ which includes two additional hyperbolic fixed points denoted by Q_- and P in Figure 5.2. Q_- occurs at the intersection with the mean-field line $p = 0$ and corresponds to the unstable fixed point of the mean-field dynamics, whereas P is associated with the effects of fluctuations. Moreover, there exists a third zero-energy curve, which includes a heteroclinic trajectory joining the mean-field fixed point at $t = -\infty$ to the fluctuational fixed point P at $t = \infty$. This heteroclinic trajectory represents the optimal (most probable) path linking the metastable fixed point to the absorbing boundary. It follows from large deviation theory [34, 38, 21, 22] that the mean escape time τ from the metastable state is, with exponential accuracy, $\tau \sim e^{N\mathcal{S}_0}$, where \mathcal{S}_0 is the zero-energy action

$$(5.11) \quad \mathcal{S}_0 = \int_{Q_+}^P pdq$$

with the integral evaluated along the heteroclinic trajectory from Q_+ to P . A similar analysis can be applied to the metastable state shown in Figure 5.1(b) for the gain function $f(q) = e^{-r/(q-\kappa)^2} \Theta(q - \kappa)$. The corresponding phase portrait is shown in Figure 5.3. Here there are three hyperbolic fixed points Q_{\pm}, Q_0 lying on the line

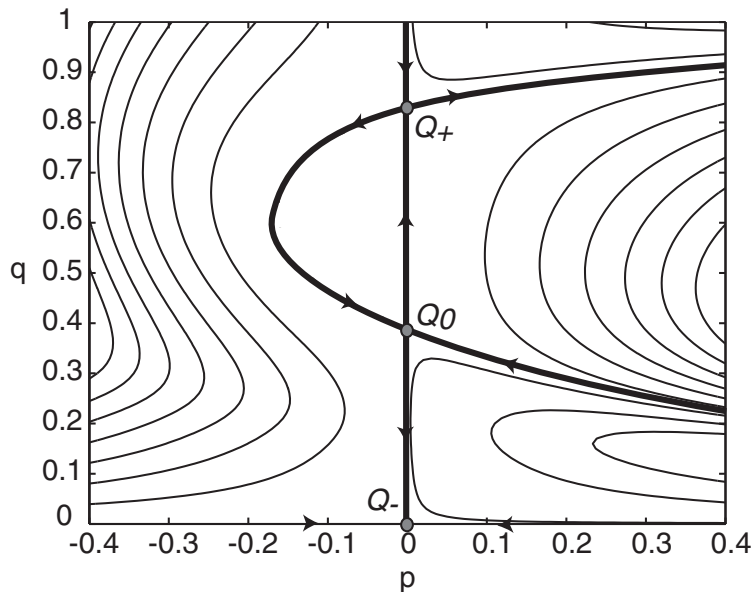


FIG. 5.3. Phase portrait of constant energy trajectories for the Hamiltonian system given by (5.7) and (5.8) with $f(q) = e^{-r/(q-\kappa)^2} \Theta(q-\kappa)$, $r = 0.1$, $\kappa = 0.5$, and $W = 3$. Zero-energy trajectories are indicated by thick curves. The hyperbolic fixed points Q_{\pm} correspond to stable equilibria of the mean-field dynamics, whereas Q_0 corresponds to an unstable equilibrium. The fixed point Q_+ is connected to Q_0 via a zero-energy heteroclinic connection, which represents the optimal path from the metastable state to the boundary of its basin of attraction.

$p = 0$, corresponding to the two stable and one unstable fixed points of the mean-field rate equation. In this case there is a non-mean-field heteroclinic connection Q_+Q_- that is the optimal path for escape from the basin of attraction of the active stable state. Evaluating the action along this path again determines the mean escape time. Note that a more explicit expression for the zero-energy action (5.11) can be obtained by performing the canonical transformation

$$(5.12) \quad \tilde{q} = (p+1)q, \quad \tilde{p} = \log(p+1).$$

The zero-energy curves are then given by solutions of the equation

$$(5.13) \quad \tilde{q}(e^{-\tilde{p}} - 1) + f(W\tilde{q})(e^{\tilde{p}} - 1) = 0.$$

