The Journal of Mathematical Neuroscience
a SpringerOpen Journal

RESEARCH   **Open Access**

Path-Integral Methods for Analyzing the Effects of Fluctuations in Stochastic Hybrid Neural Networks

Paul C. Bressloff[1]

Abstract We consider applications of path-integral methods to the analysis of a stochastic hybrid model representing a network of synaptically coupled spiking neuronal populations. The state of each local population is described in terms of two stochastic variables, a continuous synaptic variable and a discrete activity variable. The synaptic variables evolve according to piecewise-deterministic dynamics describing, at the population level, synapses driven by spiking activity. The dynamical equations for the synaptic currents are only valid between jumps in spiking activity, and the latter are described by a jump Markov process whose transition rates depend on the synaptic variables. We assume a separation of time scales between fast spiking dynamics with time constant τ_a and slower synaptic dynamics with time constant τ. This naturally introduces a small positive parameter $\epsilon = \tau_a/\tau$, which can be used to develop various asymptotic expansions of the corresponding path-integral representation of the stochastic dynamics. First, we derive a variational principle for maximum-likelihood paths of escape from a metastable state (large deviations in the small noise limit $\epsilon \to 0$). We then show how the path integral provides an efficient method for obtaining a diffusion approximation of the hybrid system for small ϵ. The resulting Langevin equation can be used to analyze the effects of fluctuations within the basin of attraction of a metastable state, that is, ignoring the effects of large deviations. We illustrate this by using the Langevin approximation to analyze the effects of intrinsic noise on pattern formation in a spatially structured hybrid network. In particular, we show how noise enlarges the parameter regime over which patterns occur, in an analogous fashion to PDEs. Finally, we carry out a $1/\epsilon$-loop expansion of the path integral, and use this to derive corrections to voltage-based mean-field equations, analogous to the modified activity-based equations generated from a neural master equation.

✉ P.C. Bressloff
   bressloff@math.utah.edu

[1]   Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, UT
       84112, USA

1 Introduction

One of the major challenges in neuroscience is developing our understanding of how noise at the molecular and cellular levels affects dynamics and information processing at the macroscopic level of synaptically coupled neuronal populations. It is well known that the spike trains of individual cortical neurons in vivo tend to be very noisy, having interspike interval (ISI) distributions that are close to Poisson [1, 2]. Indeed, one observes trial-to-trial variability in spike trains, even across trials in which external stimuli are identical. On the other hand, neurons are continuously bombarded by thousands of synaptic inputs, many of which are uncorrelated, so that an application of the law of large numbers would suggest that total input fluctuations are small. This would make it difficult to account for the Poisson-like behavior of individual neurons, even when stochastic ion channel fluctuations or random synaptic background activity is taken into account. One paradigm for reconciling these issues is the so-called balanced network [3–5]. In such networks, each neuron is driven by a combination of strong excitation and strong inhibition, which mainly cancel each other out, so that the remaining fluctuations occasionally and irregularly push the neuron over the firing threshold. Even in the absence of any external sources of noise, the resulting deterministic dynamics is chaotic and neural outputs are Poisson-like. Interestingly, there is some experimental evidence that cortical networks can operate in a balanced regime [6].

Another emergent feature of balanced networks is that they can support an asynchronous state characterized by large variability in single neuron spiking, and yet arbitrarily small pairwise correlations, even in the presence of substantial amounts of shared inputs [7]. Thus there is a growing consensus that the trial-to-trial irregularity in the spiking of individual neurons is often unimportant, and that information is typically encoded in firing rates. There is then another level of neural variability, namely, trial-to-trial variations in the firing rates themselves. Recent physiological data shows that the onset of a stimulus reduces firing-rate fluctuations in cortical neurons, while having little or no effect on the spiking variability [8]. Litwin-Kumar and Doiron have recently shown how these two levels of stochastic variability can emerge in a balanced network of randomly connected spiking neurons, in which a small amount of clustered connections induces firing-rate fluctuations superimposed on spontaneous spike fluctuations [9].

Various experimental and computational studies of neural variability thus motivate the incorporation of noise into rate-based neural network models [10]. One approach is to add extrinsic noise terms to deterministic models resulting in a neural Langevin equation [11–15]. An alternative approach is to assume that noise arises intrinsically as a collective population effect, and to describe the stochastic dynamics in terms of a neural master equation [16–20]. In the latter case, neurons are partitioned into a set of M local homogeneous populations labeled $\alpha = 1, \ldots, M$, each consisting of $\mathcal{N}$ neurons. The state of each population at time t is specified by the number

$\mathcal{N}_\alpha(t)$ of active neurons in a sliding window $(t, t + \Delta t]$, and transition rates between the discrete states are chosen so that standard rate-based models are obtained in the mean-field limit, where statistical correlations can be ignored. There are two versions of the neural master equation, which can be distinguished by the size of the sliding window width Δt. (Note that the stochastic models are keeping track of *changes* in population activity.) One version assumes that each population operates close to an asynchronous state for large $\mathcal{N}$ [18, 19], so that one-step changes in population activity occur relatively slowly. Hence, one can set $\Delta t = 1$ and take $\mathcal{N}$ to be large but finite. The other version of the neural master equation assumes that population activity is approximately characterized by a Poisson process [17, 20]. In order to maintain a one-step jump Markov process, it is necessary to take the limits $\Delta t \to 0$, $\mathcal{N} \to \infty$ such that $\mathcal{N}\Delta t = 1$. Thus, one considers the number of active neurons in an infinite background sea of inactive neurons, which is reasonable if the networks are in low activity states. (Note that it is also possible to interpret the master equation of Buice et al. in terms of activity states of individual neurons rather than populations [17, 20].)

One way to link the two versions of the neural master equation is to extend the Doi–Peliti path-integral representation of chemical master equations [21–23] to the neural case; the difference between the two versions then reduces to a different choice of scaling of the underlying action functional [18]. Buice et al. [17, 20] used diagrammatic perturbations methods (Feynman graphs) to generate a truncated moment hierarchy based on factorial moments, and thus determined corrections to mean-field theory involving coupling to two-point and higher-order cumulants. They also used renormalization group methods to derive scaling laws for statistical correlations close to criticality, that is, close to a bifurcation point of the underlying deterministic model [17]. On the other hand, Bressloff [18, 19] showed how 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 metastable state, following along analogous lines to previous work on large deviations in chemical master equations [24–26].

One limitation of both versions of the neural master equation is that they neglect the dynamics of synaptic currents. The latter could be particularly significant if the time scale τ of synaptic dynamics is larger than the window width Δt. Therefore, we recently extended the Buice et al. neural master equation by formulating the network population dynamics in terms of a stochastic hybrid system also known as a 'velocity' jump Markov process [27]. The state of each population is now described in terms of two stochastic variables $U_\alpha(t)$ and $\mathcal{N}_\alpha(t)$. The synaptic variables $U_\alpha(t)$ evolve according to piecewise-deterministic dynamics describing, at the population level, synapses driven by spiking activity. These equations are only valid between jumps in spiking activity $\mathcal{N}_\alpha(t)$, which are described by a jump Markov process whose transition rates depend on the synaptic variables. We also showed how asymptotic methods recently developed to study metastability in other stochastic hybrid systems, such as stochastic ion channels, motor-driven intracellular cargo transport, and gene networks [28–32], can be extended to analyze metastability in stochastic hybrid neural networks, in a regime where the synaptic dynamics is much slower than the spiking dynamics. In the case of ion channels, $\mathcal{N}_\alpha$ would represent the number of open channels of type α, whereas U_α would be replaced by the membrane voltage V. On the

other hand, for intracellular transport, $\mathcal{N}_\alpha$ would be the number of motors of type α actively transporting a cargo and U_α would be replaced by spatial position along the track.

In this paper we show how a path-integral representation of a stochastic hybrid neural network provides a unifying framework for a variety of asymptotic perturbation methods. The basic hybrid neural network model is described in Sect. 2, where we consider several limiting cases. In Sect. 3, we reprise the path-integral construction of Bressloff and Newby [33], highlighting certain features that were not covered in the original treatment, including the connection with large-deviation principles [34], and potential difficulties in the thermodynamic limit $\mathcal{N} \to \infty$. In Sect. 4, we derive the basic variational principle that can be used to explore maximum-likelihood paths of escape from a metastable state, and relate the theory to the underlying Hamiltonian structure of the path-integral representation. In Sect. 5, we show how the path-integral representation provides an efficient method for deriving a diffusion approximation of a stochastic hybrid neural network. Although the diffusion approximation breaks down when considering escape problems, it provides useful insights into the effects of fluctuations within the basin of attraction of a given solution. We illustrate this by using the diffusion approximation to explore the effects of noise on neural pattern formation in a spatially structured network. In particular, we show how noise expands the parameter regime over which patterns can be observed, in an analogous fashion to stochastic PDEs. Finally, in Sect. 6, we use the path-integral representation to derive corrections to voltage-based mean-field equations, along analogous lines to the analysis of activity-based mean-field equations arising from the neural master equation [17, 20].

2 Stochastic Hybrid Network Model

We first describe a stochastic neural network model that generalizes the neural master equation [17, 18, 20] by incorporating synaptic dynamics. (A more detailed derivation of the model can be found in [27].) Note that there does not currently exist a complete, rigorous derivation of population rate-based models starting from detailed biophysical models of individual neurons, although some significant progress has been made in a series of papers by Buice and Chow on generalized activity equations for theta neurons [35–37]. Therefore, the construction of the stochastic rate-based model is phenomenological in nature. However, it is motivated by the idea that finite-size effects in local populations of neurons acts as a source of intrinsic noise. Consider a set of M homogeneous populations labeled $\alpha = 1, \ldots, M$, with $\mathcal{N}$ neurons in each population. (A straightforward generalization would be for each population to consist of $\mathcal{O}(\mathcal{N})$ neurons.) The output activity of each population is taken to be a discrete stochastic variable $A_\alpha(t)$ given by

$$A_\alpha(t) = \frac{\mathcal{N}_\alpha(t)}{\mathcal{N} \Delta t}, \qquad (2.1)$$

where $\mathcal{N}_\alpha(t)$ is the number of neurons in the αth population that fired in the time interval $[t - \Delta t, t]$, and Δt is the width of a sliding window that counts spikes. The

discrete stochastic variables $\mathcal{N}_\alpha(t)$ are taken to evolve according to a one-step jump Markov process:

$$\mathcal{N}_\alpha(t) \overset{\omega_+/\tau_a}{\to} \mathcal{N}_\alpha(t) + 1, \qquad \mathcal{N}_\alpha(t) \overset{\omega_-/\tau_a}{\to} \mathcal{N}_\alpha(t) - 1 \tag{2.2}$$

with corresponding transition rates

$$\omega_+ = \mathcal{N}\Delta t\, F(U_\alpha), \qquad \omega_- = \mathcal{N}_\alpha. \tag{2.3}$$

Here F is a sigmoid firing-rate or gain function

$$F(u) = \frac{F_0}{1 + e^{-\gamma(u-\kappa)}}, \tag{2.4}$$

where γ, κ correspond to the gain and threshold, respectively, F_0 is the maximum firing rate, and $U_\alpha(t)$ is the effective synaptic current into the αth population, which evolves (for exponential synapses) according to

$$\tau\, dU_\alpha(t) = \left[-U_\alpha(t) + \sum_{\beta=1}^{M} w_{\alpha\beta} A_\beta(t) \right] dt. \tag{2.5}$$

We will assume that $\mathcal{N}$ is large but finite and take $\mathcal{N}\Delta t = 1$. In the dual limits $\mathcal{N} \to \infty$ and $\tau \to 0$, our model then reduces to the Buice et al. [17, 20] version of the neural master equation. The resulting stochastic process defined by (2.1)–(2.5) is an example of a stochastic hybrid system based on a piecewise-deterministic process. That is, the transition rate ω_+ depend on U_α, with the latter itself coupled to the associated jump Markov according to (2.5), which is only defined between jumps, during which $U_\alpha(t)$ evolves deterministically. It is important to note that the time constant τ_a cannot be identified directly with membrane or synaptic time constants. Instead, it determines the relaxation rate of a local population to the instantaneous firing rate.

