# Chapter 6
# Population Models and Neural Fields

In this chapter we show how to construct population-based models of synaptically coupled neural networks, distinguishing between voltage-based and activity-based versions [71, 168]. We also consider the important issue of how noise at the single-cell level (Poisson-like spiking statistics) relates to noise at the population level. In particular, we describe a neural master equation formulation of stochastic population dynamics. One biological motivation for population-based models is that neurons in cortex tend to be organized into tightly coupled groups known as cortical columns, in which the cells share similar functional properties (Sects. 5.1 and 8.1). Another important feature of cortex is that it is spatially organized at both the anatomical and functional levels, which can be modeled in terms of spatially structured networks of interacting populations. In the continuum limit, such models can be described in terms of neural field equations. The advantage of a continuum rather than a discrete representation of spatially structured networks is that various techniques from the analysis of PDEs presented in Part I can be adapted to the study of neural field models of cortex (Chaps. 7 and 8).

## 6.1 Population Averaging and Rate Models

Suppose that a network of synaptically coupled spiking neurons is partitioned into a set of $P$ homogeneous populations labeled $\alpha = 1,\ldots,P$, with $N_\alpha = N$ neurons in each population. (It is straightforward to relax this assumption by taking $N_\alpha = \mathcal{O}(N)$.) Let $\chi$ denote the population function that maps the single neuron index $i = 1,\ldots,N$ to the population index $\alpha$ to which neuron $i$ belongs: $\chi(i) = \alpha$. Furthermore, suppose the synaptic interactions between populations are the same for all neuron pairs. (Relaxing this assumption can lead to additional sources of stochasticity as explored in [184, 632].) Denote the sequence of firing times of the $j$th neuron by $\{T_j^m, m \in \mathbf{Z}\}$. The net synaptic current into postsynaptic neuron $i$ due to innervation by the spike train from presynaptic neuron $j$, with $\chi(i) = \alpha, \chi(j) = \beta$, is taken to have the general form $N^{-1}\sum_m \Phi_{\alpha\beta}(t - T_j^m)$, where $N^{-1}\Phi_{\alpha\beta}(t)$ represents

P.C. Bressloff, *Waves in Neural Media: From Single Neurons to Neural Fields*,
Lecture Notes on Mathematical Modelling in the Life Sciences,
DOI 10.1007/978-1-4614-8866-8_6, 

the temporal filtering effects of synaptic and dendritic processing of inputs from any neuron of population $\beta$ to any neuron of population $\alpha$; see Sect. 1.4.2. Assuming that all synaptic inputs sum linearly, the total synaptic input to the soma of the $i$th neuron is

$$u_i(t) = \sum_\beta \frac{1}{N} \sum_{j;\chi(j)=\beta} \Phi_{\alpha\beta}(t - T_j^m) = \int_{-\infty}^t \sum_\beta \Phi_{\alpha\beta}(t-t') \frac{1}{N} \sum_{j;\chi(j)=\beta} a_j(t')dt' \tag{6.1}$$

for all $\chi(i) = \alpha$, where $a_j(t) = \sum_{m\in\mathbf{Z}} \delta(t - T_j^m)$. That is, $a_j(t)$ represents the output spike train of the $j$th neuron in terms of a sum of Dirac delta functions. In order to obtain a closed set of equations, we have to determine the firing times $T_i^m$ given by Eq. (5.2), where $v_i(t)$ evolves according to conductance-based model (5.1). It follows that, since the right-hand side of (6.1) is independent of $i$, we can set $u_i(t) = u_\alpha(t)$ for all $p(i) = \alpha$ with

$$u_\alpha(t) = \sum_{\beta=1}^P \int_{-\infty}^t \Phi_{\alpha\beta}(t-t') a_\beta(t')dt', \tag{6.2}$$

and $a_\alpha(t)$ is the output activity of the $\alpha$th population:

$$a_\alpha(t) = \frac{1}{N} \sum_{j;\chi(j)=\alpha} a_j(t). \tag{6.3}$$

We now assume that each homogeneous population is close to a so-called asynchronous state, in which the spike trains of different neurons within a homogeneous population are uncorrelated (see Sect. 6.3.1). The population activity $a_\alpha$ is then approximately constant, which means that the synaptic currents are also slowly varying functions of time. It follows that $a_\alpha$ can track the input current according to $a_\alpha(t) = F(u_\alpha(t))$, where $F$ is a population gain function [2, 86, 87, 213]. Substituting this approximation into (6.2) leads to the closed system of equations

$$u_\alpha(t) = \sum_{\beta=1}^P \int_{-\infty}^t \Phi_{\alpha\beta}(t-t') F[u_\beta(t')]dt', \tag{6.4}$$

(Note, however, the asynchronous state only exists in the thermodynamic limit so that for finite $N$ we expect statistical fluctuations around the asynchronous state; see Sect. 6.4.) A rate equation identical in form to (6.4) was derived in Sect. 5.3 for an IF network with slow synapses, except that it involved single neuron labels rather than the population labels $\alpha, \beta$. As highlighted elsewhere [71, 168], Eq. (6.4) can be reduced to a system of ordinary differential equations provided that we place rather mild restrictions on the time dependence of $\Phi_{\alpha\beta}(t)$. First, suppose that $\Phi_{\alpha\beta}(t) = w_{\alpha\beta}\Phi_\alpha(t)$ where $w_{\alpha\beta}$ denotes the mean synaptic strength of connections from population $\beta$ to neuron $\alpha$ and $\Phi_\alpha(t)$ determines the time course of the input, which is assumed to depend only on properties of the postsynaptic population

$\alpha$. Furthermore, suppose that there exists a differential operator $\mathbb{L}_\alpha$ such that (see also (5.76))

$$\mathbb{L}_\alpha \Phi_\alpha(t) = \delta(t). \tag{6.5}$$

Applying the operator $\mathbb{L}_\alpha$ to both sides of equation (6.4) then leads to a system of differential equations for the population currents $u_\alpha(t)$:

$$\mathbb{L}_\alpha u_\alpha(t) = \sum_{\beta=1}^{P} w_{\alpha\beta} F_\beta(u_\beta(t)). \tag{6.6}$$

Note that we could easily convert the synaptic current $u_\alpha(t)$ into an input voltage $v_\alpha(t) = u_\alpha(t)/\sigma$ using an effective input conductance $\sigma$, for example. Thus Eq. (6.6) is often referred to as a voltage equation and forms the basis of most classical neural networks such as the Hopfield model [288]. On the other hand, if the time course of the inputs depends only on presynaptic parameters, $\Phi_{\alpha\beta}(t) = w_{\alpha\beta}\Phi_\beta(t)$, with $\Phi_\beta$ having inverse differential operator $\mathbb{L}_\beta$, then we obtain a system of differential equations for the so-called synaptic drives:

$$z_\alpha(t) = \int_{-\infty}^{t} \Phi_\alpha(t-t') F_\alpha(u_\alpha(t'))dt'. \tag{6.7}$$

That is, applying the differential operator $\mathbb{L}_\alpha$ to Eq. (6.7) and using $u_\alpha(t) = \sum_{\beta=1}^{P} w_{\alpha\beta} z_\beta(t)$ leads to the activity-based model

$$\mathbb{L}_\alpha z_\alpha(t) = F_\alpha\left(\sum_{\beta=1}^{P} w_{\alpha\beta} z_\beta(t)\right). \tag{6.8}$$

The differential operator $\mathbb{L}_\alpha$ appearing in Eqs. (6.6) and (6.8) is often taken to be first order in time:

$$\mathbb{L}_\alpha = \frac{\partial}{\partial t} + \frac{1}{\tau_\alpha}, \tag{6.9}$$

with inverse kernel $\Phi_\alpha(t) = H(t)\mathrm{e}^{-t/\tau_\alpha}$. In order to relate the effective time constant $\tau_\alpha$ to membrane and synaptic time constants, let us assume for simplicity that all synapses are sufficiently close to the soma so that the dendrite simply acts as a first-order low-pass filter and set (with $V_{\text{rest}} = 0$)

$$\Phi_{\alpha\beta}(t) = \sigma_\alpha r_{m,\alpha} V_{\text{syn},\beta} \bar{g}_{\alpha\beta} H(t) \int_0^t \mathrm{e}^{-(t-s)/\tau_{m,\alpha}} h_\beta(s)ds,$$

with

$$h_\beta(s) = \frac{\tau_{d,\beta}}{\tau_{d,\beta} - \tau_{r,\beta}} (\mathrm{e}^{-s/\tau_{d,\beta}} - \mathrm{e}^{-s/\tau_{r,\beta}}).$$