In particular, the non-mean-field solution is

$$(5.14) \quad \tilde{p} = \log \left[\frac{\tilde{q}}{f(W\tilde{q})} \right]$$

such that

$$(5.15) \quad \mathcal{S}_0 = \int_{\tilde{Q}_+}^{\tilde{P}} \log \left[\frac{\tilde{q}}{f(W\tilde{q})} \right] d\tilde{q}.$$

In the case of a single neural population, it is possible to calculate the prefactor in the expression for the lifetime τ of a metastable state by carrying out an asymptotic

expansion of the original master equation (2.13) based on a WKB approximation [41, 18, 28]. However, this asymptotic analysis becomes considerably more difficult in the case of multiple interacting populations or when parameters of the network become time dependent. On the other hand, the above Hamiltonian formulation extends to these more complicated cases. Moreover, the topological structure of the phase portrait for the single population model could provide insights into the nature of phase transitions in corresponding continuum neural field models. This is suggested by recent work of Elgart and Kamenev [23], who developed a classification scheme of phase transitions in reaction-diffusion models based on the phase portraits of the corresponding reaction Hamiltonian. For example, the phase portrait shown in Figure 5.2 would correspond to a phase transition belonging to the directed percolation (DP) universality class [42], whereas the one shown in Figure 5.3 would correspond to a first-order phase transition [23]. Whether or not a similar classification scheme can be applied to nonlocal neural field models remains to be established. Interestingly, Buice and Cowan [12] have used path integral and renormalization group methods to show how the Wilson–Cowan continuum model can exhibit a dynamical phase transition which belongs to the DP universality class. Moreover, the existence of such a phase transition is consistent with scaling laws found in many measurements of cortical activity, both *in vitro* and *in vivo* [4, 64]. Our results suggest that the occurrence of a DP phase transition requires both fine-tuning of the firing threshold and concavity of the gain function close to the absorbing state (see Figures 5.1, 5.2, and 5.3). The necessity of fine-tuning was also previously noted by Buice, Cowan, and Chow [13].

6. Discussion. In this paper we have explored the connection between two approximation methods for analyzing a master equation formulation of stochastic neurodynamics, one based on the Van Kampen system-size expansion and the other on path integrals and steepest descents. Under the assumption that each local population fluctuates about an asynchronous state, we have shown how both methods generate an expansion of the moment equations in powers of N^{-1} , where N is the size of each local population. Our analysis complements previous work on path integral methods in neural field theory [12, 13], which assumed that the mean-field represents the mean number rather than a fraction of active neurons in a population. In the latter case, there does not exist an explicit small parameter to carry out a system-size expansion so that other constraints are needed in order to truncate the moment hierarchy. In particular, moment closure is achieved by assuming that the network population dynamics is Poisson-like and by expressing the moment equations in terms of normal-ordered gain functions and cumulants. Which version of stochastic neural field theory is more appropriate is likely to depend on the particular application and whether population activity is better described by a Poisson process or by Gaussian fluctuations about a mean field. Moreover, as we highlighted in the paper, there is an ambiguity in how to construct the master equation, which directly affects the dynamical equations for second- and higher-order statistics. Such ambiguity reflects the phenomenological nature of the master equation and suggests that additional biophysical constraints are needed. In order to fully characterize the differences between the various models and how these depend on the system size N , it will be necessary to carry out direct Monte Carlo simulations of the corresponding master equations for a variety of network configurations. We will present the results of such simulations elsewhere.

Although the master equation formulation has a number of limitations, it still yields important insights into the general problem of stochastic neural field theory. First, it provides a framework for analyzing possible forms of coupling between mean

activity and higher-order statistics, which is not captured by additive white noise processes. Hence, one obtains extensions to standard neural theory, which can be used to study the effects of noise on spatially organized phenomena such as traveling waves and bumps. Second, it demonstrates that even when fluctuations scale as $N^{-1/2}$, they can be significant in the vicinity of critical points of the underlying deterministic equations. Moreover, the path integral representation of the master equation provides a powerful framework for studying the dynamics close to critical points. For example, the renormalization group can be used to extract the power law behavior of correlations arising from dynamical phase transitions [12], and these can then be compared with experimental studies of spontaneous activity in vivo and in vitro [4, 64]. Finally, as we also showed in this paper, path integral methods can be used to study large deviations or rare event statistics, which determine the optimal path for escape from the basin of attraction of a stable fixed point of the underlying neural field equations away from criticality [21, 22, 29, 28]. These optimal paths are solutions to an effective Hamiltonian dynamical system, which might prove useful in classifying dynamical phase transitions in stochastic neural field models.