Introduce the probability density

$$\text{Prob}\{U_\alpha(t) \in (u_\alpha, u_\alpha + du), \mathcal{N}_\alpha(t) = n_\alpha; \alpha = 1, \ldots, M\}$$
$$= p(\mathbf{u}, \mathbf{n}, t | \mathbf{u}_0, \mathbf{n}_0, 0)\, d\mathbf{u},$$

with $\mathbf{u} = (u_1, \ldots, u_M)$ and $\mathbf{n} = (n_1, \ldots, n_M)$. It follows from (2.1)–(2.5) that the probability density evolves according to the differential Chapman–Kolmogorov (CK) equation (dropping the explicit dependence on initial conditions)

$$\frac{\partial p}{\partial t} + \frac{1}{\tau} \sum_\alpha \frac{\partial [v_\alpha(\mathbf{u}, \mathbf{n})\, p(\mathbf{u}, \mathbf{n}, t)]}{\partial u_\alpha}$$
$$= \frac{1}{\tau_a} \sum_\alpha (\mathbb{T}_\alpha - 1)\big(n_\alpha\, p(\mathbf{u}, \mathbf{n}, t)\big)$$
$$+ \frac{1}{\tau_a} \sum_\alpha (\mathbb{T}_\alpha^{-1} - 1)\big(F(u_\alpha)\, p(\mathbf{u}, \mathbf{n}, t)\big), \tag{2.6}$$

with

$$v_\alpha(\mathbf{u}, \mathbf{n}) = -u_\alpha + \sum_\beta w_{\alpha\beta} n_\beta, \tag{2.7}$$

and $\mathbb{T}_\alpha$ the translation operator: $\mathbb{T}_\alpha^{\pm 1} f(\mathbf{n}) = f(\mathbf{n}_{\alpha\pm})$ for any function f with $\mathbf{n}_{\alpha\pm}$ denoting the configuration with n_α replaced by $n_\alpha \pm 1$. Equation (2.6) can be re-expressed in the more general form

$$\frac{\partial p}{\partial t} = -\frac{1}{\tau} \sum_{\alpha=1}^M \frac{\partial}{\partial u_\alpha} \big(v_\alpha(\mathbf{u}, \mathbf{n}) p(\mathbf{u}, \mathbf{n}, t)\big) + \frac{1}{\tau_a} \sum_{\mathbf{m}} W(\mathbf{n}, \mathbf{m}; \mathbf{u}) p(\mathbf{u}, \mathbf{m}, t). \tag{2.8}$$

The drift 'velocities' $v_\alpha(\mathbf{u}, \mathbf{n})$ for fixed $\mathbf{n}$ represent the piecewise-deterministic synaptic dynamics according to

$$\tau \frac{du_\alpha}{dt} = v_\alpha(\mathbf{u}, \mathbf{n}), \quad \alpha = 1, \ldots, M, \tag{2.9}$$

and W is defined in terms of the $\mathbf{u}$-dependent transition matrix T for the jump Markov process, that is,

$$W(\mathbf{n}, \mathbf{m}; \mathbf{u}) = T(\mathbf{n}, \mathbf{m}; \mathbf{u}) - \delta_{\mathbf{n},\mathbf{m}} \sum_{\mathbf{k}} T(\mathbf{k}, \mathbf{m}; \mathbf{u}).$$

It follows from (2.6) that W can be written as

$$W(\mathbf{n}, \mathbf{m}; \mathbf{u}) = \sum_{\alpha=1}^M W^\alpha(n_\alpha, m_\alpha; u_\alpha) \prod_{\beta \neq \alpha} \delta_{n_\beta, m_\beta}$$

with W^α the tridiagonal matrix

$$W^\alpha(n, n-1; u) = F(u), \qquad W^\alpha(n, n+1; u) = n+1,$$
$$W^\alpha(n, n; u) = -F(u) - n.$$

For fixed u, the matrix W^α is irreducible (which means that there is a non-zero probability of transitioning, possibly in more than one step, from any state to any other state in the jump Markov process). Moreover, all off-diagonal elements are non-negative. It follows that the full transition matrix $W(\mathbf{n}, \mathbf{m}; \mathbf{u})$ also has these properties and, hence, we can apply the Perron–Frobenius theorem to show that there exists a unique invariant measure for the Markov process. That is, the master equation

$$\frac{dp(\mathbf{u}, \mathbf{n}, t)}{dt} = \frac{1}{\tau_a} \sum_{\mathbf{m}} W(\mathbf{n}, \mathbf{m}; \mathbf{u}) p(\mathbf{u}, \mathbf{m}, t),$$

has a globally attracting steady state $\rho(\mathbf{u}, \mathbf{n})$ such that $p(\mathbf{u}, \mathbf{n}, t) \to \rho(\mathbf{u}, \mathbf{n})$ as $t \to \infty$.

The Perron–Frobenius theorem states that [38] a real square matrix with positive entries has a unique largest real eigenvalue (the Perron eigenvalue) and that the corresponding eigenvector has strictly positive components. If we define a new transition matrix $\widehat{W}(\mathbf{n}, \mathbf{m}; \mathbf{u})$ by

$$\widehat{W}(\mathbf{n}, \mathbf{m}; \mathbf{u}) = W(\mathbf{n}, \mathbf{m}; \mathbf{u}) + \delta_{\mathbf{n},\mathbf{m}} W^*, \quad W^* = \max_{\mathbf{m}} W(\mathbf{m}, \mathbf{m}; \mathbf{u}) + \kappa,$$

for an arbitrary $\kappa > 0$, then we can apply the Perron–Frobenius theorem directly to $\widehat{W}$ and thus to W. Since $\sum_{\mathbf{n}} W(\mathbf{n}, \mathbf{m}; \mathbf{u}) = 0$ for all $\mathbf{m}$, that is, $\eta(\mathbf{n}) = (1, 1, \ldots, 1)^T$ is a left null-vector, it follows that the Perron eigenvalue is $\lambda = 0$. The unique invariant measure then corresponds to the right null-vector of W for fixed $\mathbf{u}$:

$$\sum_{\mathbf{m}} W(\mathbf{n}, \mathbf{m}; \mathbf{u}) \rho(\mathbf{u}, \mathbf{m}) = 0. \tag{2.10}$$

The steady-state solution $\rho(\mathbf{u}, \mathbf{n})$ of (2.6) can be factorized as $\rho(\mathbf{u}, \mathbf{n}) = \prod_{\beta=1}^{M} \rho_0(u_\beta, n_\beta)$ with

$$0 = \sum_{\alpha=1}^{M} \left[\prod_{\beta \neq \alpha} \rho_0(u_\beta, n_\beta) \right] \left[J(u_\alpha, n_\alpha + 1) - J(u_\alpha, n_\alpha) \right], \tag{2.11}$$

where

$$J(u, n) = n \rho_0(u, n) - F(u) \rho_0(u, n - 1).$$

A sufficient condition for (2.11) to hold is

$$J(u, n + 1) - J(u, n) = 0.$$

Since $\rho_0(u, -1) \equiv 0$, it then follows that $J(u, 0) = 0$ and thus $J(u, n) = 0$ for all n. Hence, we obtain the positive steady-state solution

$$\rho_0(u, n) = \rho_0(u, 0) \prod_{m=1}^{n} \frac{F(u)}{m} = \rho_0(u, 0) \frac{F(u)^n}{n!}. \tag{2.12}$$

The Perron–Frobenius theorem ensures that this is the unique positive solution. The fact that the steady state factorizes is a consequence of the fact that the transition rates do not involve any coupling between populations—the only coupling appears in the drift terms of (2.6). Strictly speaking, the Perron–Frobenius theorem applies to finite-dimensional matrices, so we are assuming that $\mathcal{N}$ is finite. Nevertheless, in the thermodynamic limit $\mathcal{N} \to \infty$, the corresponding normalized density reduces to a Poisson process with rate $F(u)$:

$$\rho_0(u, n) = e^{-F(u)} \frac{F(u)^n}{n!}. \tag{2.13}$$

There are two time scales in the CK equation (2.8), the synaptic time constant τ and the time constant τ_a, which characterizes the relaxation rate of population

activity. In the limit $\tau \to 0$, (2.5) reduces to the neural master equation of Buice et al. [17, 20]. First, note that the synaptic variables $U_\alpha(t)$ are eliminated by setting $v_\alpha = 0$, that is, $U_\alpha(t) = \sum_\beta w_{\alpha\beta} A_\beta(t)$. This then leads to a pure birth-death process for the discrete variables $\mathcal{N}_\alpha(t)$. That is, let $P(\mathbf{n}, t) = \mathrm{Prob}[\mathcal{N}(t) = \mathbf{n}]$ denote the probability that the network of interacting populations has configuration $\mathbf{n} = (n_1, n_2, \ldots, n_M)$ at time t, $t > 0$, given some initial distribution $P(\mathbf{n}, 0)$. The probability distribution then evolves according to the birth-death master equation [17, 18, 20]

$$\frac{dP(\mathbf{n}, t)}{dt} = \sum_\alpha \big[(\mathbb{T}_\alpha - 1)\big(\Omega_\alpha^-(\mathbf{n}) P(\mathbf{n}, t)\big)$$
$$+ \big(\mathbb{T}_\alpha^{-1} - 1\big)\big(\Omega_\alpha^+(\mathbf{n}) P(\mathbf{n}, t)\big)\big], \tag{2.14}$$

where

$$\Omega_\alpha^+(\mathbf{n}) = \frac{1}{\tau_a} F\Big(\sum_\beta w_{\alpha\beta} n_\beta\Big), \qquad \Omega_\alpha^-(\mathbf{n}) = \frac{n_\alpha}{\tau_a}. \tag{2.15}$$

Buice et al. [20] show that the network operates in a Poisson-like regime in which the rates of the Poisson process are stochastic variables whose means evolve according to the activity-based mean-field equation

$$\tau_\alpha \frac{dA_\alpha}{dt} = -A_\alpha(t) + F\Big(\sum_\beta w_{\alpha\beta} A_\beta(t)\Big). \tag{2.16}$$

On the other hand, if $\tau_a \to 0$ for fixed τ, then we obtain deterministic voltage or current-based mean-field equations

$$\tau \frac{du_\alpha}{dt} = \sum_\mathbf{n} v_\alpha\big(\mathbf{u}(t), \mathbf{n}\big) \rho\big(\mathbf{u}(t), \mathbf{n}\big)$$
$$= -u_\alpha(t) + \sum_{\beta=1}^M w_{\alpha\beta} \sum_\mathbf{n} n_\beta \rho\big(\mathbf{u}(t), \mathbf{n}\big). \tag{2.17}$$

Since $\rho(\mathbf{u}, \mathbf{n})$ is given by a product of independent Poisson processes with rates $F(u_\alpha)$, consistent with the operating regime of the Buice et al. master equation [17, 20], it follows that

$$\langle n_\beta \rangle = F(u_\beta), \tag{2.18}$$

and (2.17) reduces to the standard voltage or current-based activity equation

$$\tau \frac{du_\alpha}{dt} = -u_\alpha(t) + \sum_{\beta=1}^M w_{\alpha\beta} F(u_\beta). \tag{2.19}$$

Note that the limit $\tau_a \to 0$ is analogous to the slow synapse approximation used by Ermentrout [39] to reduce deterministic conductance-based neuron models to voltage-based rate models. Now suppose that the network operates in the regime

$0 < \tau_a/\tau \equiv \epsilon \ll 1$. There is then a natural small parameter in the system, ϵ, which allows a variety of perturbation methods to be used:

(i) A quasi-steady-state (QSS) diffusion approximation of the stochastic hybrid system, in which the CK equation (2.6) reduces to a Fokker–Planck equation [27]. This exploits the fact that for small ϵ there are typically a large number of transitions between different firing states $\mathbf{n}$ while the synaptic currents $\mathbf{u}$ hardly change at all. This implies that the system rapidly converges to the (quasi) steady state $\rho(\mathbf{u}, \mathbf{n})$, which will then be perturbed as $\mathbf{u}$ slowly evolves.

(ii) The diffusion approximation captures the Gaussian-like fluctuations within the basin of attraction of a fixed point of the mean-field equations. However, for small ϵ this yields exponentially large errors for the transition rates between metastable states. (A similar problem arises in approximating chemical and neural master equations by a Fokker–Planck equation in the large N limit [19, 24, 40].) However, one can use a Wentzel–Kramers–Brillouin (WKB) approximation of solutions to the full CK equation to calculate the mean first passage time for escape [27].

(iii) Another way to analyze the dynamics of a stochastic hybrid network is to derive moment equations. However, for a nonlinear system, this yields an infinite hierarchy of coupled moment equations, resulting in the problem of moment closure. In the case of small ϵ, one can expand the moment equations to some finite order in ϵ.

In this paper, we show how a path-integral representation of a stochastic hybrid system provides a unifying framework for carrying out all three perturbation schemes highlighted above.