We have made explicit that the reversal potential $V_{\text{syn}}$ and synaptic rise/fall times $\tau_{r,d}$ only depend on the particular class of synapses innervated by the presynaptic population $\beta$, whereas the membrane time constant $\tau_m$, resistance $r_m$, and conductance $\sigma$ are solely properties of the postsynaptic population $\alpha$. Only the maximum

conductance $\bar{g}$ is specific to the particular synapses $\beta \to \alpha$. The various constant factors can be combined to define the synaptic weight $w_{\alpha\beta}$. In particular, $w_{\alpha\beta} \sim V_{\text{syn},\beta}$ so that the sign of $V_{\text{syn},\beta}$ (relative to the resting potential) determines whether the synapse is excitatory or inhibitory. If $\tau_m \gg \tau_r, \tau_d$, then the time course is effectively independent of the presynaptic label $\beta$, and we have the voltage-based model (6.6) with first-order operator $\mathbb{L}_\alpha$ and $\tau_\alpha = \tau_m$. On the other hand, if $\tau_d \gg \tau_m, \tau_r$, then we obtain the activity-based model with $\tau_\alpha = \tau_d$. Although the reduction to a rate-based model is a major simplification of the full conductance-based model, it is still possible to incorporate various additional physiological features.

1. *Synaptic depression.* In order to incorporate synaptic depression we need to return to the single neuron level. Equation (6.1) is modified according to

$$\begin{aligned} u_i(t) &= \sum_\beta \frac{1}{N} \sum_{j;\chi(j)=\beta} \Phi_{\alpha\beta}(t - T_j^m) q_{ij}(t - T_j^m) \\ &= \int_{-\infty}^t \sum_\beta \Phi_{\alpha\beta}(t-t') \left[ \frac{1}{N} \sum_{j;\chi(j)=\beta} a_j(t') q_{ij}(t') \right] dt' \end{aligned}$$

with $q_{ij}$ evolving according to an equation of the form (1.45), which we write as

$$\frac{dq_{ij}}{dt} = \frac{1 - q_{ij}}{\tau_q} - (1-\gamma) q_{ij}(t) a_j(t).$$

Averaging the latter equation with respect to $j, \chi(j) = \beta$, and introducing the compact notation

$$\langle f(t) \rangle_\beta = \frac{1}{N} \sum_{j;\chi(j)=\beta} f_j(t),$$

we have the pair of equations

$$u_i(t) = \int_{-\infty}^t \sum_\beta \Phi_{\alpha\beta}(t-t') \langle a(t') q_i(t') \rangle_\beta dt' \tag{6.10}$$

and

$$\frac{d\langle q_i \rangle_\beta}{dt} = \frac{1 - \langle q_i \rangle_\beta}{\tau_q} - (1-\gamma) \langle a(t) q_i(t) \rangle_\beta.$$

We now make the mean-field approximation

$$\langle a(t) q_i(t) \rangle_\beta = a_\beta(t) \langle q_i(t) \rangle_\beta \tag{6.11}$$

Since all averaged depression variables $\langle q_i(t) \rangle_\beta, i = 1, \ldots, N\}$ for fixed $\beta$ now have a common input drive $a_\beta(t)$, it follows that

$$\tau_q \frac{d(\langle q_i(t) \rangle_\beta - \langle q_{i'}(t) \rangle_\beta)}{dt} = -[\langle q_i(t) \rangle_\beta - \langle q_{i'}(t) \rangle_\beta],$$

and thus $\langle q_i(t)\rangle_\beta \to \langle q_{i'}(t)\rangle_\beta = q_\beta(t)$ for all $i,i' = 1,\ldots,N$. In other words, after an initial transient of duration $\tau_q$, we can identify all depression variables associated with a given presynaptic population $\beta$. If we now assume that $\Phi_{\alpha\beta}(t) = w_{\alpha\beta}\Phi_\beta(t)$, we can introduce the synaptic drives (6.7) and derive the modified activity-based model [25, 617, 637, 644]:

$$\mathbb{L}_\alpha z_\alpha(t) = F_\alpha\left(\sum_{\beta=1}^{P} w_{\alpha\beta}q_\beta(t)z_\beta(t)\right), \tag{6.12}$$

with

$$\frac{dq_\alpha}{dt} = \frac{1-q_\alpha(t)}{\tau_q} - (1-\gamma)q_\alpha(t)F_\alpha\left(\sum_{\beta=1}^{P} w_{\alpha\beta}q_\beta(t)z_\beta(t)\right). \tag{6.13}$$

The corresponding voltage-based model is

$$\mathbb{L}_\alpha u_\alpha(t) = \sum_{\beta=1}^{P} w_{\alpha\beta}q_\beta(t)F_\beta(u_\beta(t)). \tag{6.14}$$

with

$$\frac{dq_\alpha}{dt} = \frac{1-q_\alpha(t)}{\tau_q} - (1-\gamma)q_\alpha(t)F_\alpha(u_\alpha(t)). \tag{6.15}$$

2. *Axonal propagation delays.* In the above derivation of rate-based models, we have assumed that the spiking of a presynaptic neuron has an instantaneous effect on downstream postsynaptic neurons. This neglects the fact that action potentials take time to propagate along an axon to innervate a synaptic terminal. Let us denote the corresponding axonal delay of a synapse $\alpha \to \beta$ by $\tau_{\alpha\beta}$. The integral equation (6.2) is modified according to

$$u_\alpha(t) = \sum_{\beta=1}^{P}\int_{-\infty}^{t} \Phi_{\alpha\beta}(t-t')a_\beta(t'-\tau_{\alpha\beta})dt'. \tag{6.16}$$

The corresponding voltage-based model then takes the form of a system of delay differential equations,

$$\mathbb{L}_\alpha u_\alpha(t) = \sum_{\beta=1}^{P} w_{\alpha\beta}F_\beta(u_\beta(t-\tau_{\alpha\beta})), \tag{6.17}$$

and similarly for the activity-based model.

3. *Adaptive threshold dynamics.* Another biophysical process that can be incorporated into rate-based models is spike frequency adaptation. Spike frequency adaptation causes a neuron's firing rate to decay to a submaximal level and occurs when a potassium current, presumably activated by elevated intracellular

calcium, hyperpolarizes the membrane voltage [35, 396, 604]. This afterhyperpolarization current has a time constant of around 40–120 ms. Spike frequency adaptation can be introduced as a negative current $-c_i$ on the right-hand side of the conductance-based model equation (5.1). Assuming that $c_i(t)$ varies slowly relative to the voltage $v_i(t)$, it can be shown that $c_i$ effectively acts as an adaptive threshold that varies linearly with the firing rate [35]. In the case of a network of homogeneous populations, each neuron within a given population fires at the same mean rate so that we can identify $c_i = c_\alpha$ for all $i$ such that $\chi(i) = \alpha$. Hence, the voltage-based model becomes

$$\mathbb{L}_\alpha u_\alpha(t) = \sum_{\beta=1}^{P} w_{\alpha\beta} F_\beta(u_\beta(t) - c_\beta(t)). \tag{6.18}$$

with

$$\frac{dc_\alpha}{dt} = -\frac{c_\alpha(t)}{\tau_c} + \gamma_c F_\alpha(u_\alpha(t) - c_\alpha(t)). \tag{6.19}$$

## 6.2 E–I Oscillator Networks

One of the simplest population-based networks, which is often taken to be a fundamental module in large-scale models of cortex, is a pair of mutually coupled local populations of excitatory and inhibitory neurons known as an E–I network; see Fig. 6.1. An E–I network has the important property of exhibiting limit cycle oscillations and can thus act as a basic oscillatory element in network models of cortical phase waves, as an alternative to single spiking neurons (Sect. 5.2). An activity-based version of an E–I network takes the form (for first-order synapses)

$$\begin{aligned} \frac{da_E}{dt} &= -a_E + F\left(w_{EE}a_E - w_{EI}a_I + h_E\right) \\ \frac{da_I}{dt} &= -a_I + F\left(w_{IE}a_E - w_{II}a_I + h_I\right), \end{aligned} \tag{6.20}$$