Although we have focused on one particular formulation of stochastic neurodynamics that is based on a phenomenological master equation, our analysis raises some issues that relate to other approaches. In particular, investigating corrections to mean-field theory by carrying out a system-size expansion and truncating the moment hierarchy has been applied elsewhere to binary and integrate-and-fire networks [37, 54, 53, 69, 5]. However, as we have highlighted in this paper, any such moment closure method can at best capture only the quasi-steady-state behavior of a large network. That is, when there are multiple fixed points, the truncated moment equations fail to take into account exponentially small transitions between fixed points, which underlie the asymptotically slow approach to the true steady state of the full probabilistic model (assuming that it exists). It would be interesting to explore the issue of rare event or large deviation statistics within the context of population density methods, where failure of moment closure even occurs in the absence of multiple fixed points [52].

Appendix A. In this appendix we describe how to derive the path integral representation of the master equation (2.14), following along similar lines to Buice and Cowan [12] for the Wilson–Cowan equation and Doi and Peliti [19, 20, 60] for reaction-diffusion systems. This will make clear how the issue of normal ordering arises within the path integral framework, which plays a crucial role in differentiating between our choice of scaling and that used by Buice and Cowan [12] and Buice, Cowan, and Chow [13]. The first step is to introduce creation-annihilation operators for each population according to the commutation rule

$$(A.1) \quad [\Phi_i, \Phi_j^\dagger] \equiv \Phi_i \Phi_j^\dagger - \Phi_j^\dagger \Phi_i = \delta_{i,j}.$$

These operators generate a state space by acting on the “vacuum” state $|0\rangle$, which represents the state in which all populations are quiescent ($n_i = 0$ for all i), with $\Phi_i |0\rangle = 0$. The representation of the state with configuration \mathbf{n} is then generated according to

$$(A.2) \quad |\mathbf{n}\rangle = \prod_i \Phi_i^{\dagger n_i} |0\rangle.$$

Inner products in this state space are defined by $\langle 0|0\rangle = 1$ and the commutation relations. It follows that the dual of the vector $\Phi_i^\dagger |0\rangle$ is $\langle 0|\Phi_i$. The number of active

neurons in the i th population can be extracted by operating on a state vector with the number operator $\Phi_i^\dagger \Phi_i$ and using the commutation relations,

$$(A.3) \quad \Phi_i^\dagger \Phi_i |\mathbf{n}\rangle = n_i |\mathbf{n}\rangle.$$

The next step is to construct an operator representation of the master equation (2.14). Given the probability distribution $P(\mathbf{n}, t)$, we introduce the state vector

$$(A.4) \quad |\phi(t)\rangle = \sum_{\mathbf{n}} P(\mathbf{n}, t) \prod_i \Phi_i^{\dagger n_i} |0\rangle = \sum_{\mathbf{n}} P(\mathbf{n}, t) |\mathbf{n}\rangle,$$

where the sum is over all possible configurations \mathbf{n} . Introducing the projection state

$$(A.5) \quad |\mathbf{1}\rangle = \exp\left(\sum_i \Phi_i^\dagger\right) |0\rangle$$

with $\Phi_i |\mathbf{1}\rangle = |\mathbf{1}\rangle$, we can then express expectation values in terms of inner products. For example,

$$(A.6) \quad \langle \mathbf{1} | \Phi_i^\dagger \Phi_i | \phi(t) \rangle = \sum_{\mathbf{n}} n_i P(\mathbf{n}, t) = \langle n_i(t) \rangle.$$