3 Path-Integral Representation

3.1 One-Population Model

We now derive the path-integral representation of a stochastic hybrid neural network using the construction introduced in [33]. For ease of notation, we consider a one-population model ($M = 1$); the generalization to multiple populations is then straightforward (see Sect. 3.4). We first discretize time by dividing a given interval $[0, T]$ into N equal subintervals of size Δt such that $T = N\Delta t$ and set $u_j = u(j\Delta t)$, $n_j = n(j\Delta t)$. (Note that the infinitesimal time interval Δt used in path discretization is distinct from the width of the moving window used in the construction of the stochastic neural network; see Sect. 2. One should also take care to distinguish between the discrete time label j and the population label α.) The conditional probability density for $u_1, \ldots, u_N$, given u_0 and a particular realization of the stochastic discrete variables n_j, $j = 0, \ldots, N - 1$, is

$$P(u_1, \ldots, u_N | u_0, n_0, \ldots, n_{N-1}) = \prod_{j=0}^{N-1} \delta\big(u_{j+1} - u_j - v_{n_j}(u_j)\Delta t\big),$$

where

$$v_n(u) \equiv v(u, n) = -u + wn. \tag{3.1}$$

Inserting the Fourier representation of the Dirac delta function gives

$$P(u_1, \ldots, u_N | u_0, n_0, n_1, \ldots, n_{N-1})$$
$$= \prod_{j=0}^{N-1} \left[\int_{-\infty}^{\infty} e^{-ip_j(u_{j+1} - u_j - v_{n_j}(u_j)\Delta t)} \frac{dp_j}{2\pi} \right].$$

On averaging with respect to the intermediate states n_j, $j = 1, N - 1$, we have

$$P(u_1, \ldots, u_N | u_0, n_0)$$
$$= \left[\prod_{j=0}^{N-1} \int_{-\infty}^{\infty} \frac{dp_j}{2\pi} \right] \sum_{n_1, \ldots, n_{N-1}} \prod_{j=0}^{N-1} T_{n_{j+1}, n_j}(u_j) e^{-ip_j(u_{j+1} - u_j - v_{n_j}(u_j)\Delta t)}, \tag{3.2}$$

where

$$T_{n_{j+1}, n_j}(u_j) = \left(\delta_{n_{j+1}, n_j} + W_{n_{j+1}, n_j}(u_j) \frac{\Delta t}{\epsilon} \right) + o(\Delta t)$$

and $W_{nm}(u) \equiv W(n, m; u)$ such that

$$W_{n,n-1} = F(u), \qquad W_{nn} = -F(u) - n, \qquad W_{n,n+1} = n + 1. \tag{3.3}$$

In order to evaluate the above path integral, we introduce the eigenvalue equation

$$\sum_m \left[W_{nm}(u) + q\delta_{n,m} v_m(u) \right] R_m^{(s)}(u, q) = \lambda_s(u, q) R_n^{(s)}(u, q), \tag{3.4}$$

and let $\xi_m^{(s)}$ be the adjoint eigenvector satisfying

$$\sum_n \xi_n^{(s)}(u, q) \left[W_{nm}(u) + q\delta_{n,m} v_m(u) \right] = \lambda_s(u, q) \xi_n^{(s)}(u, q). \tag{3.5}$$

In our original construction of the path-integral representation [33], we arrived at (3.4) and its adjoint through trial and error, based on our previous work on WKB methods. It turns out that the principal eigenvalue of the linear equation (3.4) can be related to the rate function of large-deviation theory, as we explain in Sect. 3.2. A basic result from linear algebra is that $R^{(s)}$ and $\xi^{(s)}$ form a bi-orthonormal set for fixed u, q. First, rewrite (3.4) and (3.5) in the compact form

$$L^{\dagger} \xi^{(s)} = \lambda_s \xi^{(s)}, \qquad L R^{(s)} = \lambda_s R^{(s)}.$$

Defining the inner product $\langle \xi^{(s)}, R^{(s')} \rangle = \sum_n \xi_n^{(s)} R_n^{(s')}$, we see that

$$0 = \left\langle \xi^{(s)}, L R^{(s')} \right\rangle - \left\langle L^{\dagger} \xi^{(s)}, R^{(s')} \right\rangle = (\lambda_s - \lambda_{s'}) \left\langle \xi^{(s)}, R^{(s')} \right\rangle.$$

Thus, for distinct eigenvalues ($\lambda_s \neq \lambda_{s'}$) the eigenvectors $R^{(s')}$ and $\xi^{(s)}$ are orthogonal, and this can be extended to degenerate eigenvalues by Schmidt orthogonalization. Now suppose that we expand a general vector v according to $v = \sum_s c_s R^{(s)}$ for coefficients c_s. Biorthogonality implies that $c_s = \langle \xi^{(s)}, v \rangle$. Substituting back into the eigenvector expansion of v gives

$$v_n = \sum_s \langle \xi^{(s)}, v \rangle R_n^{(s)} = \sum_s \sum_m \xi_m^{(s)} \xi_m^{(s)} R_n^{(s)},$$

which leads to the completeness relation

$$\sum_s \xi_m^{(s)}(u, q) R_n^{(s)}(u, q) = \delta_{m,n} \tag{3.6}$$

for all u, q.

Now suppose that we insert multiple copies of the identity (3.6) into the path integral (3.2) with $q = q_j$ at the $(j + 1)$th time step. That is, taking

$$T_{n_{j+1}n_j}(u_j) = \sum_m T_{mn_j}(u_j) \delta_{m,n_{j+1}}$$

$$= \sum_{s_j,m} R_{n_{j+1}}^{(s_j)}(u_j, q_j) \xi_m^{(s_j)}(u_j, q_j) \left(\delta_{n_j,m} + A_{mn_j}(u_j) \frac{\Delta t}{\epsilon} \right)$$

$$= \sum_{s_j} \left(1 + [\lambda_{s_j}(u_j, q_j) - q_j v_{n_j}(u_j)] \frac{\Delta t}{\epsilon} \right) R_{n_{j+1}}^{(s_j)}(u_j, q_j) \xi_{n_j}^{(s)}(u_j, q_j)$$

$$\sim \sum_{s_j} \exp\left([\lambda_{s_j}(x_j, q_j) - q_j v_{n_j}(x_j)] \frac{\Delta t}{\epsilon} \right) R_{n_{j+1}}^{(s_j)}(x_j, q_j) \xi_{n_j}^{(s_j)}(x_j, q_j)$$

we find that

$$P(u_1, \ldots, u_N | u_0, n_0)$$

$$= \left[\prod_{j=0}^{N-1} \int_{-\infty}^{\infty} \frac{dp_j}{2\pi} \right] \sum_{n_1,\ldots,n_{N-1}} \sum_{s_j} \exp\left(\left[\lambda_{s_j}(u_j, q_j) - i\epsilon p_j \frac{u_{j+1} - u_j}{\Delta t} \right] \frac{\Delta t}{\epsilon} \right)$$

$$\times \exp\left([i\epsilon p_j v_{n_j}(u_j) - q_j v_{n_j}(u_j)] \frac{\Delta t}{\epsilon} \right) R_{n_{j+1}}^{(s_j)}(u_j, q_j) \xi_{n_j}^{(s_j)}(u_j, q_j), \tag{3.7}$$

to leading order in $O(\Delta u, \Delta t)$. It is important to note that the total path integral is independent of the q_j, since performing the summations over s_j recovers the Kronecker deltas. Let us now introduce the probability density

$$P(u_N, n_N | u_0, n_0) = \left[\prod_{j=1}^{N-1} \int_{-\infty}^{\infty} du_j \right] P(u_1, \ldots, u_N, n_N | u_0, n_0). \tag{3.8}$$

Substituting for P using (3.2) and (3.7), leads to

$$
\begin{aligned}
&P(u_N, n_N | u_0, n_0) \\
&= \left[\prod_{j=1}^{N-1} \int_{-\infty}^{\infty} du_j \right] \left[\prod_{j=0}^{N-1} \int_{-\infty}^{\infty} \frac{dp_j}{2\pi} \right] \\
&\quad \times \sum_{n_1,\ldots,n_{N-1}} \sum_{s_0,\ldots,s_{N-1}} \left[\prod_{j=0}^{N-1} R_{n_{j+1}}^{(s_j)}(u_j, q_j) \xi_{n_j}^{(s_j)}(u_j, q_j) \right] \\
&\quad \times \exp\left(\sum_j \left[\lambda_{s_j}(u_j, q_j) - i\epsilon p_j \frac{u_{j+1} - u_j}{\Delta t} \right] \frac{\Delta t}{\epsilon} \right) \\
&\quad \times \exp\left(\left[i\epsilon p_j v_{n_j}(u_j) - q_j v_{n_j}(u_j) \right] \frac{\Delta t}{\epsilon} \right).
\end{aligned}
$$

By inserting the eigenfunction products and using the Fourier representation of the Dirac delta function, we have introduced sums over the discrete labels s_j and new phase variables p_j. However, this allows us to obtain a simple action principle in the limit $\epsilon \to 0$. Since the path integral is ultimately independent of the q_j, we are free to set $q_j = i\epsilon p_j$ for all j, thus eliminating the final exponential factor. (Fixing the q_j is analogous to gauge-fixing in field theory.) This choice means that we can perform the summations with respect to the intermediate discrete states n_j using the orthogonality relation

$$
\sum_n R_n^{(s)}(u_j, q_{j-1}) \xi_n^{(s')}(u_{j+1}, q_j) = \delta_{s,s'} + O(\Delta u, \Delta q).
$$

We thus obtain the result that $s_j = s$ for all j, which means that we can then take the continuum limit of (3.9) to obtain the following path integral from $u(0) = u_0$ to $u(\tau) = u$ (after performing the change of variables $i\epsilon p_j \to p_j$, that is, performing a contour deformation in the complex p-plane):

$$
\begin{aligned}
P(u, n, \tau | u_0, n_0, 0) = \sum_s \int_{u(0)=u_0}^{u(\tau)=u} \exp\left(-\frac{1}{\epsilon} \int_0^\tau \left[p\dot{u} - \lambda_s(u, p) \right] dt \right) \\
\times R_n^{(s)}(u, p(\tau)) \xi_{n_0}^{(s)}(u_0, p(0)) \mathcal{D}[p]\mathcal{D}[u]. \quad (3.9)
\end{aligned}
$$

Applying the Perron–Frobenius theorem to the linear operator on the left-hand side of (3.4) for fixed u and q, shows that there exists a real, simple Perron eigenvalue. We assume that the eigenvalues are ordered such that $\lambda_0 > \operatorname{Re}(\lambda_1) \geq \operatorname{Re}(\lambda_2) \ldots$ with λ_0 the Perron eigenvalue. Since λ_0 is the only eigenvalue with a positive eigenfunction, we require on physical grounds that the initial and final states are only non-vanishing for $s = 0$. It follows that the sum over s in (3.9) projects to the single term $s = 0$. Also note that the factor $R_n^{(0)}(u, p(\tau)) \xi_{n_0}^{(0)}(u_0, p(0))$ in (3.9) essentially projects on to stochastic trajectories that start in the discrete state n_0 and terminate in the discrete

state n. We will ignore any restrictions on these discrete states and simply consider the probability density (for fixed $u(0) = u_0$)

$$P(u, t) = \int_{\mathbf{u}(0)=\mathbf{u}_0}^{\mathbf{u}(\tau)=\mathbf{u}} e^{-S[u,p]/\epsilon} \mathcal{D}[p]\mathcal{D}[u], \qquad (3.10)$$

with the action

$$S[u, p] = \int_0^\tau \left[p\dot{u} - \lambda_0(u, p) \right] dt, \qquad (3.11)$$

and λ_0 the Perron eigenvalue of the linear equation

$$\sum_m \left[W_{nm}(u) + p\delta_{n,m} v_m(u) \right] R_m^{(0)}(u, p) = \lambda_0(u, p) R_n^{(0)}(u, p). \qquad (3.12)$$

Formally comparing to classical mechanics, we have a path integral in a phase space (u, p) consisting of a dynamical variable $u(t)$ and its 'conjugate momentum' $p(t)$ with the Perron eigenvalue $\lambda_0(u, p)$ interpreted as a Hamiltonian. This underlying Hamiltonian structure of a stochastic hybrid system has also been identified using large-deviation theory [34, 41]; see below.

3.2 Large-Deviation Principles

It is important to point out that the formal derivation of the path integral (3.10), see also [33], involves a few steps that have not been justified rigorously. First, we 'gauge fix' the path integral by setting $q_j = \epsilon p_j$ with p_j pure imaginary. However, when we carry out steepest descents, we assume that the dominant contribution to the path integral in the complex p-plane occurs for real p_j. (There is an assumption as regards analytic continuation.) This then allows us to apply the Perron–Frobenius theorem to the linear operator of the eigenvalue equation. Second, we have not established that the discrete path integral converges to a well-defined functional measure in the continuum limit. Nevertheless, it turns out that the resulting action $S[u, p]$ is identical to one obtained using large-deviation theory [41–43]. This connection has recently been established by Bressloff and Faugeras [34]. We briefly summarize the main results here.

Following [34], we take as our starting point a Lagrangian large-deviation principle of Faggionato et al. [42, 43], which applies to a wide class of stochastic hybrid systems. Here we state the LDP for the particular one-population neural model. Let $\mathcal{M}_+([0, T])$ denote the space of non-negative finite measures on the interval $[0, T]$ and take the K-dim. vector $\{\psi(t)\}_{t\in[0,T]}$ to be an element of the product space $\mathcal{M}_+([0, T])^\Gamma$ where $\Gamma = \{0, \ldots, \mathcal{N}\}$ and $K = \mathcal{N} + 1$. In other words, for each $t \in [0, T]$, $\psi(t) = (\psi_1(t), \ldots, \psi_K(t))$ such that

$$\psi_n(t) \geq 0, \quad \sum_{n\in\Gamma} \psi_n(t) = 1.$$

A particular realization of the stochastic process, $\{(x(t), n(t))\}_{t \in [0,T]}$, then lies in the product space $C([0, T]) \times \mathcal{M}_+([0, T])^\Gamma$ with

$$\psi_n(t) = 1_{\{n(t)=n\}} \equiv \begin{cases} 1, & \text{if } n(t) = n, \\ 0, & \text{if } n(t) \neq n, \end{cases} \tag{3.13}$$

and

$$u(t) = u_0 + \int_0^t \sum_{n \in \Gamma} \psi_n(s) v_n\big(u(s)\big) \, ds. \tag{3.14}$$

Let $\mathcal{Y}_{u_0}$ denote the subspace of $C([0, T]) \times \mathcal{M}_+([0, T])^\Gamma$ for which (3.14) holds but ψ is now a general element of $\mathcal{M}_+([0, T])^\Gamma$. Such a space contains the set of trajectories of the stochastic hybrid system with $\psi_n(t)$ given by (3.13) and $n(t)$ evolving according to the Markov chain. Finally, take P_{u_0,n_0}^ϵ to be the probability density functional or law of the set of trajectories in $\mathcal{Y}_{u_0}$. The following large-deviation principle then holds [42, 43]:

For any $(u, \psi) \in C([0, T]) \times [0, 1]^\Gamma$ define

$$j(u, \psi) = \sup_{z \in (0,\infty)^\Gamma} \sum_{(n,n') \in \Gamma \times \Gamma} \psi_n W_{n'n}(u) \left[1 - \frac{z_{n'}}{z_n} \right]. \tag{3.15}$$

Then, for any given path $\{(u(t), \psi(t))\}_{t \in [0,T]} \in \mathcal{Y}_{u_0}$,

$$\mathbb{P}_{u_0,n_0}^\epsilon \Big[\big\{ \big(u(t), \psi(t)\big) \big\}_{t \in [0,T]} \Big] \sim e^{-J_T(\{(u(t),\psi(t))\}_{t\in[0,T]})/\epsilon}, \tag{3.16}$$

where the rate function $J_T : \mathcal{Y}_{u_0} \to [0, \infty)$ is given by

$$J_T\Big(\big\{\big(u(t), \psi(t)\big)\big\}_{t \in [0,T]}\Big) = \int_0^T j\big(u(t), \psi(t)\big) \, dt. \tag{3.17}$$

Here the symbol $\sim$ means asymptotic logarithmic equivalence in the limit $\epsilon \to 0$.