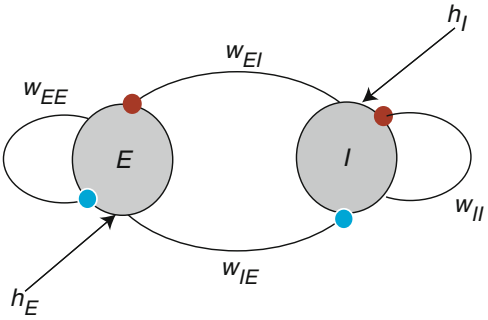


**Fig. 6.1** Two-population E–I network

where $h_E, h_I$ represent constant background inputs. For simplicity, we set $\tau_E = \tau_I = 1$. The bifurcation structure of the two-population model given by Eq. (6.20) has been analyzed in detail elsewhere [55]. An equilibrium $(a_E^*, a_I^*)$ is obtained as a solution of the pair of equations

$$\begin{aligned} a_E^* &= F\left(w_{EE}a_E^* - w_{EI}a_I^* + h_E\right) \\ a_I^* &= F\left(w_{IE}a_E^* - w_{II}a_I^* + h_I\right). \end{aligned} \tag{6.21}$$

These can be inverted to yield

$$\begin{aligned} h_E &= F^{-1}(a_E^*) - w_{EE}a_E^* + w_{EI}a_I^* \\ h_I &= F^{-1}(a_I^*) - w_{IE}a_E^* + w_{II}a_I^*. \end{aligned} \tag{6.22}$$

As a further simplification, take the gain function $F$ to be the simple sigmoid $F(u) = (1+\mathrm{e}^{-u})^{-1}$. Using the fact that the sigmoid function then satisfies $F' = F(1-F)$ and applying the fixed-point equations allows one to represent the associated Jacobian in the form

$$\Delta = \begin{pmatrix} -1 + w_{EE}a_E^*(1-a_E^*) & -w_{EI}a_E^*(1-a_E^*) \\ w_{IE}a_I^*(1-a_I^*) & -1 - w_{II}a_I^*(1-a_I^*) \end{pmatrix}.$$

An equilibrium will be stable provided that the eigenvalues $\lambda_\pm$ of $\Delta$ have negative real parts, where

$$\lambda_\pm = \frac{1}{2}\left(\mathrm{Tr}\,\Delta \pm \sqrt{[\mathrm{Tr}\,\Delta]^2 - 4\mathrm{Det}\,\Delta}\right). \tag{6.23}$$

This leads to the stability conditions $\mathrm{Tr}\,\Delta < 0$ and $\mathrm{Det}\,\Delta > 0$. In order to construct a phase diagram in the $(h_E, h_I)$-plane for a fixed weight matrix $\mathbf{w}$, we express $a_I^*$ as a function of $a_E^*$ by imposing a constraint on the eigenvalues $\lambda_\pm$ and then substitute the resulting function into Eq. (6.22). This yields bifurcation curves in the $(h_E, h_I)$-plane that are parameterized by $a_E^*$, $0 < a_E^* < 1$; see Fig. 6.2. For example, the constraint

$$\mathrm{Tr}\,\Delta \equiv -2 + w_{EE}a_E^*(1-a_E^*) - w_{II}a_I^*(1-a_I^*) = 0 \tag{6.24}$$

with $\mathrm{Det}\,\Delta > 0$ determines Hopf bifurcation curves where a pair of complex conjugate eigenvalues cross the imaginary axis. Since the trace is a quadratic function of $a_E^*, a_I^*$, we obtain two Hopf branches. Similarly, the constraint $\mathrm{Det}\,\Delta = 0$ with $\mathrm{Tr}\,\Delta < 0$ determines saddle–node or fold bifurcation curves where a single real eigenvalue crosses zero. The saddle–node curves have to be determined numerically, since the determinant is a quartic function of $a_E^*, a_I^*$.

Now consider a network of synaptically coupled E–I modules or subnetworks, with each module labeled by the discrete index $n$:

$$\frac{da_E^n}{dt} = -\frac{a_E^n}{\tau_E} + F\left(\sum_m \left[w_{EE}(n,m)a_E^m - w_{EI}(n,m)a_I^m + h_E\right]\right) \tag{6.25a}$$

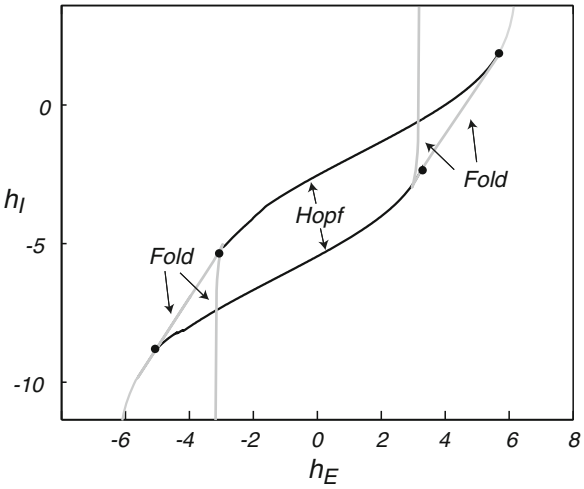


**Fig. 6.2** Phase diagram of two-population Wilson–Cowan model (6.20) for fixed set of weights $w_{EE} = 11.5, w_{IE} = w_{EI} = 10, , w_{II} = -2$. The *dots* correspond to Takens–Bogdanov bifurcation points

$$\frac{da_I^n}{dt} = -\frac{a_I^n}{\tau_I} + F\left(\sum_m [w_{IE}(n,m)a_E^m - w_{II}(n,m)a_I^m + h_I]\right), \tag{6.25b}$$

Suppose that synaptic interactions within a local E–I network are stronger than those between E–I networks. That is, we write

$$w_{ab}(n,m) = w_{ab}\delta_{nm} + \varepsilon J_{ab}(n,m)(1-\delta_{nm})$$

with $\varepsilon \ll 1$. Substituting into (6.25) and Taylor expanding to $\mathcal{O}(\varepsilon)$ then gives

$$\begin{aligned}\frac{da_E^n}{dt} &= -a_E^n + F\left([w_{EE}a_E^n - w_{EI}a_I^n + h_E]\right) \\ &\quad + \varepsilon F'\left([w_{EE}a_E^n - w_{EI}a_I^n + h_E]\right)\sum_{m\neq n}[J_{EE}(n,m)a_E^m - J_{EI}(n,m)a_I^m]\end{aligned} \tag{6.26a}$$

$$\begin{aligned}\frac{da_I^n}{dt} &= -a_I^n + F\left([w_{IE}a_E^n - w_{II}a_I^n + h_I]\right) \\ &\quad + \varepsilon F'\left([w_{IE}a_E^n - w_{II}a_I^n + h_E]\right)\sum_{m\neq n}[J_{IE}(n,m)a_E^m - J_{II}(n,m)a_I^m]\end{aligned} \tag{6.26b}$$

One can view the synaptic interaction between E–I modules as weak perturbations of the underlying limit cycle oscillators, which suggests carrying out a phase reduction of (6.26) along the lines of Sects. 1.2 and 5.2. In order to proceed, it is first necessary to determine the phase-resetting curve (PRC) of an individual E–I network. Suppose that each uncoupled E–I network operates in a parameter regime where the mean-field equations (6.20) support a stable limit cycle. For concreteness, take a point in parameter space between the two Hopf curves in Fig. 6.2, namely,

$(h_E, h_I) = (0, -4)$. A plot of the oscillation in phase space is shown in Fig. 6.3(a) and the components $Z_E, Z_I$ of the corresponding PRC are shown in Fig. 6.3(b). Note that both components are approximately sinusoidal so that the E–I network acts as a type II oscillator.