Differentiating the state vector $|\phi(t)\rangle$ with respect to t and using the master equation (2.14) gives

$$(A.7) \quad \begin{aligned} \tau_s \partial_t |\phi(t)\rangle &= \sum_{\mathbf{n}} \partial_t P(\mathbf{n}, t) |\mathbf{n}\rangle \\ &= \sum_i \left[(n_i + 1) G_i(\mathbf{n}_{i+}/N) P(\mathbf{n}_{i+}, t) - n_i G_i(\mathbf{n}/N) P(\mathbf{n}, t) \right. \\ &\quad \left. + N \left(H_i(\mathbf{n}_{i-}/N) P(\mathbf{n}_{i-}, t) - H_i\left(\frac{\mathbf{n}}{N}\right) \right) P(\mathbf{n}, t) \right] |\mathbf{n}\rangle \\ &= \sum_i \left[(\Phi_i - \Phi_i^\dagger \Phi_i) G_i(\Psi/N) + N(\Phi_i^\dagger - 1) H_i(\Psi/N) \right] |\phi(t)\rangle, \end{aligned}$$

where $\Psi_i = \Phi_i^\dagger \Phi_i$. It is convenient to express all expectation values by taking the inner product with respect to the vacuum state $|0\rangle$ rather than the projection state $|\mathbf{1}\rangle$. This can be achieved by moving $\exp(\sum_i \Phi_i^\dagger)$ all the way to the right in expectation values. From the commutation relation, this is equivalent to performing the shift $\Phi_i^\dagger \rightarrow \Phi_i^\dagger + 1$ in all operator expressions. Thus

$$(A.8) \quad \langle n_i(t) \rangle = \langle 0 | \Phi_i^\dagger \Phi_i + \Phi_i | \phi(t) \rangle.$$

Hence, performing such a shift in (A.7) we see that the operator representation of the master equation (2.14) is

$$(A.9) \quad \partial_t |\phi(t)\rangle = -\mathcal{H} |\phi(t)\rangle$$

with

$$(A.10) \quad \mathcal{H} = \alpha \sum_i \left[\Phi_i^\dagger \Phi_i G_i(\Psi/N) - N \Phi_i^\dagger H_i(\Psi/N) \right],$$

$\Psi_i = \Phi_i^\dagger \Phi_i + \Phi_i$, and $\alpha = \tau_s^{-1}$. Note that the operator equation (A.9) can be converted to a PDE describing the evolution of the corresponding generating function $G(\mathbf{z}, t)$ of the probability distribution $P(\mathbf{n}, t)$. This follows from setting

$$G(\mathbf{z}, t) \equiv \langle 0 | e^{\sum_i z_i \Phi_i} | \phi(t) \rangle = \sum_{\mathbf{n}} \left[\prod_i z_i^{n_i} \right] P(\mathbf{n}, t),$$

and converting the operators Φ_i, Φ_i^\dagger to $d/dz_i, z_i$. Both sets of operators satisfy the same commutation relations.

Formally speaking, the solution to the operator version of the master equation (A.9) can be written as

$$(A.11) \quad |\phi(t)\rangle = e^{-\mathcal{H}t} |\phi(0)\rangle = e^{-\mathcal{H}t} \mathcal{I}_0 |0\rangle,$$

where \mathcal{I}_0 is the operator that generates the initial state. The expectation value of some physical quantity such as the number $n_i(t)$ can now be expressed as

$$(A.12) \quad \langle n_i(t) \rangle = \langle 0 | (\Phi_i^\dagger \Phi_i + \Phi_i) e^{-\mathcal{H}t} \mathcal{I}_0 |0\rangle.$$

For simplicity, we will assume that the initial state is given by a Poisson distribution so that

$$(A.13) \quad \mathcal{I}_0 |0\rangle = \prod_i e^{-\bar{n}_i} \sum_k \frac{\bar{n}_i^k}{k!} \Phi_i^{\dagger k} |0\rangle = \exp \left(\sum_i (-\bar{n}_i + \bar{n}_i \Phi_i^\dagger) \right) |0\rangle.$$

After performing the shift $\Phi_i^\dagger \rightarrow \Phi_i^\dagger + 1$, we see that

$$(A.14) \quad \mathcal{I}_0 = \exp \left(\sum_i \bar{n}_i \Phi_i^\dagger \right).$$