A key idea behind the LDP is that a slow dynamical process coupled to the fast Markov chain on Γ rapidly samples the different discrete states of Γ according to some non-negative measure ψ. In the limit $\epsilon \to 0$, one has $\psi \to \rho$, where ρ is the ergodic measure of the Markov chain. On the other hand, for small but non-zero ϵ, ψ is itself distributed according to the LDP (3.16), whereby one averages the different functions $v_n(x)$ over the measure ψ to determine the dynamics of the slow system. In our population model, we are interested in the synaptic current u (for a current-based or voltage-based model). Eliminating $\psi(t)$ using a contraction principle then leads to the following LDP for $\{u(t)\}_{t \in [0,T]}$ alone [41, 43]:

Given an element $\{u(t)\}_{t \in [0,T]} \in C([0, T])$, we have

$$\mathbb{P}_{u_0,n_0}^\epsilon \Big[\big\{ u(t) \big\}_{t \in [0,T]} \Big] \sim e^{-J_T(\{u(t)\}_{t\in[0,T]})/\epsilon},$$

where the rate function $J_{[0,T]} : C([0, T], \Omega) \to [0, \infty)$ is given by

$$J_T\Big(\big\{u(t)\big\}_{t \in [0,T]}\Big) = \inf_{\{\psi(t)\}_{t \in [0,T]} : \dot{u}(t) = \sum_n v_n(u)\psi_n} J_T\Big(\big\{\big(u(t), \psi(t)\big)\big\}_{t \in [0,T]}\Big). \tag{3.18}$$

Roughly speaking, one can understand the contraction principle in terms of steepest descents. That is

$$\mathbb{P}^{\epsilon}_{x_0,n_0}\big[\{u(t)\}_{t\in[0,T]}\big] = \int \mathbb{P}^{\epsilon}_{u_0,j_0}\big[\{(u(t),\psi(t))\}_{t\in[0,T]}\big]D[\psi]$$

$$\sim \int e^{-J_T(\{(u(t),\psi(t))\}_{t\in[0,T]})/\epsilon}\,D[\psi],$$

where $D[\psi]$ is the appropriate functional measure on $\mathcal{M}_+([0,T])^{\Gamma}$. The path integral is then dominated by the infimum of the rate function in the limit $\epsilon \to 0$.

In [34], it is proven that the rate function (3.18) can be written in the form of an action

$$J_T\big(\{u(t)\}_{t\in[0,T]}\big) = \int_0^T L(u,\dot{u})\,dt, \tag{3.19}$$

with Lagrangian given by

$$L(u,\dot{u}) = \mu(u,\dot{u})\dot{u} - \lambda_0\big(u(t),\mu(u,\dot{u})\big), \tag{3.20}$$

where $\lambda_0(x,\mu)$ is the Perron eigenvalue of the linear equation

$$\sum_m W_{nm}(u)R_m^{(0)} + \mu v_n(u)R_n^{(0)} = \lambda_0 R_n^{(0)},$$

and $\mu = \mu(u,\dot{u})$ the solution of the equation

$$\dot{u} = \frac{\partial\lambda_0}{\partial\mu} \equiv \sum_m R_m^{(0)}(u,\mu)\xi_m^{(0)}(u,\mu)v_m(u) \tag{3.21}$$

with $\xi^{(0)}$ the adjoint eigenvector of $R^{(0)}$. Note that μ is a Lagrange multiplier which is introduced in order to impose the constraint $\dot{u} = \sum_m v_n(u)\psi_n$ when evaluating the infimum of (3.18). Given the Lagrangian L, we can determine a corresponding Hamiltonian H according to the Fenchel–Legendre transformation

$$H(u,p) = \sup_y\big[(p-\mu(u,y))y + \lambda\big(u,\mu(u,y)\big)\big]. \tag{3.22}$$

Minimizing the right-hand side yields the equation

$$p - \mu(u,y) + \left[\frac{\partial\lambda}{\partial\mu} - y\right]\frac{\partial\mu}{\partial y} = 0. \tag{3.23}$$

Since $\partial_\mu\lambda = y$, we see that $p = \mu$ i.e., we can identify the Lagrange multiplier μ in the construction of the Lagrangian as the conjugate momentum p of the Hamiltonian

$$H = \lambda_0(u,p), \tag{3.24}$$

where $\lambda_0(u,p)$ is the Perron eigenvalue of the linear equation (3.12). It follows that the action obtained from a large-deviation principle is identical to the action (3.11) derived using formal path-integral methods.

3.3 Calculation of Perron Eigenvalue

In our previous work [32, 33], we obtained an explicit solution for the Perron eigenvalue and the associated positive eigenvector by taking the number of discrete states in each population to be infinite, that is, $\mathcal{N} \to \infty$. However, the classical Perron–Frobenius theorem applies to finite-dimensional Markov processes. One consequence of this is that the Perron eigenvalue develops singularities in the thermodynamic limit. In order to explore this issue, let us return to the one-population eigenvalue equation (3.12), which takes the explicit form

$$0 = F(u)R_{n-1}^{(0)}(u, p) - \left[\lambda_0 + F(u) + pu + n(1 - wp)\right]R_n^{(0)}(u, p)$$
$$+ (n + 1)R_{n+1}^{(0)}(u, p). \tag{3.25}$$

In the infinite-dimensional case, one can formally solve this equation using the trial positive solution

$$R_n^{(0)}(u, p) = \frac{\Lambda(u, p)^n}{n!}, \quad \Lambda(u) = \lambda_0 + F(u) + pu. \tag{3.26}$$

This yields the following equation relating Λ and p:

$$\left[\frac{F(u)}{\Lambda} - 1\right]n + \Lambda - F(u) - \lambda_0 = -p(-u + wn).$$

We now collect terms independent of n and linear in n, respectively, to obtain the pair of equations

$$p = -\frac{1}{w}\left[\frac{F(u)}{\Lambda} - 1\right], \quad \Lambda = F(u) + pu + \lambda_0.$$

It follows that

$$\Lambda = \frac{F(u)}{1 - wp}, \quad \lambda_0 = pw\frac{F(u)}{1 - wp} - pu. \tag{3.27}$$

There is clearly a singularity at $p = 1/w$ such that $\Lambda(u, p) < 0$ for $p > 1/w$, contradicting the requirement that the eigenfunction $R_n^{(0)}$ is positive.

The origin of the singularity can be understood by considering a large, but finite population size $\mathcal{N}$. The Perron–Frobenius theorem then holds but the solution of the eigenvalue equation becomes non-trivial. The basic difficulty arises because the above ansatz for $R_n^{(0)}$ does not satisfy the boundary condition at $n = \mathcal{N}$. That is, setting $n = \mathcal{N} - 1$ and $n = \mathcal{N}$ in (3.25) with $R_{\mathcal{N}+1}^{(0)} = 0$ gives

$$F(u)R_{\mathcal{N}-2}^{(0)}(u, p) - \left[\Lambda(u) + (\mathcal{N} - 1)(1 - wp)\right]R_{\mathcal{N}-1}^{(0)}(u, p) + \mathcal{N}R_{\mathcal{N}}^{(0)}(u, p) = 0$$

and

$$F(u)R_{\mathcal{N}-1}^{(0)}(u, p) - \left[\Lambda(u) - F(u) + \mathcal{N}(1 - wp)\right]R_{\mathcal{N}}^{(0)}(u, p) = 0.$$

Assuming that $R_n^{(0)} = \Lambda^n/n!$ for $0 \le n < \mathcal{N}$, with Λ given by (3.26) and $p < 1/w$ (positive solution), we see that the first equation is satisfied by taking $R_{\mathcal{N}}^{(0)} = \Lambda^{\mathcal{N}}/\mathcal{N}!$. However, the second equation requires

$$R_{\mathcal{N}}^{(0)} = \frac{F(u)}{\Lambda(u) - F(u) + \mathcal{N}(1 - wp)} R_{\mathcal{N}-1}^{(0)}(u).$$

In the large-N limit with $p < 1/w$, we set

$$R_{\mathcal{N}}^{(0)} \to \frac{F(u)}{\mathcal{N}(1 - wp)} R_{\mathcal{N}-1}^{(0)}(u) = \frac{\Lambda(u)}{\mathcal{N}} R_{\mathcal{N}-1}^{(0)}(u).$$

This shows that the given ansatz is a good approximation to the eigensolution for large $\mathcal{N}$ and $p < 1/w$. Clearly, the given ansatz breaks down as p crosses $p = 1/w$. Although the Perron–Frobenius theorem guarantees a unique positive solution for finite $\mathcal{N}$, it does not have a simple expression in the large N limit. In conclusion, our expression (3.27) for the Perron eigenvalue only holds for $p < 1/w$. This does not affect our subsequent analysis because we evaluate the path integral in regions for which $p < 1/w$.

3.4 Multi-population Model

Following along identical lines to the one-population model, we can derive a path-integral representation of the solution of the multi-population CK equation (2.6):

$$p(\mathbf{u}, \tau) = \int_{\mathbf{u}(0)=\mathbf{u}_0}^{\mathbf{u}(\tau)=\mathbf{u}} \mathcal{D}[\mathbf{p}]\mathcal{D}[\mathbf{u}] \exp\left(-\frac{1}{\epsilon} S[\mathbf{u}, \mathbf{p}]\right) \qquad (3.28)$$

with the action

$$S[\mathbf{u}, \mathbf{p}] = \int_0^\tau \left[\sum_{\alpha=1}^M p_\alpha \dot{u}_\alpha - \lambda_0(\mathbf{u}, \mathbf{p})\right] dt. \qquad (3.29)$$

Here λ_0 is the Perron eigenvalue of the following linear operator equation (cf. (3.4)):

$$\sum_{\mathbf{m}} W(\mathbf{n}, \mathbf{m}; \mathbf{u}) R^{(0)}(\mathbf{u}, \mathbf{p}, \mathbf{m})$$

$$= \left[\lambda_0(\mathbf{u}, \mathbf{p}) - \sum_{\alpha=1}^M p_\alpha v_\alpha(\mathbf{u}, \mathbf{n})\right] R^{(0)}(\mathbf{u}, \mathbf{p}, \mathbf{n}), \qquad (3.30)$$

and $\xi^{(0)}$ is the corresponding adjoint eigenvector. For sufficiently small p_αs, (3.30) can be solved for the Perron eigenvalue in the thermodynamic limit $\mathcal{N} \to \infty$ using the ansatz

$$R^{(0)}(\mathbf{u}, \mathbf{p}, \mathbf{n}) = \prod_{\alpha=1}^M \frac{\Lambda_\alpha(\mathbf{u}, \mathbf{p})^{n_\alpha}}{n_\alpha!}. \qquad (3.31)$$

Substituting into (3.30) and using the explicit expressions for W and v_α, we find that

$$\sum_{\alpha=1}^{M}\left(\left[\frac{F(u_\alpha)}{\Lambda_\alpha}-1\right]n_\alpha + \Lambda_\alpha - F(u_\alpha)\right) - \lambda_0$$

$$= -\sum_{\alpha=1}^{M} p_\alpha\left[-u_\alpha + \sum_{\beta} w_{\alpha\beta} n_\beta\right]. \tag{3.32}$$

Collecting terms in n_α for each α yields

$$\frac{F(u_\alpha)}{\Lambda_\alpha} - 1 = -\sum_{\beta=1}^{M} p_\beta w_{\beta\alpha}, \tag{3.33}$$

and collecting terms independent of all n_α gives

$$\lambda_0 = \sum_{\alpha=1}^{M}\left[\Lambda_\alpha - F(u_\alpha) - u_\alpha p_\alpha\right]. \tag{3.34}$$

Solving for each Λ_α in terms of $\mathbf{p}$, we have

$$\lambda_0(\mathbf{u},\mathbf{p}) \equiv \sum_{\alpha=1}^{M}\left[\frac{F(u_\alpha)}{1-\sum_{\beta=1}^{M} p_\beta w_{\beta\alpha}} - u_\alpha p_\alpha - F(u_\alpha)\right]. \tag{3.35}$$
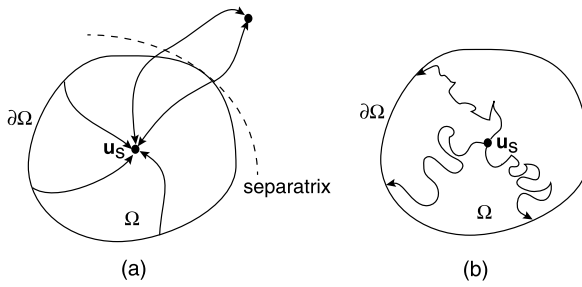
As in the one-population model, the Perron eigenvalue has singularities, reflecting the possible breakdown of the Perron–Frobenius theorem in the thermodynamic limit.