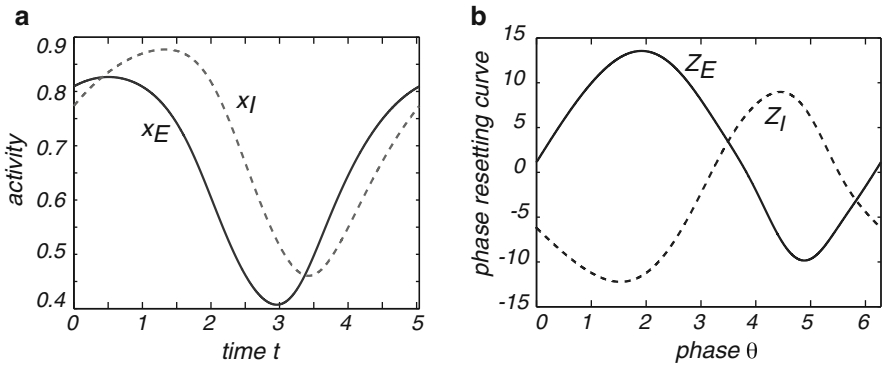


**Fig. 6.3** (**a**) Trajectories along the limit cycle of an E–I network: $x_E(t)$ (*solid curve*) and $x_I(t)$ (*dashed curve*). Parameters are $w_{EE} = 11.5, w_{IE} = -w_{EI} = 10$, $w_{II} = -2$, $h_E = 0$, and $h_I = -4$. Also $F(u) = 1/(1+\mathrm{e}^{-u})$. (**b**) Corresponding components $Z_E$ and $Z_I$ of the phase-resetting curve

The next step is to rewrite (6.26) in the more compact form

$$\frac{d\mathbf{a}^n}{dt} = -\mathbf{a}^n + \mathbf{f}(\mathbf{a}^n) + \varepsilon \mathbf{g}(\mathbf{a}^n)^T \sum_{m \neq n} \mathbf{J}(n,m)\mathbf{a}^m, \tag{6.27}$$

where $\mathbf{a}^n = (a_E^n, a_I^n)^T$, $\mathbf{f} = (f_E, f_I)$, $\mathbf{g} = (g_E, g_I)$,

$$f_p(\mathbf{a}) = F(w_{pE}a_E - w_{pI}a_I + h_p), \quad g_p(\mathbf{a}) = F'(w_{pE}a_E - w_{pI}a_I + h_p), \quad p = E, I.$$

Phase reduction now proceeds using the method of isochrones as described in Sect. 1.2. Let $\Theta(\mathbf{a})$ represent the isochronal mapping in a neighborhood of an uncoupled limit cycle oscillator and denote the natural frequency by $\omega_0$. Then

$$Z_E(\theta) = \frac{\partial \Theta(\mathbf{a}^*(\theta))}{\partial a_E}, \quad Z_I(\theta) = \frac{\partial \theta(\mathbf{a}^*(\theta))}{\partial a_I}, \tag{6.28}$$

where $\mathbf{a}^*$ is a point on the limit cycle. Applying the phase reduction procedure to (6.27) with $\theta_n = \Theta(\mathbf{a}^n)$, we have

$$\frac{d\theta_n}{dt} = \omega_0 + \varepsilon \sum_{p,q=E,I} Z_p(\theta_n) g_p(\theta_n) \sum_m J_{pq}(n,m) a_q^*(\theta_m). \tag{6.29}$$

Here, all quantities are evaluated on the limit cycles so that $g_p(\theta_n) = g_p(a^*(\theta_n))$ etc. Finally, averaging over one period $\Delta_0 = 2\pi/\omega_0$ gives the phase equations

$$\frac{d\theta_n}{dt} = \omega_0 + \varepsilon \sum_{p,q=E,I} \left[ \sum_m J_{pq}(n,m) H_{pq}(\theta_m - \theta_n) \right], \tag{6.30}$$

with phase interaction functions

$$H_{pq}(\phi) = \frac{1}{2\pi}\int_0^{2\pi} Z_p(\theta-\phi)g_p(\theta-\phi)a_q^*(\theta)d\theta. \tag{6.31}$$

Given the phase equations, one can investigate the existence and stability of phase-locked states, including phase waves, along identical lines to the analysis of waves on a chain in Sect. 5.2.2. Note that phase-coupled E–I networks have been used to study the phenomenon of stimulus-induced oscillations and synchrony in primary visual cortex model, where each E–I network is interpreted as a cortical column consisting of reciprocally coupled populations of excitatory and inhibitory neurons [239, 560].

## 6.3 Firing Rates, Asynchronous States, and Spiking Statistics

### *6.3.1 The Asynchronous State in a Homogeneous Spiking Network*

One of the major assumptions in the derivation of rate-based population models in Sect. 6.1 was that each homogeneous population is in an asynchronous state. Here we consider the existence and stability of an asynchronous state in a large, globally coupled network of IF neurons [2, 213]. Consider the following synaptically coupled network of nonlinear IF neurons (see also Sect. 5.3):

$$\frac{dv_i}{dt} = G(v_i) + \frac{\varepsilon}{N}\sum_{j=1}^{N}\int_{-\infty}^{\infty}\Phi(t'-\tau_d)\sum_m \delta(t-t'-T_j^m)dt', \tag{6.32}$$

with $\Phi$ given by the alpha function (5.23), $\tau_d$ a discrete axonal delay and $\varepsilon$ determines the strength of coupling. We take a threshold $v_\kappa = 1$ and a reset $v_r = 0$. In the case of global coupling, we can express the sum over delta functions in terms of the population activity variable

$$a(t) = \sum_{j=1}^{N}\sum_m \delta(t-T_j^m) \tag{6.33}$$

so that

$$\frac{dv_i}{dt} = G(v_i) + \varepsilon u(t), \tag{6.34}$$

where $u(t) = \int_{-\infty}^{\infty}\Phi(t'-\tau_d)a(t-t')dt'$. Now suppose that there exists an asynchronous state $a(t) = a_0$. (Strictly speaking, such a state only exists in the thermodynamic limit $N \to \infty$.) Since $\Phi$ is normalized to unity it follows that $u(t) = a_0$ as well (ignoring transients). An implicit equation for $a_0$ is then obtained by integrating equation (6.34) between successive firing times:

$$\frac{1}{a_0} = \int_0^1 \frac{du}{G(u) + \varepsilon a_0}. \tag{6.35}$$

We will assume that there exists a unique solution to this equation for given $G$ and $\varepsilon$. If $G$ depends on a uniform external input $I$, then this yields the population gain function $F$ with $a_0 = F(I)$.

In order to study the stability of the asynchronous state, it is convenient to carry out the change of variables

$$y_i = a_0 \int_0^{v_i} \frac{du}{G(u) + \varepsilon a_0}, \tag{6.36}$$

with $0 < y_i < 1$ such that Eq. (6.34) becomes

$$\frac{dy_i}{dt} = a_0 + \Gamma(y_i)[u(t) - a_0] \tag{6.37}$$

and

$$\Gamma(y) = \frac{a_0 \varepsilon}{G(y) + a_0 \varepsilon}. \tag{6.38}$$

We also incorporate the effects of synaptic noise by including an additive white noise term $\xi_i(t)$,

$$\frac{dy_i}{dt} = a_0 + \Gamma(y_i)[u(t) - a_0] + \xi_i(t), \tag{6.39}$$

with

$$\langle \xi_i(t) \rangle = 0, \quad \langle \xi_i(t) \xi_j(t') \rangle = \sigma^2 \delta_{ij} \delta(t - t'). \tag{6.40}$$

(Note that diffusive fluctuations of the membrane potential due to stochastic background activity would lead to an additive white noise term in Eq.(6.34) rather than in Eq. (6.40). The corresponding stochastic equation for $y_i$ would then involve multiplicative noise, which is much harder to analyze.) In the presence of noise the variable $y_i$ can become negative so $-\infty < y_i < 1$. The Langevin equation (6.40) has an associated Fokker–Planck equation

$$\frac{\partial}{\partial t} p(y,t) = -\frac{\partial}{\partial y} J(y,t), \tag{6.41}$$

where $J(y,t)$ is the probability flux

$$J(y,t) = [a_0 + \Gamma(y)[u(t) - a_0]]\, p(y,t) - \frac{\sigma^2}{2} \frac{\partial}{\partial y} p(y,t). \tag{6.42}$$