In order to convert the operator form of the expectation value into a path integral, we divide the time interval $[0, t]$ into \mathcal{N} intervals of length $\Delta t = t/\mathcal{N}$ and set $t_r = r\Delta t$, $r = 0, 1, \dots, \mathcal{N}$. We then introduce coherent states of the form

$$(A.15) \quad |\varphi(t)\rangle = \exp \left(-\frac{1}{2} \sum_i \tilde{\varphi}_i(t) \varphi_i(t) \right) \exp \left(\sum_i \varphi_i(t) \Phi_i^\dagger \right) |0\rangle,$$

such that $\varphi_i(t)$ is the complex-valued eigenvalue of the annihilation operator Φ_i , with complex conjugate $\tilde{\varphi}_i$. Coherent states satisfy the completeness relation

$$(A.16) \quad \int \prod_i \frac{d\varphi_i d\tilde{\varphi}_i}{\pi^M} |\varphi\rangle \langle \varphi| = 1.$$

At each discrete time step we insert a complete set of coherent states using the completeness relation (A.16) so that the expectation value becomes

$$(A.17) \quad \langle n_i(t) \rangle = \langle 0 | (\Phi_i^\dagger \Phi_i + \Phi_i) |\varphi(t)\rangle \left[\prod_{r=0}^{\mathcal{N}-1} \langle \varphi(t_{r+1}) | (1 - \mathcal{H}\Delta t) | \varphi(t_r) \rangle \right] \langle \varphi(t_0) | \mathcal{I}_0 |0\rangle,$$

where we have made use of the formula

$$(A.18) \quad e^{-\mathcal{H}t} = \lim_{\mathcal{N} \rightarrow \infty} (1 - \mathcal{H}\Delta t)^{\mathcal{N}}.$$

The crucial observation is that at the r th time step we can replace the annihilation and creation operators Φ_i and Φ_i^\dagger in \mathcal{H} by the corresponding coherent state eigenvalues $\tilde{\varphi}_i(t_{r+1})$ and $\varphi_i(t_r)$, provided that we normal order the operator \mathcal{H} . Finally, taking the limits $\mathcal{N} \rightarrow \infty$ and $\Delta t \rightarrow 0$, we obtain the following path integral representation of the expectation value:

$$(A.19) \quad \langle n_i(t) \rangle = \int \prod_j D\varphi_j \int \prod_j D\tilde{\varphi}_j \varphi_i(t) e^{-S[\varphi, \tilde{\varphi}]},$$

where S is given by the action (4.3) after making the change of variable $\varphi \rightarrow \Phi, \tilde{\varphi} \rightarrow \tilde{\Phi}$ with Φ and $\tilde{\Phi}$ now treated as numbers rather than operators. Note that since $\langle 0 | \Phi_i^\dagger = 0$ we have replaced $\Phi_i^\dagger \Phi_i + \Phi_i$ by Φ_i in the first inner product on the right-hand side of (A.17).

Acknowledgements. P.C.B. would like to thank Michael Buice (NIH) for many helpful comments regarding the path integral representation, and for highlighting the importance of normal ordering in matching the loop expansion with the system-size expansion.