4 A Variational Principle and Optimal Paths of Escape

It is clear from the formal structure of the path integral (3.28) that each synaptic variable u_α has a 'conjugate momentum' p_α with $\lambda_0(\mathbf{u},\mathbf{p})$ the corresponding 'Hamiltonian' H. Applying steepest descents to the path integral for small ϵ yields a variational principle in which maximum-likelihood paths minimize the action (3.29). As is well known from classical mechanics, the least action principle leads to Hamilton's equations

$$\dot{\mathbf{u}} = \nabla_{\mathbf{p}} H(\mathbf{u},\mathbf{p}), \qquad \dot{\mathbf{p}} = -\nabla_{\mathbf{u}} H(\mathbf{u},\mathbf{p}), \tag{4.1}$$

describing a 'classical particle' moving in the phase space $(\mathbf{u},\mathbf{p})$. What is the physical interpretation of the solutions to Hamilton's equations? In order to address this question, suppose that the underlying deterministic mean-field equation (2.19) has a stable fixed point $\mathbf{u}_s$ with some basin of attraction Ω, as illustrated in Fig. 1. If the system starts within Ω, then on relatively short time scales we expect the system to rapidly converge to $\mathbf{u}_s$ along a classical deterministic trajectory, with noise generating Gaussian-like fluctuations about this trajectory. However, on a longer time scale, a rare event (large fluctuation) will generate a path of escape from $\mathbf{u}_s$ to the boundary

Fig. 1 **a** Deterministic trajectories of a multistable dynamical system. The subset Ω is contained within the basin of attraction of a fixed point $\mathbf{u}_s$. The boundary of the basin of attraction consists of separatrices, which are also solution trajectories. Paths that start in a neighborhood of a separatrix are attracted by different fixed points, depending whether they begin on the left or right of the separatrix. **b** Random trajectories of the stochastic system. Escape from the domain Ω occurs when a random trajectory hits the boundary $\partial\Omega$

of Ω. It turns out that both classical trajectories and the maximum-likelihood paths of escape correspond to zero energy solutions of Hamilton's equations of motion; this follows from the fact that the action vanishes at fixed points of the deterministic mean-field equation. We will illustrate this by considering the simpler one-population model.

Setting $\lambda_0 = 0$ in the eigenvalue equation (3.12) gives

$$\sum \left[W_{nm}(u) + p\delta_{n,m}v_m(u) \right] R_m^{(0)}(u, p) = 0 \tag{4.2}$$

with $R_m^{(0)}$ required to be a positive function. One solution is $p = 0$ and $R_m^{(0)}(u, 0) = \rho_m(u)$ with

$$\sum_m W_{nm}(u)\rho_m(u) = 0.$$

Differentiating the eigenvalue equation with respect to p and then setting $p = 0$, $\lambda_0 = 0$ shows that

$$\left. \frac{\partial \lambda_0(u, p)}{\partial p} \right|_{p=0} \rho_n(u) = v_n(u)\rho_n(u) + \sum_m W_{nm}(u) \left. \frac{\partial R_m^{(0)}(u, p)}{\partial p} \right|_{p=0}.$$

Summing both sides with respect to n and using $\sum_n W_{nm} = 0$,

$$\left. \frac{\partial \lambda_0(u, p)}{\partial p} \right|_{p=0} = \sum_n v_n(u)\rho_n(u).$$

Similarly, one finds that $\partial \lambda_0(u, p)/\partial u$ vanishes at $p = 0$. Hence, Hamilton's equations

$$\dot{u} = \frac{\partial \lambda_0(u, p)}{\partial p}, \qquad \dot{p} = -\frac{\partial \lambda_0(u, p)}{\partial u}$$

reduce to

$$\dot{u} = \sum_n v_n(u)\rho_n(u), \qquad \dot{p} = 0 \quad \text{for } p = 0.$$

It follows that $(u_s, 0)$ is a fixed point in the full phase space with the line $p = 0$ a stable manifold. Along this manifold, u converges to u_s according to the scalar version of the mean-field equation (2.19). From the explicit expression for $\lambda_0(u, p)$, (3.27), we see that there exists another zero energy solution given by

$$p = \phi(u) \equiv \frac{1}{w} - \frac{F(u)}{u},$$

which is the unique, non-trivial solution of the equation

$$\sum_m \big[W_{nm}(u) + \phi(u)\delta_{n,m} v_m(u) \big] \psi_m(u) = 0 \tag{4.3}$$

for positive functions $\psi_m(u)$. (Note that $p < 1/w$ so that we do not have to worry about the singular nature of the Perron eigenvalue in the limit $\mathcal{N} \to \infty$.) It corresponds to the trajectory along the unstable manifold of $(u_s, 0)$ and is the optimal path of escape from u_s. Along this optimal path $\lambda_0 = 0$, so that the corresponding action is given by the quasipotential

$$\Phi(u) = \int_{-\infty}^t \big[p\dot{u} - \lambda_0(u, p) \big] dt \bigg|_{p=\phi(u)} = \int_{-\infty}^t \phi(u)\dot{u}\, dt = \int_{u_s}^u \phi(y)\, dy \tag{4.4}$$

and

$$P \sim \mathrm{e}^{-\Phi(u)/\epsilon}. \tag{4.5}$$

A similar situation holds for the higher-dimensional case, except that there are now multiple maximum-likelihood paths of escape from a metastable state [27, 33].

5 The Diffusion Approximation and Neural Pattern Formation

Another useful application of the multi-population path integral (3.28) is that it provides a direct method for obtaining a Gaussian or diffusion approximation of the stochastic hybrid system, equivalent to the one obtained using the more complicated QSS reduction [27]. Performing the rescaling $\mathbf{p} \to i\mathbf{p}/\epsilon$ gives

$$P(\mathbf{u}, t) = \int_{\mathbf{u}(0)=\mathbf{u}_0}^{\mathbf{u}(\tau)=\mathbf{u}} D[\mathbf{u}] D[\mathbf{p}]$$
$$\times \exp\left(-\int_0^\tau i \sum_\alpha p_\alpha \left[\dot{u}_\alpha + u_\alpha - \sum_\beta \frac{w_{\alpha\beta} F(u_\beta)}{1 - i\epsilon \sum_\gamma w_{\gamma\beta} p_\gamma} \right] dt \right).$$

The Gaussian approximation involves Taylor expanding the Lagrangian to first order in ϵ, which yields a quadratic in p:

$$P(u,t) = \int_{\mathbf{u}(0)=\mathbf{u}_0}^{\mathbf{u}(\tau)=\mathbf{u}} D[\mathbf{u}]D[\mathbf{p}]\exp\left(\int_0^\tau i\sum_\alpha p_\alpha\left(\dot{u}_\alpha + u_\alpha - \sum_\beta w_{\alpha\beta}F(u_\beta)\right)\right)$$

$$\times \exp\left(-i\epsilon\int_0^\tau \sum_{\alpha,\gamma} p_\alpha \mathcal{Q}_{\alpha\gamma}(\mathbf{u})p_\gamma \, dt\right),$$

where $\mathcal{Q}_{\alpha\gamma}(\mathbf{u}) = \sum_\beta w_{\alpha\beta}F(u_\beta)w_{\gamma\beta}$. Performing the Gaussian integration then yields

$$P(\mathbf{u},t) = \int D[\mathbf{u}]e^{-\mathcal{A}[\mathbf{u}]/\epsilon},$$

with action functional

$$\mathcal{A}[\mathbf{u}] = \frac{1}{4}\int_0^\tau \sum_{\alpha,\beta}\left(\dot{u}_\alpha(t) - V_\alpha(\mathbf{u}(t))\right)\mathcal{Q}_{\alpha\beta}^{-1}(\mathbf{u})\left(\dot{u}_\beta(t) - V_\beta(\mathbf{u}(t))\right)dt, \qquad (5.1)$$

where $V_\alpha(\mathbf{u}) = -u_\alpha + \sum_\beta w_{\alpha\beta}F(u_\beta)$. This path integral is identical in form to the Onsager–Machlup path-integral representation [44] of solutions to the Langevin equation

$$dU_\alpha(t) = V_\alpha(\mathbf{U})\,dt + \sqrt{2\epsilon}\sum_\beta w_{\alpha\beta}\sqrt{F(u_\beta)}\,dW_\beta(t), \qquad (5.2)$$

where the $W_\alpha(t)$ are independent Wiener processes. Since there is no additional Jacobian factor in the Onsager–Machlup path integral, it follows that the Langevin equation is of the Ito form. As we have discussed extensively elsewhere [27, 33], the diffusion or Gaussian approximation breaks down when solving escape problems. On the other hand, it provides useful information when analyzing the effects of fluctuations within the basin of attraction of a metastable state. For example, it is well known within the context of PDEs that fluctuations can enlarge the parameter regime over which time-periodic (limit cycles) or spatially periodic (Turing patterns) can occur. A similar phenomenon exists for stochastic hybrid neural networks. We will illustrate this by considering Turing-like instabilities in a spatially structured hybrid neural network under the diffusion approximation.

5.1 Noise-Induced Pattern Formation

Consider a system of coupled homogeneous neural populations that are distributed on a regular d-dimensional lattice $\mathcal{L}$, with lattice spacing $\Delta\alpha$ and site index $\alpha \in \mathcal{L}$. Following recent studies of stochastic pattern formation in RD systems [45–51], we investigate the occurrence of stochastic neural patterns by linearizing the spatially discrete Langevin equation (5.2) about a homogeneous stationary solution u_0 of the mean-field equation (2.19) and calculating the resulting power spectrum using discrete Fourier transforms. In order to reflect the homogeneous structure of the weights

we also set

$$w_{\alpha\alpha'} = w(|\alpha - \alpha'|).$$

Substituting

$$U_\alpha(t) = u_0 + \sqrt{\epsilon}\,\Phi_\alpha(t)$$

into (5.2) and Taylor expanding to first order in Φ gives the multivariate Ornstein–Uhlenbeck process

$$\frac{d\Phi_\alpha}{dt} = \sum_{\alpha'} J_0(|\alpha - \alpha'|)\Phi_{\alpha'}(t) + \sum_{\alpha'} B_0(|\alpha - \alpha'|)\xi_{\alpha'}(t), \qquad (5.3)$$

with

$$J_0(|\alpha - \alpha'|) = \left.\frac{\partial V_\alpha}{\partial U_{\alpha'}}\right|_{u_0} = -\delta_{\alpha,\alpha'} + w(|\alpha - \alpha'|)F'(u_0), \qquad (5.4)$$

and

$$B_0(|\alpha - \alpha'|) = w(|\alpha - \alpha'|)\sqrt{2F(u_0)}. \qquad (5.5)$$

Considerable insight into the behavior of the system can now be obtained by transforming to Fourier space [45, 51]. For simplicity, consider a 1D lattice with periodic boundary conditions, $u_{\alpha+N} = u_\alpha$ for $\alpha = 1, \ldots, N$ and set the lattice spacing $\Delta\alpha = 1$. Introduce the discrete Fourier transforms

$$\widehat{\Phi}(k) = \sum_\alpha e^{-ik\alpha}\Phi_\alpha, \qquad \Phi_\alpha = \frac{1}{N}\sum_k e^{ik\alpha}\widehat{\Phi}(k)$$

with $k = 2\pi m/N$, $m = 0, \ldots, N-1$. Using the following result for convolutions:

$$\sum_{\alpha,\alpha'} e^{-ik\alpha} J(\alpha - \alpha')\Phi_{\alpha'} = \widehat{J}(k)\widehat{\Phi}(k),$$

the discrete Fourier transform of the Langevin equation is

$$\frac{d\widehat{\Phi}(k,t)}{dt} = \widehat{J}_0(k)\widehat{\Phi}(k,t) + \widehat{B}_0(k)\widehat{\xi}(k,t) \qquad (5.6)$$

with

$$\widehat{J}_0(k) = -1 + \widehat{w}(k)F'(u_0), \qquad \widehat{B}_0(k) = \widehat{w}(k)\sqrt{2F(u_0)} \qquad (5.7)$$

and $\langle\widehat{\xi}(k,t)\rangle = 0$,

$$\langle\widehat{\xi}(k,t)\widehat{\xi}(k',t')\rangle = \delta_{k,-k'}\delta(t-t').$$

Note that the homogeneous equation

$$\frac{d\widehat{\Phi}(k,t)}{dt} = \widehat{J}_0(k)\widehat{\Phi}(k,t) \qquad (5.8)$$

determines the stability of the fixed point u_0 in the absence of noise. It follows that the deterministic system is stable provided that $\widehat{J}_0(k) < 0$ for all k. Suppose that the gain $\mu = F'(u_0)$ is treated as a bifurcation parameter. Clearly if $\widehat{w}(k)$ is bounded and μ is sufficiently small, then $\widetilde{J}_0(k) < 0$ for all k. However, if $\max_k\{\widehat{w}(k)\} = \widehat{w}(k_c) > 0$ then the fixed point becomes marginally stable at $\mu = \mu_c = 1/\widehat{w}(k_c)$, resulting in the growth of a spatially periodic pattern of wavenumber k_c as μ crosses μ_c. A standard neural mechanism for inducing a Turing-like instability is to have a combination of short-range excitation and longer-range inhibition [52, 53]. This can implemented in the 1D scalar model by taking w to be the difference-of-Gaussians

$$w_D(x) = e^{-x^2/2} - Ae^{-x^2/2\sigma^2}. \tag{5.9}$$

(More precisely, in order to match the periodic boundary conditions, we should take $w(\alpha) = \sum_n w_D(\alpha - nN)$.)