This is supplemented by the boundary conditions arising from reset:

$$p(1,t) = 0, \quad J(1,t) = a(t), \tag{6.43}$$

$$p(0^+,t) = p(0^-,t), \quad J(0^+,t) - J(0^-,t) = a(t). \tag{6.44}$$

We also require $p(-\infty,t)=0$ and $J(-\infty,t)=0$. The steady-state solution of the Fokker–Planck equation is $J(y,t)=a_0$ and $p(y,t)=p_0(y)$ with

$$p_0(y)=\begin{cases} e^{2a_0y/\sigma^2}-e^{2a_0(y-1)/\sigma^2}, & y<0 \\ 1-e^{2a_0(y-1)/\sigma^2}, & 0<y<1 \end{cases}. \tag{6.45}$$

The stability of the steady state can be determined by setting

$$p(y,t)=p_0(y)+\rho(y)e^{\lambda t}, \quad a(t)=a_0+a_1e^{\lambda t} \tag{6.46}$$

and expanding to first order in $\rho, a_1$. This gives the eigenvalue equation

$$\lambda\rho(y)=\frac{\sigma^2}{2}\frac{\partial^2}{\partial y^2}\rho(y)-a_0\frac{\partial}{\partial y}\rho(y)-a_1\tilde{\Phi}(\lambda)\frac{\partial}{\partial y}[\Gamma(y)p_0(y)], \tag{6.47}$$

where $\tilde{\Phi}(\lambda)$ is the Laplace transform

$$\tilde{\Phi}(\lambda)=\int_0^\infty \Phi(t-\tau_d)e^{-\lambda t}=\frac{\alpha^2}{(\lambda+\alpha)^2}e^{\tau_d\lambda}. \tag{6.48}$$

Defining the function

$$h(y)=a_1\tilde{\Phi}(\lambda)\frac{\partial}{\partial y}[\Gamma(y)p_0(y)], \tag{6.49}$$

we can write Eq. (6.47) as the inhomogeneous equation

$$[\mathscr{L}\rho(y)\equiv\left[\frac{\sigma^2}{2}\frac{\partial^2}{\partial y^2}-a_0\frac{\partial}{\partial y}\right]\rho(y)=\lambda\mathbf{1}]\rho(y)+h(y;\lambda). \tag{6.50}$$

This inhomogeneous equation can be solved in terms of the associated one-dimensional Green's function satisfying $[\mathscr{L}-\lambda\mathbf{1}]\mathscr{G}(y,y';\lambda)=\delta(y-y')$ and $\mathscr{G}(1,y',\lambda)=0$:

$$\rho(y)=\int_{-\infty}^{1}\mathscr{G}(y,y';\lambda)h(y';\lambda)dy'-a_1\mathscr{G}(y,0;\lambda) \tag{6.51}$$

with

$$\mathscr{G}(y,y';\lambda)=\begin{cases} A\left(e^{\mu_+(\lambda)[y-1]}-e^{-\mu_-(\lambda)[y-1]}\right)e^{\mu_-(\lambda)[y'-1]}, & y'<y<1 \\ A\left(e^{\mu_-(\lambda)[y'-1]}-e^{-\mu_+(\lambda)[y'-1]}\right)e^{\mu_+(\lambda)[y-1]}, & y<y', \end{cases} \tag{6.52}$$

where

$$A=\frac{2}{\sigma^2}\frac{1}{\mu_++\mu_-}, \tag{6.53}$$

$$\mu_\pm(\lambda)=\frac{1}{\sigma^2}\left[\sqrt{a_0^2+2\lambda\sigma^2}\pm a_0\right]. \tag{6.54}$$

Note that the term $a_1\mathscr{G}(y,0;\lambda)$ ensures that the flux discontinuity at $y=0$ is satisfied. Finally, an implicit equation for the eigenvalues $\lambda$ can be obtained by substituting Eq. (6.49) into (6.51) and imposing the boundary condition $J(1,t)=a(t)$, which corresponds to the following first-order condition:

$$-\frac{\sigma^2}{2}\frac{\partial}{\partial y}\rho(y,t)|_{y=1}=a_1. \tag{6.55}$$

The resulting characteristic equation is [2]

$$\left(e^{\mu_-(\lambda)}-1\right)=\mu_-(\lambda)\tilde{\Phi}(\lambda)\int_{-\infty}^{1}p_0(y)\Gamma(y)e^{\mu_-(\lambda)y}dy. \tag{6.56}$$

In the zero noise limit $\sigma\to 0$, we have $\mu_-(\lambda)\to\lambda/a_0$ and $p_0(y)\to 1$ for $0<y<1$ and is zero otherwise. Thus, Eq. (6.56) becomes

$$\left(e^{\lambda/a_0}-1\right)=\frac{\lambda}{a_0}\tilde{\Phi}(\lambda)\int_{0}^{1}\Gamma(y)e^{\lambda y/a_0}dy. \tag{6.57}$$

In the weak coupling regime, solutions of Eq. (6.57) are of the form $\lambda=2\pi ina_0+\Lambda_n$ for integer $n$ with $\Lambda_n=\mathscr{O}(\varepsilon)$. The term $\Lambda_n$ can be calculated by performing a perturbation expansion in the coupling $\varepsilon$. The lowest-order contribution is simply determined by setting $\lambda=2\pi ina_0$ on the right-hand side of Eq. (6.57). In the case of a linear IF model with $G(u)=I-u$, we have $\Gamma(y)=e^{y/a_0}$ so that

$$\Lambda_n=\varepsilon\left(\frac{2\pi ina_0}{1+2\pi ina_0}\right)\tilde{\Phi}(2\pi ina_0)+\mathscr{O}(\varepsilon^2). \tag{6.58}$$

We then have the following stability results in the absence of noise [2, 213]:

(i) For zero axonal delays ($\tau_d=0$) and excitatory coupling ($\varepsilon>0$), the asynchronous state is stable with respect to excitation of the $n$th mode if and only if $\alpha<\alpha_n$ where

$$\alpha_n=-1+\sqrt{1+4n^2\pi^2a_0^2} \tag{6.59}$$

Hence, it is stable for sufficiently slow synapses, that is, $\alpha<\alpha_1$. The asynchronous state is always unstable in the case of inhibitory coupling since the condition for stability with respect to the $n$th harmonic is now $\alpha>\alpha_n$, which cannot be satisfied for all $n$.

(ii) The asynchronous state is almost always unstable for nonzero delays (in the noise-free case).

(iii) For large $n$, $|\Lambda_n|\sim 1/n^2$ so that higher harmonics grow or decay slowly.

Note that although the zero delay case is a singular limit in the absence of noise, it becomes non-singular for arbitrarily small amounts of noise, where instabilities

with respect to higher harmonics are suppressed [2, 213]. One finds that for sufficiently high noise levels the asynchronous state is always stable. Reducing the noise for fixed delay induces an instability due to excitation of one of the harmonic modes with frequency $\omega \approx \omega_n = 2\pi n a_0$. A bifurcation at $\omega \approx \omega_1$ implies that the period of the resulting collective oscillation is identical to the period of the individual oscillators. Higher harmonics correspond to instabilities of the asynchronous state that lead to the formation of cluster states [213, 231]: each neuron fires with mean rate $a_0$, but the population of neurons splits up into several groups that fire in sequence so that the overall activity exhibits faster oscillations. Finally, note that fast oscillations are also found in sparsely connected random networks [87].

### *6.3.2 Balanced Networks and Neural Variability*

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 [179, 593]. Indeed, one observes trial-to-trial variability in spike trains, even across trials in which external stimuli are identical. This raises a number of important issues. First, 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 (Sect. 1.5) or random synaptic background activity is taken into account. For example, in the homogeneous spiking network analyzed above, the spike trains of individual neurons can be quite regular even when the population activity is asynchronous. Conversely, irregularity in spiking at the single-cell level can coexist with regular firing rates at the population level. However, there is growing evidence that noise can play a significant role at the network level. For example, noise appears to be present during perceptual decision making [656] and bistable perception, the latter being exemplified by perceptual switching during binocular rivalry [436, 578, 663]; see Sect. 8. Noise also contributes to the generation of spontaneous activity during resting states [148, 152]. At the level of large-scale neural systems, as measured with functional MRI (fMRI) imaging, this ongoing spontaneous activity reflects the organization of a series of highly coherent functional networks that may play an important role in cognition. Another issue is the possible computational role of pairwise and higher-order statistical correlations between the spike trains of distinct neurons, that is, whether or not it is necessary to go beyond firing rates.