REFERENCES

- [1] L. F. ABBOTT AND C. VAN VRESSWIJK, *Asynchronous states in networks of pulse-coupled oscillators*, Phys. Rev. E, 48 (1993), pp. 1483–1490.
- [2] S. AMARI, *Dynamics of pattern formation in lateral inhibition type neural fields*, Biol. Cybern., 27 (1977), pp. 77–87.
- [3] D. J. AMIT AND N. BRUNEL, *Model of global spontaneous activity and local structured activity during delay periods in cerebral cortex*, Cereb. Cortex, 7 (1997), pp. 237–252.
- [4] 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.
- [5] S. EL BOUSTANI AND A. DESTEXHE, *A master equation formalism for macroscopic modeling of asynchronous irregular activity states*, Neural Comput., 21 (2009), pp. 46–100.
- [6] P. C. BRESSLOFF, *Traveling fronts and wave propagation failure in an inhomogeneous neural network*, Phys. D, 155 (2001), pp. 83–100.
- [7] P. C. BRESSLOFF, *Pattern formation in visual cortex*, in Les Houches 2003: Methods and Models in Neurophysics, C. C. Chow, B. Gutkin, D. Hansel, C. Meunier, and J. Dalibard, eds., Elsevier, New York, 2005, pp. 477–574.
- [8] P. C. BRESSLOFF, J. D. COWAN, M. GOLUBITSKY, P. J. THOMAS, AND M. WIENER, *Geometric visual hallucinations, euclidean symmetry and the functional architecture of striate cortex*, Philos. Trans. R. Soc. Lond. B, 356 (2001), pp. 299–330.
- [9] N. BRUNEL, *Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons*, J. Comput. Neurosci., 8 (2000), pp. 183–208.
- [10] N. BRUNEL AND V. HAKIM, *Fast global oscillations in networks of integrate-and-fire neurons with low firing rates*, Neural Comput., 11 (1999), pp. 1621–1671.
- [11] M. BUICE, *private communication* (2009).
- [12] M. BUICE AND J. D. COWAN, *Field-theoretic approach to fluctuation effects in neural networks*, Phys. Rev. E, 75 (2007), p. 051919.
- [13] M. BUICE, J. D. COWAN, AND C. C. CHOW, *Systematic fluctuation expansion for neural network activity equations*, to appear in Neural Comput.
- [14] D. CAI, L. TAO, M. SHELLY, AND D. W. MCLAUGHLIN, *An effective kinetic representation of fluctuation-driven neuronal networks with application to simple and complex cells in visual cortex*, Proc. Natl. Acad. Sci. USA, 101 (2004), pp. 7757–7562.
- [15] M. COHEN AND S. GROSSBERG, *Absolute stability and global pattern formation and parallel storage by competitive neural networks*, IEEE Trans. Syst. Man. Cybern., 13 (1983), pp. 815–826.