Spectral theory can now be used to determine the effects of noise on pattern formation. First, Fourier transforming the Langevin equation (5.6) with respect to time gives

$$\Lambda(k, \Omega)\widehat{\Phi}(k, \Omega) = \widehat{B}_0(k)\widehat{\xi}(k, \Omega)$$

with

$$\Lambda(k, \Omega) = -i\Omega - \widehat{J}_0(k)$$

and

$$\langle\widehat{\xi}(k, \Omega)\rangle = 0, \qquad \langle\widehat{\xi}(k, \Omega)\widehat{\xi}(k', \Omega')\rangle = \delta_{k,-k'}\delta(\Omega + \Omega').$$

It follows that

$$\langle\widehat{\Phi}(k, \Omega)\widehat{\Phi}(k', \Omega')\rangle = \left\langle\left[\frac{\widehat{B}_0(k)}{\Lambda(k, \Omega)}\widehat{\xi}(k, \Omega)\right]\left[\frac{\widehat{B}_0(k')}{\Lambda(k', \Omega')}\widehat{\xi}(k', \Omega')\right]\right\rangle$$

$$= \delta_{k,-k'}\delta(\Omega + \Omega')\frac{|B_0(k)|^2}{|\Lambda(k, \Omega)|^2}.$$

Defining the power spectrum by

$$\langle\widehat{\Phi}(k, \Omega)\widehat{\Phi}(k', \Omega')\rangle = S(k, \Omega)\delta_{k,-k'}\delta(\Omega + \Omega'),$$
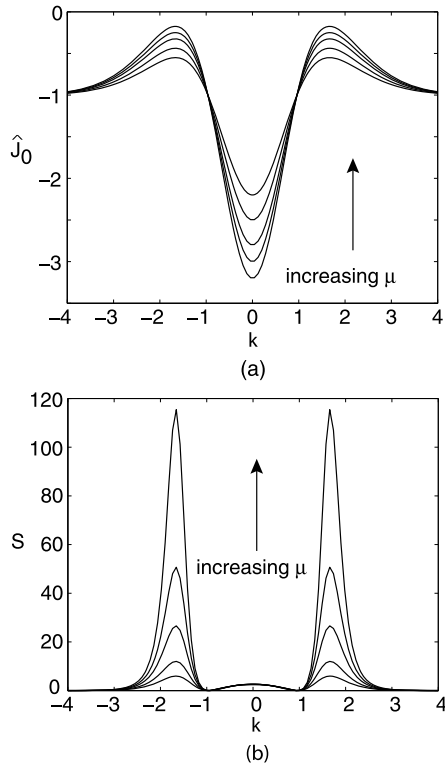
we deduce that

$$S(k, \Omega) = \frac{|B_0(k)|^2}{|\Lambda(k, \Omega)|^2}. \tag{5.10}$$

From the deterministic theory, we know that the system undergoes a Turing instability (stationary patterns) rather than a Turing–Hopf instability (oscillatory patterns) so we can set $\Omega = 0$ and determine conditions under which $S(k, 0)$ has a peak at a non-zero, finite value of k, which is an indication of a stochastic pattern. Substituting the explicit expression for $\Lambda(k, 0)$ and $B_0(k)$, we have

$$S(k, 0) = \frac{2\widehat{w}(k)^2 F(u_0)}{[-1 + \widehat{w}(k)F'(u_0)]^2} = \frac{2F(u_0)}{F'(u_0)^2}\left[1 + \lambda(k)^{-1}\right]^2. \tag{5.11}$$

Fig. 2 *Stochastic pattern formation in a scalar neural network.* **a** Plot of Fourier transformed weight distribution as a function of wavenumber k for various values of gain $\mu = f'(u_0)$: $\mu = 0.12, 0.15, 0.18, 0.2, 0.22$. **b** Sketch of corresponding power spectra $S(k, 0)$, showing the peak in the spectrum at the critical wavenumber k_c for $\mu < \mu_c$. Parameter values are $\sigma^2 = 4$, $A = 0.5$, $F_0 = 20$, $\eta = 2$, $\kappa = 0$

Suppose that $\mu \equiv F'(u_0) < \mu_c$ so the system is below the deterministic critical point for a Turing instability. Clearly $S(k, 0)$ becomes singular as $\mu \to \mu_c$, consistent with the fixed point becoming unstable. The main new result is that $S(k, 0)$ has a peak at the critical wavenumber k_c for all μ, $0 < \mu < \mu_c = \widehat{w}(k_c)^{-1}$. This follows from the fact that $\lambda(k) < 0$ for all k in the subcritical regime with $\min_k\{|\lambda(k)|\} = |\lambda(k_c)|$. Hence, $S(k, 0)$ will have a peak at $k = k_c$ provided that

$$0 < \left|\lambda(k_c)\right| \equiv 1 - \mu\widehat{w}(k_c) < 1 \quad \Longrightarrow \quad \mu < \mu_c. \tag{5.12}$$

This is illustrated in Fig. 2.

5.2 Continuum Limit

The above stochastic model of a spatially structured lattice of neural populations can be reduced to a stochastic neural field by taking a continuum limit. A heuristic derivation proceeds as follows. Suppose that there is a uniform density ρ of populations distributed in $\mathbb{R}^d$. We then reinterpret u_α as the mean current averaged over the $\rho\Delta\alpha^d$ populations in the infinitesimal volume $\Delta\alpha^d$ centered at the lattice point $\alpha \in \mathcal{L}$. If an individual population in the set of populations centered at α is labeled by the pair (α, j), then

$$u_\alpha = \frac{\sum_j u_{\alpha,j}}{\rho\Delta\alpha^d}.$$

We will assume that the weights are slowly varying on the length-scale $\Delta\alpha$ so that $w_{\alpha j, \alpha' j'} = w_{\alpha\alpha'}$. (Relaxing this assumption can lead to additional sources of stochasticity as explored in [12, 14].) The deterministic mean-field (2.19) for an individual population becomes

$$\tau \frac{du_{\alpha j}}{dt} = -u_{\alpha j}(t) + \sum_{\alpha'} w_{\alpha\alpha'} \sum_{j'} F(u_{\alpha' j'}).$$

Averaging with respect to j gives

$$\tau \frac{du_{\alpha}}{dt} = -u_{\alpha}(t) + \rho\Delta\alpha^d \sum_{\alpha'} w_{\alpha\alpha'} \frac{1}{\rho\Delta\alpha^d} \sum_{j'} F(u_{\alpha' j'})$$

$$= -u_{\alpha}(t) + \rho\Delta\alpha^d \sum_{\alpha'} w_{\alpha\alpha'} F(u_{\alpha})$$

under the approximation that all local populations are in a similar state so

$$\frac{1}{\rho\Delta\alpha^d} \sum_{j'} F(u_{\alpha', j'}) \approx F\left(\frac{1}{\rho\Delta\alpha^d} \sum_{j'} u_{\alpha', j'} \right).$$

Effectively, we are scaling the population firing-rate function by a factor $\rho\Delta\alpha^d$. Finally, setting $\alpha = \mathbf{x}$, $u_{\alpha}(t) = u(\mathbf{x}, t)$, $\rho w_{\alpha\alpha'} = w(\mathbf{x}, \mathbf{x}')$, and taking the continuum limit $\Delta\alpha \to 0$ yields the deterministic neural field equation

$$\tau \frac{\partial u}{\partial t} = -u(\mathbf{x}, t) + \int_{\mathbb{R}^d} w(\mathbf{x}, \mathbf{x}') F(u(\mathbf{x}', t)) \, d\mathbf{x}'. \tag{5.13}$$

Applying a similar analysis to the diffusion matrix, we have

$$Q_{\alpha j, \alpha' j'} = \sum_{\alpha_1, j_1} w_{\alpha\alpha_1} w_{\alpha'\alpha_2} F(u_{\alpha_1, j_1}) = \rho\Delta\alpha^d \sum_{\alpha_1} w_{\alpha\alpha_1} w_{\alpha'\alpha_2} F(u_{\alpha_1}).$$

Hence, $\mathbf{Q}$ is independent of the local population labels j, j' and the Langevin equation (5.2) becomes

$$\frac{dU_{\alpha j}}{dt} = \overline{V}_{\alpha j}(\mathbf{U}) + \sqrt{2\epsilon\rho\Delta\alpha^d} \sum_{\alpha'} B_{\alpha\alpha'}(\mathbf{U}) \xi_{\alpha'}(t). \tag{5.14}$$

Averaging with respect to j and taking the continuum limit yields the following neural field model with spatiotemporal Gaussian white noise:

$$\frac{dU(\mathbf{x}, t)}{dt} = -U(\mathbf{x}, t) + \int_{\mathbb{R}^d} w(\mathbf{x}, \mathbf{x}') F(U(\mathbf{x}', t)) \, d\mathbf{x}'$$

$$+ \sqrt{\epsilon} \int_{\mathbb{R}^d} B(\mathbf{x}, \mathbf{x}') \xi(\mathbf{x}', t) \, d\mathbf{x}', \tag{5.15}$$

where

$$B(\mathbf{x}, \mathbf{x}') = w(\mathbf{x}, \mathbf{x}')\sqrt{2F(U(\mathbf{x}', t))} \tag{5.16}$$

and $\langle \eta(\mathbf{x}, t) \rangle = 0$,

$$\langle \eta(\mathbf{x}, t)\eta(\mathbf{x}', t') \rangle = \delta(\mathbf{x} - \mathbf{x}')\delta(t - t'). \tag{5.17}$$

For finite $\Delta \alpha$, we have introduced the scaling $\eta_\alpha(t)/\sqrt{\Delta \alpha^d} = \eta(\mathbf{x}, t)$.

From a numerical perspective, any computer simulation would involve rediscretizing space and then solving a time-discretized version of the resulting stochastic neural field equation. On the other hand, in order to investigate analytically the effects of noise on spatiotemporal dynamics such as traveling waves, it is more useful to work directly with stochastic neural fields. One can then adapt various PDE methods for studying noise in spatially extended systems [15, 54–58]. Finally, note that a large-deviation principle for a stochastic neural field with additive noise has been developed in [59].

6 Generating Functionals and the $1/\epsilon$ Loop Expansion

One step beyond the Gaussian approximation is to consider corrections to the mean-field equation (2.19), which couple the mean synaptic current with higher-order moments. As demonstrated previously for neural master equations [17, 18, 20], path integrals provide a systematic method for generating the hierarchy of moment equations. We will illustrate this by calculating the lowest-order correction to mean-field theory based on coupling to second-order correlations. One could then take investigate the bifurcation structure of the higher-dimensional dynamical system along analogous lines to Touboul and Ermentrout [13]. However, certain caution must be exercised, since one does not keep track of the validity of the truncated moment equations. Note that the path-integral methods used in this section were originally introduced within the context of stochastic processes by Martin–Siggia–Rose [60], and have previously been applied to stochastic neural networks by Sompolinsky et al. [61, 62] and Buice et al. [17, 20].