One paradigm for exploring these various issues is the so-called balanced network [567, 645, 650]. 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 [391]. 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 [519]. 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. (Such rates could be a measure of population activity as in Sect. 6.1 or correspond to the rates of individual neurons modeled as inhomogeneous Poisson processes; see below.) There is then another level of neural variability, namely, trail-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 spiking variability [119]; see Fig. 6.4. Litwin-Kumar and Doiron have recently shown how these two levels of stochastic variability can emerge in a chaotic balanced network of randomly connected spiking neurons, in which a small amount of clustered connections induces firing rate fluctuations superimposed on spontaneous spike fluctuations [387].

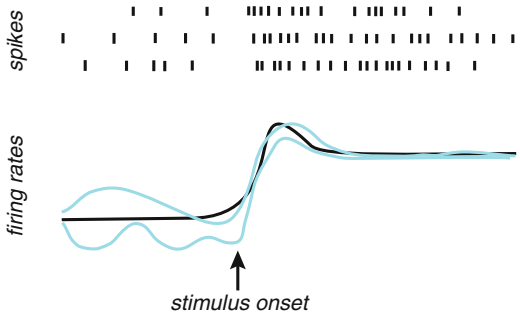


**Fig. 6.4** Neurons can exhibit double stochasticity. Trial-to-trial variability in spike trains (shown as raster plots) can coexist with trial-to-trial variability in the firing rate (shown as continuous *gray curves*). Trial-averaged firing rate is the *black curve*. Following stimulus onset, the variability in the firing rate is reduced

### *6.3.3 Spike Statistics and the Poisson Process*

Given the ubiquity of Poisson processes in spiking neuron models and its relevance to neural master equations (see Sect. 6.4), it is useful to consider these processes in a little more detail. Suppose that a neuron emits a sequence of spikes at times $t_1, t_2, \ldots, t_n$. One way to characterize this sequence statistically is in terms of the probability density $\rho(t_1, \ldots t_n)$ of finding such a sequence over many experimental trials. In other words, the probability of having a sequence of $n$ spikes in the interval $[0, T]$ with the $i$th spike falling between the times $t_i$ and $t_i + \Delta t$

is $P[t_1,\ldots t_n] = \rho(t_1,\ldots t_n)\Delta t^n$. In principle, the probability of an event occurring, namely, the firing of the next spike, could depend on the entire history of past spikes. If this dependence only extends to the previous spike so that the interspike intervals are statistically independent, then the stochastic process is said to be a *renewal process*. If there is no dependence at all on preceding spikes so that the firing times are themselves independent, then we have a *Poisson process*.

Consider a sequence of spikes generated by a homogeneous Poisson process, that is, one with a time-independent mean firing rate $r$. Divide a given time interval $T$ into $M$ bins of size $\Delta t = T/M$ and assume that $\Delta T$ is small enough so that the probability of finding two spikes within any one bin can be neglected. Then the probability $P_T[n]$ of finding $n$ spikes over the interval $T$ is given by

$$P_T[n] = \lim_{\Delta t\to 0} \frac{M!}{(M-n)!n!}(r\Delta t)^n(1-r\Delta t)^{M-n}$$

This consists of the probability $(r\Delta t)^n$ of finding n spikes in $n$ specific bins multiplied by the probability $(1-r\Delta t)^{M-n}$ of not finding spikes in the remaining bins. The binomial factor is the number of ways of choosing $n$ out of $M$ bins with spikes. Using the approximation $M-n \approx M = T/\Delta t$ and defining $\varepsilon = -r\Delta t$, we have that

$$\lim_{\Delta t\to 0}(1-r\Delta t)^{M-n} = \lim_{\varepsilon\to 0}\left((1+\varepsilon)^{1/\varepsilon}\right)^{-rT} = \mathrm{e}^{-rT}.$$

For large $M$, $M!/(M-n)! \approx M^n = (T/\Delta t)^n$, so that we obtain the Poisson distribution

$$P_T[n] = \frac{(rT)^n}{n!}\mathrm{e}^{-rT}$$

Given that there are $n$ independent spikes over the interval $[0,T]$, the probability that these spikes lie within specified bins of size $\Delta t$ is $n!(\Delta t/T)^n$. Hence

$$\rho(t_1,\ldots t_n) = n!\left(\frac{1}{T}\right)^n P_T[n] = r^n\mathrm{e}^{-rT} \tag{6.60}$$

Also note that the fastest way to generate a sequence of Poisson spikes for constant $r$ is to iterate the firing times $t_{n+1} = t_n - \log(x_{rand})/r$ with $x_{rand}$ uniformly distributed over $[0,1]$.

A simple method for calculating the moments of the Poisson distribution is to introduce the moment-generating function

$$g(s) = \sum_{n=0}^{\infty} P_T[n]\mathrm{e}^{sn}$$

Differentiating with respect to $s$ shows that

$$\left.\frac{d^k g(s)}{ds^k}\right|_{s=0} = \langle n^k\rangle$$

The generating function for the Poisson process can be evaluated explicitly as

$$g(s) = \exp(-rT)\exp(rT\mathrm{e}^{s})$$

from which we deduce that

$$\langle n \rangle = rT, \quad \sigma_n^2 = rT$$

Another useful quantity is the interspike interval (ISI) distribution. Suppose that a spike was last fired at time $T^n$. The probability of a homogeneous Poisson process generating the next spike in the interval $T^n + \tau \leq T^{n+1} \leq T^n + \tau + \Delta\tau$ is equal to the probability that no spike is fired for a time $\tau$, which is $P_\tau[0] = \mathrm{e}^{-r\tau}$ multiplied by the probability $r\Delta\tau$ of generating a spike within the following interval $\Delta\tau$:

$$\mathbf{Pr}[\tau \leq T^{n+1} - T^n \leq \tau + \Delta\tau] = r\Delta\tau\mathrm{e}^{-r\tau}$$

The ISI probability density is thus an exponential, $\rho(\tau) = r\mathrm{e}^{-r\tau}$. It follows that the mean interspike interval is

$$\langle \tau \rangle = \int_0^\infty r\mathrm{e}^{-r\tau}\tau d\tau = \frac{1}{r}$$

and the variance is

$$\sigma_\tau^2 = \int_0^\infty r\mathrm{e}^{-r\tau}\tau^2 d\tau - \langle \tau \rangle^2 = \frac{1}{r^2}$$

The ratio of the standard deviation to the mean is called the *coefficient of variation*

$$C_V = \frac{\sigma_\tau}{\langle \tau \rangle}$$

It follows that for a homogeneous Poisson process $C_V = 1$.

How well do Poisson statistics describe experimentally measured spike trains? One often finds that for ISIs longer than about 10 msec, the ISI distribution is indeed approximately exponential. However, for shorter intervals, there is a rapid decrease in the distribution reflecting the fact that neurons are refractory immediately after

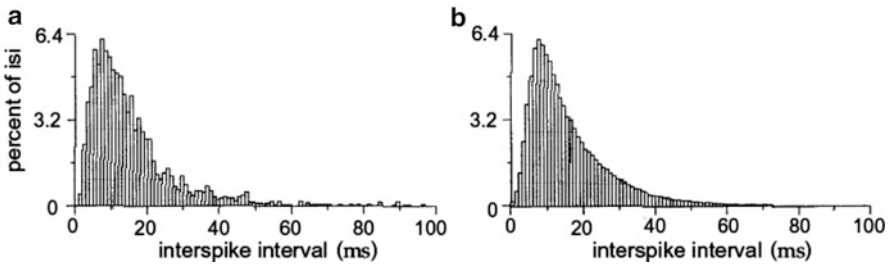


**Fig. 6.5** (**a**) Interspike interval distribution from a neuron from the MT cortical visual area of a macaque monkey responding to a moving image. (**b**) Interspike interval generated with a Poisson model with a stochastic refractory period. Redrawn from [20]

firing. This is illustrated in Fig. 6.5. The data can be fitted more accurately by a gamma distribution

$$p[\tau] = \frac{r(r\tau)^k \mathrm{e}^{-t\tau}}{k!} \tag{6.61}$$

Alternatively, one can introduce a refractory period into a standard Poisson model. Note that $C_V$ values extracted from the spike trains of cortical neurons also take values around unity, provided that the mean interspike interval is not too small [593].