- [16] S. COOMBES, *Waves, bumps and patterns in neural field theories*, Biol. Cybern., 93 (2005), pp. 91–108.
- [17] S. COOMBES AND M. R. OWEN, *Evans functions for integral neural field equations with Heaviside firing rate function*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 574–600.
- [18] C. R. DOERING, K. V. SARGSYAN, L. M. SANDER, AND E. VANDEN-ELINDEN, *Asymptotics of rare events in birth-death processes bypassing the exact solutions*, J. Phys. Cond. Matter, 19 (2007), p. 065145.
- [19] M. DOI, *Second quantization representation for classical many-particle systems*, J. Phys. A, 9 (1976), pp. 1465–1477.
- [20] M. DOI, *Stochastic theory of diffusion controlled reactions*, J. Phys. A, 9 (1976), pp. 1479–1495.
- [21] M. I. DYKMAN, E. MORI, J. ROSS, AND P. M. HUNT, *Large fluctuations and optimal paths in chemical kinetics*, J. Chem. Phys. A, 100 (1994), pp. 5735–5750.
- [22] V. ELGART AND A. KAMENEV, *Rare event statistics in reaction-diffusion systems*, Phys. Rev. E, 70 (2004), p. 041106.
- [23] V. ELGART AND A. KAMENEV, *Classification of phase transitions in reaction-diffusion systems*, Phys. Rev. E, 74 (2006), p. 041101.
- [24] G. B. ERMENTROUT, *Neural networks as spatio-temporal pattern-forming systems*, Rep. Progr. Phys., 61 (1998), pp. 353–430.
- [25] G. B. ERMENTROUT AND J. COWAN, *A mathematical theory of visual hallucination patterns*, Biol. Cybern., 34 (1979), pp. 137–150.
- [26] G. B. ERMENTROUT AND J. D. COWAN, *Large scale spatially organized activity in neural nets*, SIAM J. Appl. Math., 38 (1980), pp. 1–21.
- [27] G. B. ERMENTROUT AND J. B. MCLEOD, *Existence and uniqueness of traveling waves for a neural network*, Proc. R. Soc., 123A (1993), pp. 461–478.
- [28] C. ESCUDERO AND A. KAMANEV, *Switching rates of multistep reactions*, Phys. Rev. E, 79 (2009), p. 041149.
- [29] C. ESCUDERO AND J. A. RODRIGUEZ, *Persistence of instanton connections in chemical reactions with time-dependent rates*, Phys. Rev. E, 77 (2008), p. 011130.
- [30] O. FAUGERAS, F. GRIMBERT, AND J. J. SLOTINE, *Absolute stability and complete synchronization in a class of neural fields models*, SIAM J. Appl. Math, 69 (2008), pp. 205–250.
- [31] S. E. FOLIAS AND P. C. BRESSLOFF, *Breathing pulses in an excitatory neural network*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 378–407.
- [32] S. E. FOLIAS AND P. C. BRESSLOFF, *Stimulus-locked traveling waves and breathers in an excitatory neural network.*, SIAM J. Appl. Math., 65 (2005), pp. 2067–2092.
- [33] S. E. FOLIAS AND P. C. BRESSLOFF, *Breathers in two-dimensional neural media*, Phys. Rev. Lett., 95 (2005), p. 208107.
- [34] M. I. FREIDLIN AND A. D. WENTZELL, *Random Perturbations of Dynamical Systems*, Springer, New York, 1984.
- [35] W. GERSTNER AND W. KISTLER, *Spiking Neuron Models*, Cambridge University Press, Cambridge, 2002.
- [36] 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.
- [37] I. GINZBURG AND H. SOMPOLINSKY, *Theory of correlations in stochastic neural networks*, Phys. Rev. E, 50 (1994), pp. 3171–3191.
- [38] R. GRAHAM AND T. TEL, *Weak-noise limit of Fokker-Planck models and nondifferentiable potentials for dissipative dynamical systems*, Phys. Rev. A, 31 (1985), pp. 1109–1122.
- [39] R. GRAHAM AND T. TEL, *Kinetic theory of coupled oscillators*, Phys. Rev. Lett., 98 (2007), p. 054101.
- [40] E. J. HILDEBRAND, M. A. BUICE, AND C. C. CHOW, *Correlations, fluctuations, and stability of a finite-size network of coupled oscillators*, Phys. Rev. E, 76 (2007), p. 031118.
- [41] R. HINCH AND S. J. CHAPMAN, *Exponentially slow transitions on a Markov chain: The frequency of calcium sparks*, Eur. J. Appl. Math., 16 (2005), pp. 427–446.
- [42] H. HINRICHSEN, *Non-equilibrium critical phenomena and phase transitions into absorbing states*, Adv. Phys., 49 (2000), pp. 815–958.
- [43] J. J. HOPFIELD, *Neurons with graded response have collective computational properties like those of two-state neurons*, Proc. Natl. Acad. Sci. USA, 81 (1984), pp. 3088–3092.
- [44] X. HUANG, W. C. TROY, Q. YANG, H. MA, C. R. LAING, S. J. SCHIFF, AND J. Y. WU, *Spiral waves in disinhibited mammalian neocortex*, J. Neurosci., 24 (2004), pp. 9897–9902.
- [45] D. H. HUBEL AND T. N. WIESEL, *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex*, J. Neurosci., 3 (1962), pp. 1116–1133.
- [46] V. K. JIRSA AND H. HAKEN, *A derivation of a macroscopic field theory of the brain from the quasi-microscopic neural dynamics.*, Phys. D, 99 (1997), pp. 503–526.