6.1 Generating Functional and Moments

First note that the average synaptic current U_α is given by

$$\langle\!\langle U_\alpha(t_1) \rangle\!\rangle = \int u_\alpha(t_1) p[\mathbf{u}, t_1] \, d\mathbf{u} = \int D[\mathbf{u}]D[\mathbf{p}] u_\alpha(t_1) e^{-S[\mathbf{u}, \mathbf{p}]/\epsilon} \tag{6.1}$$

and two-point correlations are

$$\langle\!\langle U_\alpha(t_1)U_\beta(t_2) \rangle\!\rangle = \int D[\mathbf{u}]D[\mathbf{p}] u_\alpha(t_1)u_\beta(t_2) e^{-S[\mathbf{u}, \mathbf{p}]/\epsilon}. \tag{6.2}$$

Another important characterization of the system is how the mean synaptic current responds to small external inputs. Suppose that we add a small external source term $h_\alpha(t)$ onto the right-hand side of the deterministic rate equation (2.19). Linearizing about the time-dependent solution of the unperturbed equation ($\mathbf{h} \equiv 0$)

leads to the following (non-autonomous) linear equation for the perturbed solution $u_\alpha(t) = u_\alpha^h(t) - u_\alpha^0(t)$:

$$\tau \frac{du_\alpha}{dt} = -u_\alpha + \sum_\beta w_{\alpha\beta} F'(u_\beta^0) u_\beta + h_\alpha(t). \tag{6.3}$$

Introducing the Green's function or propagator $G_{\alpha\beta}^0(t, t')$ according to the adjoint equation

$$-\tau \frac{dG_{\alpha\gamma}^0(t, t')}{dt'} = -G_{\alpha\gamma}^0(t, t') + \sum_\beta w_{\alpha\beta} F'(u_\beta^0) G_{\beta\gamma}^0(t, t')$$

$$+ \delta_{\alpha,\gamma} \delta(t - t'), \tag{6.4}$$

we can express the linear response as

$$u_\alpha(t) = \int^t \sum_\gamma G_{\alpha\gamma}^0(t, t') h_\gamma(t') \, dt'. \tag{6.5}$$

In other words, in terms of functional derivatives

$$\frac{\delta u_\alpha(t)}{\delta h_\gamma(t')} = G_{\alpha\gamma}^0(t, t'). \tag{6.6}$$

Now suppose that we add a source term to the path-integral representation. This corresponds to adding a term $\int \sum_\gamma h_\gamma(t) p_\gamma(t) \, dt$ to the action (3.29). It follows that the associated Green's function for the full stochastic model is given by

$$\mathcal{G}_{\alpha\gamma}(t, t') \equiv \frac{\delta \langle\!\langle U_\alpha(t) \rangle\!\rangle}{\delta h_\gamma(t')} = \langle\!\langle U_\alpha(t) P_\gamma(t') \rangle\!\rangle. \tag{6.7}$$

The above analysis motivates the introduction of the generating functional

$$Z[\mathbf{J}, \widetilde{\mathbf{J}}] = \int D[\mathbf{u}] D[\mathbf{p}]$$

$$\times \exp\left(-\frac{S[\mathbf{u}, \mathbf{p}]}{\epsilon} + \int dt \sum_\alpha \left[u_\alpha(t) \widetilde{J}_\alpha(t) + J_\alpha(t) p_\alpha(t) \right] \right). \tag{6.8}$$

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

$$\langle\!\langle U_\alpha(t) \rangle\!\rangle = \frac{\delta}{\delta \widetilde{J}_\alpha(t)} Z[\mathbf{J}, \widetilde{\mathbf{J}}] \Big|_{\mathbf{J}=\widetilde{\mathbf{J}}=0},$$

$$\langle\!\langle U_\alpha(t) U_\beta(t') \rangle\!\rangle = \frac{\delta}{\delta \widetilde{J}_\alpha(t)} \frac{\delta}{\delta \widetilde{J}_\beta(t)} Z[\mathbf{J}, \widetilde{\mathbf{J}}] \Big|_{\mathbf{J}=\widetilde{\mathbf{J}}=0},$$

$$\langle\!\langle U_\alpha(t) P_\beta(t') \rangle\!\rangle = \frac{\delta}{\delta \widetilde{J}_\alpha(t)} \frac{\delta}{\delta J_\beta(t)} Z[\mathbf{J}, \widetilde{\mathbf{J}}] \Big|_{\mathbf{J}=\widetilde{\mathbf{J}}=0}.$$

6.2 Effective Action and Corrections to Mean-Field Equations

Let us rescale the currents according to $\mathbf{J} \to \mathbf{J}/\epsilon$ and $\widetilde{\mathbf{J}} \to \widetilde{\mathbf{J}}/\epsilon$ so that we can apply a loop expansion of the path integral (6.8), which is a diagrammatic method for carrying out an ϵ expansion based on steepest descents or the saddle-point method. First, we introduce the exact means

$$v_\alpha = \langle\!\langle U_\alpha \rangle\!\rangle, \qquad \widetilde{v}_k = \langle\!\langle P_\alpha \rangle\!\rangle,$$

and we shift the variables by

$$u_\alpha(t) \to u_\alpha(t) + v_\alpha(t), \qquad p_\alpha(t) = p_\alpha(t) + \widetilde{v}_\alpha(t).$$

Expanding the action in (6.8) to second order in the shifted variables $\mathbf{u}$, $\mathbf{p}$ yields an infinite-dimensional Gaussian integral, which can be formally evaluated to give

$$Z[\mathbf{J}, \widetilde{\mathbf{J}}] \approx \mathrm{Det}\big[\mathcal{D}[\mathbf{v}, \widetilde{\mathbf{v}}]\big]^{-1/2}$$
$$\times \exp\left(-\frac{S[\mathbf{v}, \widetilde{\mathbf{v}}]}{\epsilon} + \frac{1}{\epsilon}\int dt \sum_\alpha \big[v_\alpha(t)\widetilde{J}_\alpha(t) + J_\alpha(t)\widetilde{v}_\alpha(t)\big]\right),$$

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

$$\mathcal{D}[\mathbf{v}, \widetilde{\mathbf{v}}]_{r\alpha, s\beta}(t, t') = \frac{\delta^2 S}{\delta \mathbf{u}_\alpha^r(t)\delta \mathbf{u}_\beta^s(t')}\bigg|_{\mathbf{u}=\mathbf{v}, \mathbf{p}=\widetilde{\mathbf{v}}}. \tag{6.9}$$

We have introduced the vectors $\mathbf{u}^r$, $r = 1, 2$ with $\mathbf{u}^1 = \mathbf{u}$, $\mathbf{u}^2 = \mathbf{p}$. Using the following identity for a matrix $\mathbf{M}$:

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

we obtain the $\mathcal{O}(\epsilon)$ approximation

$$Z[\mathbf{J}, \widetilde{\mathbf{J}}] \approx \mathrm{e}^{-S_{\mathrm{eff}}[\mathbf{v}, \widetilde{\mathbf{v}}]/\epsilon}\mathrm{e}^{\int dt \sum_\alpha [v_\alpha(t)\widetilde{J}_\alpha(t) + J_\alpha(t)\widetilde{v}_\alpha(t)]/\epsilon}, \tag{6.10}$$

where

$$S_{\mathrm{eff}}[\mathbf{v}, \widetilde{\mathbf{v}}] = S[\mathbf{v}, \widetilde{\mathbf{v}}] + \frac{\epsilon}{2}\mathrm{Tr}\log\big[\mathcal{D}[\mathbf{v}, \widetilde{\mathbf{v}}]\big]. \tag{6.11}$$

In order to use the above expansion to determine corrections to the mean-field equations, it is first necessary to introduce a little more formalism. First, consider the Legendre transformation

$$\Gamma[\mathbf{v}, \widetilde{\mathbf{v}}] = W[\mathbf{J}, \widetilde{\mathbf{J}}] + \int dt \sum_\alpha \big[v_\alpha(t)\widetilde{J}_\alpha(t) + J_\alpha(t)\widetilde{v}_\alpha(t)\big], \tag{6.12}$$

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

$$v_\alpha(t) = -\frac{\delta W}{\delta \widetilde{J}_\alpha(t)}, \qquad \widetilde{v}_\alpha(t) = -\frac{\delta W}{\delta J_\alpha(t)}, \tag{6.13}$$

it follows from functionally differentiating (6.12) that

$$\widetilde{J}_\alpha(t) = \frac{\delta \Gamma}{\delta \nu_\alpha(t)}, \qquad J_\alpha(t) = \frac{\delta \Gamma}{\delta \widetilde{\nu}_\alpha(t)}. \tag{6.14}$$

Dynamical equations for the physical mean fields $\nu_\alpha(t)$ are then generated by setting $\mathbf{J} = 0 = \widetilde{\mathbf{J}}$ in (6.14). Another useful result is obtained by functionally differentiating (6.13) with respect to the mean fields $\boldsymbol{\nu}, \widetilde{\boldsymbol{\nu}}$:

$$\delta(t - t')\delta_{r,s}\delta_{\alpha,\beta} = \frac{\delta \nu_\alpha^r(t)}{\delta \nu_\beta^s(t')} = -\frac{\delta^2 W}{\delta J_\alpha^r(t)\delta \nu_\beta^s(t')}$$

$$= -\sum_{q=1,2}\sum_{\gamma=1}^{M} \int \frac{\delta^2 W}{\delta J_\alpha^r(t)\delta J_\gamma^q(\tau)} \frac{\delta J_\gamma^q(\tau)}{\delta \nu_\beta^s(t')}\, d\tau,$$

where $\nu_\alpha^1 = \nu_\alpha$, $\nu_\alpha^2 = \widetilde{\nu}_\alpha$ and $J_\alpha^1 = \widetilde{J}_\alpha$, $J_\alpha^2 = J_\alpha$. Differentiating (6.14) with respect to $\mathbf{J}, \widetilde{\mathbf{J}}$ then shows that

$$\sum_{q=1,2}\sum_{\gamma=1}^{M} \int \frac{\delta^2 W}{\delta J_\alpha^r(t)\delta J_\gamma^q(\tau)} \frac{\delta^2 \Gamma}{\delta \nu_\gamma^q(\tau)\delta \nu_\beta^s(t')}\, d\tau = -\delta_{r,s}\delta_{k,l}\delta(t - t').$$

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

$$\widehat{\mathcal{D}}[\boldsymbol{\nu}, \widetilde{\boldsymbol{\nu}}]_{r,s}(\alpha, t; \beta, t') = \frac{\delta^2 \Gamma}{\delta \nu_\alpha^r(t)\delta \nu_\beta^s(t')}, \tag{6.15}$$

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

$$C_{r\alpha,s\beta}(t, t') \equiv -\frac{\delta^2 W}{\delta J_\alpha^r(t)\delta J_\beta^s(t')} = \left[\langle\!\langle U_\alpha^r(t) U_\beta^s(t') \rangle\!\rangle - \langle\!\langle U_\alpha^r(t) \rangle\!\rangle \langle\!\langle U_\beta^s(t') \rangle\!\rangle \right].$$

It now follows from (6.10) and (6.12) that $\Gamma[\boldsymbol{\nu}, \widetilde{\boldsymbol{\nu}}] = S_{\text{eff}}[(\boldsymbol{\nu}, \widetilde{\boldsymbol{\nu}})] + \mathcal{O}(\epsilon^2)$. Moreover, (6.9) and (6.15) imply that $\widehat{\mathcal{D}}[\boldsymbol{\nu}, \widetilde{\boldsymbol{\nu}}] = \mathcal{D}[\boldsymbol{\nu}, \widetilde{\boldsymbol{\nu}}] + \mathcal{O}(\epsilon)$, that is, we can take $\mathcal{D}[\boldsymbol{\nu}, \widetilde{\boldsymbol{\nu}}]$ to be the inverse of the two-point covariance matrix. The first-order correction to the mean-field equation (2.19) is then obtained from (6.14) after setting $\mathbf{J} = \widetilde{\mathbf{J}} = \widetilde{\boldsymbol{\nu}} = 0$:

$$0 = \frac{\delta \Gamma[\boldsymbol{\nu}, \widetilde{\boldsymbol{\nu}}]}{\delta \nu_\alpha(t)}\bigg|_{\widetilde{\boldsymbol{\nu}}=0} = \frac{\delta S[\mathbf{u}, \mathbf{p}]}{\delta p_\alpha(t)}\bigg|_{\mathbf{u}=\boldsymbol{\nu},\mathbf{p}=0} + \frac{\epsilon}{2}\operatorname{Tr}\mathcal{D}[\mathbf{u}, \mathbf{p}]^{-1}\frac{\delta \mathcal{D}[\mathbf{u}, \mathbf{p}]}{\delta p_\alpha(t)}\bigg|_{\mathbf{u}=\boldsymbol{\nu},\mathbf{p}=0},$$

with

$$\operatorname{Tr}\mathcal{D}[\mathbf{u}, \mathbf{p}]^{-1}\frac{\delta \mathcal{D}[\mathbf{u}, \mathbf{p}]}{\delta p_\alpha(t)}\bigg|_{\mathbf{u}=\boldsymbol{\nu},\mathbf{p}=0}$$

$$= \int dt' \int dt'' \sum_{r\beta,s\gamma} C_{r\beta,s\gamma}(t', t'') \frac{\delta}{\delta p_\alpha(t)}\frac{\delta^2 S[\mathbf{u}, \mathbf{p}]}{\delta u_\beta^r(t')u_\gamma^s(t'')}\bigg|_{\mathbf{u}=\boldsymbol{\nu},\mathbf{p}=0}.$$

The functional derivative in the above equation forces $t = t' = t''$ (see also [20]). Since the only non-vanishing, equal-time two-point correlation function when $\mathbf{p} = 0$ is for $r = s = 1$, it follows that

$$\operatorname{Tr}\mathcal{D}[\mathbf{u},\mathbf{p}]^{-1}\frac{\delta\mathcal{D}[\mathbf{u},\mathbf{p}]}{\delta p_\alpha(t)}\bigg|_{\mathbf{u}=\mathbf{v},\mathbf{p}=0} = \sum_{\beta,\gamma}C_{\beta\gamma}(t)S^\alpha_{\gamma\beta}(t),$$

where

$$C_{\beta\gamma}(t) = \big[\langle\!\langle U_\beta(t)U_\gamma(t)\rangle\!\rangle - \langle\!\langle U_\beta(t)\rangle\!\rangle\langle\!\langle U_\gamma(t)\rangle\!\rangle\big],$$

and

$$\frac{\delta}{\delta p_\alpha(t)}\frac{\delta^2 S[\mathbf{u},\mathbf{p}]}{\delta u_\beta(t')u_\gamma(t'')}\bigg|_{\mathbf{u}=\mathbf{v},\mathbf{p}=0} = S^\alpha_{\gamma\beta}(t)\delta(t-t')\delta(t-t'').$$