It is possible to generalize the above Poisson model to the case of a time-dependent rate $r(t)$. The simplest way to analyze this inhomogeneous Poisson process is to consider the probability distribution $\rho(t_1, \ldots t_n)$. This is given by the product of the probabilities $r(t_i)\Delta t$ that the neuron fires within the time intervals $t_i \leq t \leq t_i + \Delta t$ and the probabilities of not firing during the interspike intervals. The latter is given by

$$\mathbf{Pr}[\text{no spikes in } (t_i, t_{i+1})] = \prod_{m=1}^{M} (1 - r(t_i + m\Delta t)\Delta t)$$

where we have partitioned the interval $(t_i, t_{i+1})$ into $M$ bins of size $\Delta t$. Taking the logarithm,

$$\log \mathbf{Pr}[\text{no spikes in } (t_i, t_{i+1})] = \sum_{m=1}^{M} \log(1 - r(t_i + m\Delta t)\Delta t) \approx -\sum_{m=1}^{M} r(t_i + m\Delta t)\Delta t$$

Taking the limit $\Delta t \to 0$ and exponentiating again shows that

$$\mathbf{Pr}[\text{no spikes in } (t_i, t_{i+1})] = \exp\left(-\int_{t_i}^{t_{i+1}} r(t)dt\right)$$

Hence

$$\rho(t_1, \ldots t_n) = \prod_{i=1}^{n} r(t_i) \exp\left(-\int_0^T r(t)dt\right) \tag{6.62}$$

In the case of a time-dependent rate, one generates $x_{rand}$ at each time step and a spike is fired if $r(t)\Delta t > x_{rand}$.

## 6.4 Stochastic Population Models

The discussion of neural variability in Sect. 6.3 motivates the incorporation of noise directly into rate-based models, without explicitly modeling spike variability. One approach is to introduce noise into a rate-based network model using extrinsic noise sources [21, 84, 184, 297]. An alternative approach is to assume that noise arises intrinsically as a collective population effect. A number of methods involve carrying out some form of dimension reduction of a network of synaptically coupled spiking neurons. These include population density methods [395, 468, 477], mean-field the-

ories [21, 86, 87], and Boltzmann-like kinetic theories [90, 100, 517]. However, such methods tend to consider either fully connected or sparsely connected networks and simplified models of spiking neurons such as the integrate-and-fire (IF) model. Nevertheless, as discussed in Sect. 6.3, one important feature of spiking networks is that they can support an asynchronous state in which individual neurons exhibit Poisson-like statistics, whereas the total population activity can be expressed in terms of a characteristic activation or gain function [2, 86, 87, 213, 519]. Formally speaking, the asynchronous state only exists in the thermodynamic limit $N \to \infty$, where $N$ determines the size of the population. This then suggests a possible source of intrinsic noise at the network level arises from fluctuations about the asynchronous state due to finite size effects [57, 220, 412, 425, 597]; this is distinct from intrinsic noise at the single-cell level due to channel fluctuations and it is assumed that the latter is negligible at the population level. The presence of finite size effects has motivated the development of a closer analogy between intrinsic noise in biochemical networks (including ion channel kinetics) and neural networks [68, 69, 72], based on extensions of the neural master equation introduced by Buice et al. [91, 92]; see also [473]. In this section, we describe the master equation framework for stochastic neural population dynamics.

### 6.4.1 Neural Master Equations

One way to incorporate intrinsic noise at the population level is to treat the output activity of a local homogeneous population as a discrete stochastic variable $A_\alpha(t)$ rather than the instantaneous firing rate $a_\alpha = F(u_\alpha)$ [68, 91, 92]:

$$A_\alpha(t) = \frac{N_\alpha(t)}{N\Delta t}, \tag{6.63}$$

where $N_\alpha(t)$ is the number of neurons in the $\alpha$th population that fired in the time interval $[t-\Delta t, t]$ and $\Delta t$ is the width of a sliding window that counts spikes. The discrete stochastic variables $N_\alpha(t)$ are taken to evolve according to a one-step jump Markov process:

$$N_\alpha(t) \to N_\alpha(t) \pm 1 : \quad \text{transition rate } \omega_\pm(U_\alpha(t), N_\alpha(t)), \tag{6.64}$$

with the synaptic current $U_\alpha(t)$ given by (for exponential synapses)

$$\tau dU_\alpha(t) = \left[ -U_\alpha(t) + \sum_{\beta=1}^{M} w_{\alpha\beta} A_\beta(t) \right] dt, \tag{6.65}$$

where, for convenience, we have rescaled the weights according to $w_{\alpha\beta} \to w_{\alpha\beta}/\tau$. The transition rates are taken to be (cf. [68])

$$\omega_+(u_\alpha,n_\alpha)\to\omega_+(u_\alpha)=\frac{N\Delta t}{\tau_a}F(u_\alpha),\quad \omega_-(u_\alpha,n_\alpha)\to\omega_-(n_\alpha)=\frac{n_\alpha}{\tau_a}. \tag{6.66}$$

We see that the transition rate $\omega_+$ depends on $U_\alpha$, with the latter itself coupled to the associated jump Markov according to Eq. (6.65), which is only defined between jumps, during which $U_\alpha(t)$ evolves deterministically. Hence, the resulting stochastic process defined by Eqs. (6.63), (6.64), (6.65) and (6.66) provides an example of a stochastic hybrid system at the population neuron level, just as stochastic ion channel gating provides an example at the molecular level (Sect. 1.5). It is important to note that the time constant $\tau_a$ 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.

**A. Case $\tau\to 0$ (fast synapses).** In the limit $\tau\to 0$, Eq. (6.65) implies that the continuous variables $U_\alpha(t)$ can be eliminated by setting $U_\alpha(t)=\sum_\beta w_{\alpha\beta}A_\alpha(t)$. This then leads to a pure birth–death process for the discrete variables $N_\alpha(t)$. That is, let $P(\mathbf{n},t)=\text{Prob}[\mathbf{N}(t)=\mathbf{n}]$ 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 $P(\mathbf{n},0)$ with $0\le n_\alpha\le N$. The probability distribution then evolves according to the birth–death master equation [68, 91, 92]

$$\frac{dP(\mathbf{n},t)}{dt}=\sum_\alpha\left[(\mathbb{T}_\alpha-1)\left(\omega_\alpha^-(\mathbf{n})P(\mathbf{n},t)\right)+(\mathbb{T}_\alpha^{-1}-1)\left(\omega_\alpha^+(\mathbf{n})P(\mathbf{n},t)\right)\right], \tag{6.67}$$

where

$$\omega_\alpha^+(\mathbf{n})=\frac{N\Delta t}{\tau_a}F\left(\sum_\beta w_{\alpha\beta}n_\alpha/N\Delta t\right),\quad \omega_\alpha^-(\mathbf{n})=\frac{n_\alpha}{\tau_a}, \tag{6.68}$$

and $\mathbb{T}_\alpha$ is a 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 $n_\alpha$ replaced by $n_\alpha\pm 1$. Equation (6.67) is supplemented by the boundary conditions $P(\mathbf{n},t)\equiv 0$ if $n_\alpha=N+1$ or $n_\alpha=-1$ for some $\alpha$. The neural master equation (6.67) has two versions depending on the choice of $\Delta t$. First, fixing $\Delta t=1$ leads to the Bressloff version of the master equation. For large but finite $N$, the master equation (6.67) can be approximated by a Fokker–Planck equation using a Kramers–Moyal or system-size expansion, so that the population activity $A_\alpha$ evolves according to a Langevin equation [68]. Introduce the rescaled variables $a_\alpha=n_\alpha/N$ and corresponding transition rates $\Omega_\alpha^-(\mathbf{a})=a_\alpha/\tau_a$ and $\Omega_\alpha^+(\mathbf{a})=\tau_\alpha^{-1}F\left(\sum_\beta w_{\alpha\beta}a_\beta\right)$. Carrying out a Kramers–Moyal expansion to second order then leads to the multivariate FP equation

$$\frac{\partial P(\mathbf{a},t)}{\partial t}=-\sum_{\alpha=1}^M\frac{\partial}{\partial a_\alpha}\left[\mathscr{A}_\alpha(\mathbf{a})P(\mathbf{a},t)\right]+\frac{1}{2N}\sum_{\alpha=1}^M\frac{\partial^2}{\partial a_\alpha^2}\left[\mathscr{B}_\alpha(\mathbf{a})P(\mathbf{a},t)\right] \tag{6.69}$$

with

$$\mathscr{A}_\alpha(\mathbf{a})=\Omega_\alpha^+(\mathbf{a})-\Omega_\alpha^-(\mathbf{a}),\quad \mathscr{B}_\alpha(\mathbf{a})=\Omega_\alpha^+(\mathbf{a})+\Omega_\alpha^-(\mathbf{a}). \tag{6.70}$$