- [47] N. G. VAN KAMPEN, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1992.
- [48] Z. P. KILAPTRICK AND P. C. BRESSLOFF, *Spatially structured oscillations in a two-dimensional neuronal network with synaptic depression*, to appear in *J. Comput. Neurosci.*
- [49] C. R. LAING, *Spiral waves in nonlocal equations*, *SIAM J. Appl. Dyn. Syst.*, 4 (2005), pp. 588–606.
- [50] C. R. LAING AND C. C. CHOW, *Stationary bumps in networks of spiking neurons*, *Neural Comput.*, 13 (2001), pp. 1473–1494.
- [51] C. R. LAING, W. C. TROY, B. GUTKIN, AND G. B. ERMENTROUT, *Multiple bumps in a neuronal model of working memory*, *SIAM J. Appl. Math.*, 63 (2002), pp. 62–97.
- [52] C. LY AND D. TRANCHINA, *Critical analysis of a dimension reduction by a moment closure method in a population density approach to neural network modeling*, *Neural Comput.*, 19 (2007), pp. 2032–2092.
- [53] M. MATTIA AND P. DEL GUIDICE, *Population dynamics of interacting spiking neurons*, *Phys. Rev. E*, 66 (2002), p. 051917.
- [54] C. MEYER AND C. VAN VREESWIJK, *Temporal correlations in stochastic networks of spiking neurons*, *Neural Comput.*, 14 (2002), pp. 369–404.
- [55] P. I. NUNEZ, *Neocortical Dynamics and Human EEG Rhythms*, Oxford University Press, New York, 1995.
- [56] D. NYKAMP AND D. TRANCHINA, *A population density method that facilitates large-scale modeling of neural networks: Analysis and application to orientation tuning*, *J. Comp. Neurosci.*, 8 (2000), pp. 19–50.
- [57] T. OHIRA AND J. D. COWAN, *Stochastic neurodynamics and the system size expansion*, in *Proceedings of the First International Conference on Mathematics of Neural Networks*, S. Ellacott and I. J. Anderson, eds., Academic Press, New York, 1997, pp. 290–294.
- [58] A. OMURTAG, B. W. KNIGHT, AND L. SIROVICH, *On the simulation of large populations of neurons*, *J. Comput. Neurosci.*, 8 (2000), pp. 51–63.
- [59] M. R. OWEN, C. R. LAING, AND S. COOMBES, *Bumps and rings in a two-dimensional neural field: Splitting and rotational instabilities*, *New J. Phys.*, 9 (2007), p. 378.
- [60] L. PELITI, *Path integral approach to birth-death processes on a lattice*, *J. Phys.*, 46 (1985), pp. 1469–1483.
- [61] D. J. PINTO AND G. B. ERMENTROUT, *Spatially structured activity in synaptically coupled neuronal networks: I. Traveling fronts and pulses*, *SIAM J. Appl. Math.*, 62 (2001), pp. 206–225.
- [62] D. PINTO AND G. B. ERMENTROUT, *Spatially structured activity in synaptically coupled neuronal networks: II. Lateral inhibition and standing pulses*, *SIAM J. Appl. Math.*, 62 (2001), pp. 226–243.
- [63] D. J. PINTO, J. C. BRUMBERG, D. J. SIMONS, AND G. B. ERMENTROUT, *A quantitative population model of whisker barrels: Re-examining the Wilson-Cowan equations*, *J. Comput. Neurosci.*, 3 (1996), pp. 247–264.
- [64] D. PLENZ AND T. C. THIAGARAJAN, *The organizing principles of neuronal avalanches: Cell assemblies in the cortex?*, *Trends Neurosci.*, 30 (2007), pp. 101–110.
- [65] A. V. RANGAN, G. KOVACIC, AND D. CAI, *Kinetic theory for neuronal networks with fast and slow excitatory conductances driven by the same spike train*, *Phys. Rev. E*, 77 (2008), p. 041915.
- [66] A. RENART, N. BRUNEL, AND X.-J. WANG, *Mean-field theory of irregularly spiking neuronal populations and working memory in recurrent cortical networks*, in *Computational Neuroscience: A Comprehensive Approach*, J. Feng, ed., CRC Press, Boca Raton, FL, 2004, pp. 431–490.
- [67] P. A. ROBINSON, C. J. RENNIE, J. J. WRIGHT, H. BAHRAMALI, E. GORDON, AND D. I. ROWE, *Prediction of EEG spectra from neurophysiology*, *Phys. Rev. E*, 63 (2001), p. 021903.
- [68] W. R. SOFTKY AND C. KOCH, *The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs*, *J. Neurosci.*, 13 (1993), pp. 334–350.
- [69] H. SOULA AND C. C. CHOW, *Stochastic dynamics of a finite-size spiking neural network*, *Neural Comput.*, 19 (2007), pp. 3262–3292.
- [70] P. TASS, *Cortical pattern formation during visual hallucinations.*, *J. Biol. Phys.*, 21 (1995), pp. 177–210.
- [71] W. C. TROY AND V. SHUSTERMAN, *Patterns and features of families of traveling waves in large-scale neuronal networks*, *SIAM J. Appl. Dyn. Syst.*, 6 (2007), pp. 263–292.
- [72] H. R. WILSON AND J. D. COWAN, *Excitatory and inhibitory interactions in localized populations of model neurons*, *Biophys. J.*, 12 (1972), pp. 1–23.
- [73] H. R. WILSON AND J. D. COWAN, *A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue*, *Kybernetik*, 13 (1973), pp. 55–80.
- [74] L. ZHANG, *On stability of traveling wave solutions in synaptically coupled neuronal networks*, *Differential Integral Equations*, 16 (2003), pp. 513–536.