Evaluating the functional derivative of the action S given by (3.29) and (3.35) finally yields the lowest-order correction to the mean-field equation (2.19), which could not be obtained from the Langevin equation (5.2):

$$\tau\frac{du_\alpha}{dt} = -u_\alpha(t) + \sum_{\beta=1}^M w_{\alpha\beta}F(u_\beta) - \frac{\epsilon}{2}\sum_\beta w_{\alpha\beta}C_{\beta\beta}(t)F''\big(u_\beta(t)\big) + \mathcal{O}(\epsilon^2). \quad (6.16)$$

It is also possible to derive a corresponding dynamical equation for the two-point correlation function by extending the definition of the effective action along the lines of Buice et al. [20]. However, the lowest-order equation for $\mathbf{C}$ can be obtained from (5.2). One finds that

$$\begin{aligned}
\frac{dC_{\alpha\beta}}{dt} &= \mathcal{Q}_{\alpha\beta}(\mathbf{u}) + \sum_\gamma\left[\frac{\partial\overline{V}_\alpha(\mathbf{u})}{\partial u_\gamma}C_{\gamma\beta} + C_{\alpha\gamma}\frac{\partial\overline{V}_\beta(\mathbf{u})}{\partial u_\gamma}\right]\\
&= \sum_\gamma w_{\alpha\gamma}F(u_\gamma)w_{\beta\gamma} + \sum_\gamma\big[-\delta_{\alpha,\gamma} + w_{\alpha\gamma}F'(u_\gamma)\big]C_{\gamma\beta}\\
&\quad + \sum_\gamma\big[-\delta_{\beta,\gamma} + w_{\beta\gamma}F'(u_\gamma)\big]C_{\alpha\gamma}\\
&= -2C_{\alpha\beta} + \sum_\gamma w_{\alpha\gamma}F(u_\gamma)w_{\beta\gamma}\\
&\quad + \sum_\gamma F'(u_\gamma)[w_{\alpha\gamma}C_{\gamma\beta} + w_{\beta\gamma}C_{\alpha\gamma}]. \quad (6.17)
\end{aligned}$$

The corrections to mean-field theory for a stochastic hybrid neural network differ significantly from those derived for the Buice et al. master equation [17, 20]. There are two primary sources of such differences. One arises from the fact that the mean equation is in 'Amari form' (with the weight matrix outside the nonlinearity). This accounts for all the difference in (6.16) for the mean, which would otherwise be identical to that of Buice et al., and the last term involving C in (6.17). The other

difference is in the non-homogeneous source term for the C equation, which appears as $\sum_\gamma w_{\alpha\gamma} F(u_\gamma) w_{\beta\gamma}$. Whereas the Buice et al. correlations are determined by multiple network motifs (with the lowest order being the direct connection $w_{\alpha\beta}$ from β to α), our result for the hybrid model indicates that the source term is given by divergent motifs indicating common input from a third population (population $\gamma \rightarrow$ populations α, β).

7 Discussion

In conclusion, we have constructed a path-integral representation of solutions to a stochastic hybrid neural network, and shown how this provides a unifying framework for carrying out various perturbation schemes for analyzing the stochastic dynamics, namely, large deviations, diffusion approximations, and corrections to mean-field equations. We highlighted the fact that the path-integral action can be expressed in terms of a Hamiltonian, which is given by the Perron eigenvalue of an appropriately defined linear operator. The latter depends on the transition rates and drift terms of the underlying hybrid system. The resulting action is consistent with that obtained using large-deviation theory.

In terms of the theory of stochastic neural networks, our hybrid model extends the neural master equation to include the effects of synaptic currents. In the limit of fast synapses one recovers the neural master equation, which can be viewed as a stochastic version of the 'Wilson–Cowan' rate equations (with the weight matrix inside the nonlinearity). On the other hand, in the case of slow synapses, one obtains a stochastic version of the 'Amari' rate equations. This leads to significant differences in the corrections to the mean-field equations. Finally, it should be noted that the path-integral formulation presented here can be applied to more general stochastic hybrid systems such as stochastic ion channels, molecular motors, and gene networks [28–32]. Thus one can view our path-integral construction as the hybrid analog of the Doi–Peliti path integral for master equations.

Competing Interests

The author declares that they have no competing interests.

References

1. Softky WR, Koch C. Cortical cell should spike regularly but do not. Neural Comput. 1992;4:643–6.
2. Faisal AA, Selen LPJ, Wolpert DM. Noise in the nervous system. Nat Rev Neurosci. 2008;9:292.
3. Shadlen MN, Newsome WT. Noise, neural codes and cortical organization. Curr Opin Neurobiol. 1994;4:569–79.
4. van Vreeswijk C, Sompolinsky H. Chaotic balanced state in a model of cortical circuits. Neural Comput. 1998;10:1321–71.

5. Vogels TP, Abbott LF. Signal propagation and logic gating in networks of integrate-and-fire neurons. J Neurosci. 2005;25:786–95.
6. London M, Roth A, Beeren L, Hausser M, Latham PE. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. Nature. 2010;466:123–7.
7. Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A, Harris KD. The asynchronous state in cortical circuits. Science. 2010;327:587–90.
8. Churchland MM, et al. Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci. 2010;13:369–78.
9. Litwin-Kumar A, Doiron B. Slow dynamics and high variability in balanced cortical networks with clustered connections. Nat Neurosci. 2012;15:1498–505.
10. Bressloff PC. Spatiotemporal dynamics of continuum neural fields. J Phys A. 2012;45:033001.
11. Hutt A, Longtin A, Schimansky-Geier L. Additive noise-induces Turing transitions in spatial systems with application to neural fields and the Swift–Hohenberg equation. Physica D. 2008;237:755–73.
12. Faugeras O, Touboul J, Cessac B. A constructive mean-field analysis of multi-population neural networks with random synaptic weights and stochastic inputs. Front Comput Neurosci. 2009;3:1.
13. Touboul JD, Ermentrout GB. Finite-size and correlation-induced effects in mean-field dynamics. J Comput Neurosci. 2011;31:453–84.
14. Touboul J, Hermann G, Faugeras O. Noise-induced behaviors in neural mean field dynamics. SIAM J Appl Dyn Syst. 2012;11:49–81.
15. Bressloff PC, Webber MA. Front propagation in stochastic neural fields. SIAM J Appl Dyn Syst. 2012;11:708–40.
16. Ohira T, Cowan JD. Stochastic neurodynamics and the system size expansion. In: Ellacott S, Anderson IJ, editors. Proceedings of the first international conference on mathematics of neural networks. San Diego: Academic Press. 1997. p. 290–4.
17. Buice M, Cowan JD. Field-theoretic approach to fluctuation effects in neural networks. Phys Rev E. 2007;75:051919.
18. Bressloff PC. Stochastic neural field theory and the system-size expansion. SIAM J Appl Math. 2009;70:1488.
19. Bressloff PC. Metastable states and quasicycles in a stochastic Wilson–Cowan model of neuronal population dynamics. Phys Rev E. 2010;85:051903.
20. Buice M, Cowan JD, Chow CC. Systematic fluctuation expansion for neural network activity equations. Neural Comput. 2010;22:377.
21. Doi M. Second quantization representation for classical many-particle systems. J Phys A. 1976;9:1465–77.
22. Doi M. Stochastic theory of diffusion controlled reactions. J Phys A. 1976;9:1479–95.
23. Peliti L. Path integral approach to birth–death processes on a lattice. J Phys. 1985;46:1469–83.
24. Dykman MI, Mori E, Ross J, Hunt PM. Large fluctuations and optimal paths in chemical kinetics. J Chem Phys. 1994;100:5735.
25. Elgart V, Kamenev A. Rare event statistics in reaction–diffusion systems. Phys Rev E. 2004;70:041106.
26. Escudero C, Kamenev A. Switching rates of multistep reactions. Phys Rev E. 2009;79:041149.
27. Bressloff PC, Newby JM. Metastability in a stochastic neural network modeled as a velocity jump Markov process. SIAM J Appl Dyn Syst. 2013;12:1394–435.
28. Keener JP, Newby JM. Perturbation analysis of spontaneous action potential initiation by stochastic ion channels. Phys Rev E. 2011;84:011918.
29. Newby JM, Keener JP. An asymptotic analysis of the spatially inhomogeneous velocity-jump process. SIAM J Multiscale Model Simul. 2011;9:735–65.
30. Newby JM. Isolating intrinsic noise sources in a stochastic genetic switch. Phys Biol. 2012;9:026002.
31. Newby JM, Bressloff PC, Keener JP. Breakdown of fast–slow analysis in an excitable system with channel noise. Phys Rev Lett. 2013;111:128101.
32. Bressloff PC, Newby JM. Stochastic hybrid model of spontaneous dendritic NMDA spikes. Phys Biol. 2014;11:016006.
33. Bressloff PC, Newby JM. Path integrals and large deviations in stochastic hybrid systems. Phys Rev E. 2014;89:042701.
34. Bressloff PC, Faugeras O. On the Hamiltonian structure of large deviations in stochastic hybrid systems. Submitted 2015.
35. Buice M, Chow CC. Beyond mean field theory: statistical field theory for neural networks. J Stat Mech Theory Exp. 2013;2013:P03003.

36. Buice M, Chow CC. Dynamic finite size effects in spiking neural networks. PLoS Comput Biol. 2013;9:e1002872.
37. Buice M, Chow CC. Generalized activity equations for spiking neural network dynamics. Front Comput Neurosci. 2013;7:162.
38. Grimmett GR, Stirzaker DR. Probability and random processes. 3rd ed. Oxford: Oxford University Press; 2001.
39. Ermentrout GB. Reduction of conductance-based models with slow synapses to neural nets. Neural Comput. 1994;6:679–95.
40. Haangi P, Grabert H, Talkner P, Thomas H. Bistable systems: master equation versus Fokker–Planck modeling. Z Phys B. 1984;28:135.
41. Kifer Y. Large deviations and adiabatic transitions for dynamical systems and Markov processes in fully coupled averaging. Mem Am Math Soc. 2009;201(944):1–129.
42. Faggionato A, Gabriell D, Crivellari MR. Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors. Markov Process Relat Fields. 2010;16:497–548.
43. Faggionato A, Gabrielli D, Ribezzi Crivellari M. Non-equilibrium thermodynamics of piecewise deterministic Markov processes. J Stat Phys. 2009;137:259–304.
44. Graham R, Tel T. On the weak-noise limit of Fokker–Planck models. J Stat Phys. 1984;35:729–48.
45. Lugo CA, McKane AJ. Quasi-cycles in a spatial predator–prey model. Phys Rev E. 2008;78:051911.
46. Biancalani T, Fanelli D, Di Patt F. Stochastic Turing patterns in the Brusselator model. Phys Rev E. 2010;81:046215.
47. Butler TC, Goldenfeld N. Robust ecological pattern formation induced by demographic noise. Phys Rev E. 2009;80:030902(R).
48. Butler TC, Goldenfeld N. Fluctuation-driven Turing patterns. Phys Rev E. 2011;84:011112.
49. Woolley TE, Baker RE, Gaffney EA, Maini PK. Stochastic reaction and diffusion on growing domains: understanding the breakdown of robust pattern formation. Phys Rev E. 2011;84:046216.
50. Schumacher LJ, Woolley TE, Baker RE. Noise-induced temporal dynamics in Turing systems. Phys Rev E. 2013;87:042719.
51. McKane AJ, Biancalani T, Rogers T. Stochastic pattern formation and spontaneous polarization: the linear noise approximation and beyond. Bull Math Biol. 2014;76:895–921.
52. Ermentrout GB, Cowan JD. A mathematical theory of visual hallucination patterns. Biol Cybern. 1979;34:137–50.
53. Bressloff PC, Cowan JD, Golubitsky M, Thomas PJ, Wiener M. Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. Philos Trans R Soc Lond B. 2001;356:299–330.
54. Webber M, Bressloff PC. The effects of noise on binocular rivalry waves: a stochastic neural field model: invited contribution. J Stat Mech. 2013;3:P03001.
55. Kilpatrick ZP, Ermentrout GB. Wandering bumps in stochastic neural fields. SIAM J Appl Dyn Syst. 2013;12:61–94.
56. Kilpatrick ZP. Coupling layers regularizes wave propagation in stochastic neural fields. Phys Rev E. 2014;89:022706.
57. Faugeras O, Inglis J. Stochastic neural field theory: a rigorous footing. J Math Biol. 2015. doi:10.1007/s00285-014-0807-6.
58. Kruger M, Stannat W. Front propagation in stochastic neural fields: a rigorous mathematical framework. 2014. arXiv:1406.2675v1.
59. Kuehn C, Reidler MG. Large deviations for nonlocal stochastic neural fields. J Math Neurosci. 2014;4:1.
60. Martin PC, Siggia ED, Rose HA. Statistical dynamics of classical systems. Phys Rev A. 1973;8:423–37.
61. Sompolinsky H, Zippelius A. Dynamic theory of the spin glass phase. Phys Rev Lett. 1981;47:359.
62. Crisanti A, Sompolinsky H. Dynamics of spin systems with randomly asymmetric bonds: Ising spins and Glauber dynamics. Phys Rev A. 1988;37:4865.