The solution to the Fokker–Planck equation (6.69) determines the probability density function for a corresponding stochastic process $\mathbf{A}(t)=(A_1(t),\ldots,A_M(t))$, which evolves according to a neural Langevin equation or SDE of the form

$$dA_\alpha = \mathscr{A}_\alpha(\mathbf{A})dt + \frac{1}{\sqrt{N}} b_\alpha(\mathbf{A})dW_\alpha(t). \tag{6.71}$$

with $b_\alpha(\mathbf{x})^2 = \mathscr{B}_\alpha(\mathbf{x})$. Here $W_\alpha(t)$ denotes an independent Wiener process such that

$$\langle dW_\alpha(t)\rangle = 0, \quad \langle dW_\alpha(t)dW_\beta(s)\rangle = \delta_{\alpha,\beta}\delta(t-s)dtds. \tag{6.72}$$

In the thermodynamic limit $N \to \infty$, we recover the activity-based mean-field equation

$$\tau_\alpha \frac{da_\alpha}{dt} = \mathscr{A}_\alpha(\mathbf{a}) = -a_\alpha(t) + F(\sum_\beta w_{\alpha\beta}a_\alpha(t)). \tag{6.73}$$

Note that the multiplicative noise in Eq. (6.71) is interpreted in the sense of Ito, which follows from the form of the FP equation (6.69); see Sect. 1.7.

A rigorous probabilistic treatment of the thermodynamic limit of the neural master equation has also been developed [89], extending previous work on chemical master equations [351]. However, as we noted within the context of stochastic ion channels in Sect. 1.5, the diffusion approximation breaks down when considering noise-induced transitions between multistage states. It is then necessary to use a WKB approximation of the master equation together with matched asymptotics [69]. The second version of the neural master equation, which was introduced by Buice et al. [91, 92], is obtained by taking the limit $N \to \infty, \Delta t \to 0$ such that $N\Delta t = 1$. In this case there is no small parameter that allows one to construct a Langevin approximation to the master equation. Nevertheless, it is possible to determine the moment hierarchy of the master equation using path integral methods or factorial moments, based on the observation that the network operates in a Poisson-like regime. The role of the sliding window size $\Delta t$ is crucial in understanding the difference between the two versions of the master equation. First, it should be emphasized that the stochastic models are keeping track of *changes* in population spiking activity. If the network is operating close to an asynchronous state for large $N$, then one-step changes in population activity could occur relatively slowly, so there is no need to take the limit $\Delta t \to 0$. On the other hand, if population activity is characterized by a Poisson process, then it is necessary to take the limit $\Delta t \to 0$ in order to maintain a one-step process. However, given the existence of an arbitrarily small time scale $\Delta t$, it is no longer clear that one is justified in ignoring synaptic dynamics by taking the limit $\tau \to 0$ in Eq. (6.65).

**B. Case $\tau \gg \tau_a > 0$ (stochastic hybrid system).** Now suppose that $\tau > 0$ in the full stochastic model given by (6.63)–(6.66), with $N \to \infty$, $\Delta t \to 0$ such that $N\Delta t = 1$. Denote the random state of the full model at time $t$ by $\{(U_\alpha(t), N_\alpha(t)); \alpha = 1,\ldots,M\}$. Introduce the corresponding probability density

$$\text{Prob}\{U_\alpha(t) \in (u_\alpha, u_\alpha + du), N_\alpha(t) = n_\alpha; \alpha = 1, \ldots, M\} = p(\mathbf{u}, \mathbf{n}, t) d\mathbf{u}, \quad (6.74)$$

with $\mathbf{n} = (n_1, \ldots, n_M)$ and $\mathbf{u} = (u_1, \ldots, u_M)$. It follows from Eqs. (6.63)–(6.66) that the probability density evolves according to the Chapman–Kolmogorov (CK) equation [82]

$$\begin{aligned} &\frac{\partial p}{\partial t} + \frac{1}{\tau} \sum_\alpha \frac{\partial [v_\alpha(\mathbf{u}, \mathbf{n}) p(\mathbf{u}, \mathbf{n}, \mathbf{t})]}{\partial u_\alpha} \\ &\quad = \frac{1}{\tau_a} \sum_\alpha \left[ (\mathbb{T}_\alpha - 1)(\omega_-(n_\alpha) p(\mathbf{u}, \mathbf{n}, t)) + (\mathbb{T}_\alpha^{-1} - 1)(\omega_+(u_\alpha) p(\mathbf{u}, \mathbf{n}, t)) \right], \end{aligned} \quad (6.75)$$

with

$$\omega_+(u_\alpha) = F(u_\alpha), \quad \omega_-(n_\alpha) = n_\alpha, \quad v_\alpha(\mathbf{u}, \mathbf{n}) = -u_\alpha + \sum_\beta w_{\alpha\beta} n_\beta. \quad (6.76)$$

Equation (6.75) can be rewritten in the more compact form (see also (1.143))

$$\frac{\partial p}{\partial t} = -\frac{1}{\tau} \sum_{\alpha=1}^{M} \frac{\partial}{\partial u_\alpha} (v_\alpha(\mathbf{u}, \mathbf{n}) p(\mathbf{u}, \mathbf{n}, t)) + \frac{1}{\tau_a} \sum_{\mathbf{m}} A(\mathbf{n}, \mathbf{m}; \mathbf{u}) p(\mathbf{u}, \mathbf{m}, t). \quad (6.77)$$

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

$$\tau \frac{du_\alpha}{dt} = v_\alpha(\mathbf{u}, \mathbf{n}), \quad \alpha = 1, \ldots, M, \quad (6.78)$$

and $A$ represents the $\mathbf{u}$-dependent transition matrix for the jump Markov process.

For fixed $\mathbf{u}$, the matrix $A(\mathbf{n}, \mathbf{m}; \mathbf{u})$ is irreducible (which means that there is a nonzero probability of transitioning, possibly in more than one step, from any state to any other state in the jump Markov process) and has a simple zero eigenvalue. In particular, $\sum_{\mathbf{n}} A(\mathbf{n}, \mathbf{m}; \mathbf{u}) = 0$ for all $\mathbf{m}$, that is, $\mathbf{n} = (1, 1, \ldots, 1)^T$ is the left null vector of $A$. The Perron–Frobenius theorem (for large but finite $N$) ensures that all other eigenvalues of $A$ are negative and the continuous-time Markov process for fixed $\mathbf{u}$,

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

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 steady-state equation is

$$\begin{aligned} 0 &= \sum_{\mathbf{m}} A(\mathbf{n}, \mathbf{m}; \mathbf{u}) \rho(\mathbf{u}, \mathbf{m}) \\ &= \sum_{\alpha=1}^{M} \left[ (n_\alpha + 1) \rho(\mathbf{u}, \mathbf{n} + \mathbf{e}_\alpha) - n_\alpha \rho(\mathbf{u}, \mathbf{n}) + F(u_\alpha)(\rho(\mathbf{u}, \mathbf{n} - \mathbf{e}_\alpha) - \rho(\mathbf{u}, \mathbf{n})) \right], \end{aligned}$$

where $[\mathbf{e}_\alpha]_\beta = \delta_{\alpha,\beta}$. The solution can be factorized as $\rho(\mathbf{u},\mathbf{n}) = \prod_{\beta=1}^{M} \rho_1(u_\beta, n_\beta)$ with

$$0 = \sum_{\alpha=1}^{M} \left[ \prod_{\beta\neq\alpha} \rho_1(u_\beta, n_\beta) \right] [J(u_\alpha, n_\alpha + 1) - J(u_\alpha, n_\alpha)],$$

where

$$J(u,n) = n\rho_1(u,n) - F(u)\rho_1(u,n-1).$$

Since $\rho_1(u,-1) \equiv 0$, it follows that $J(u,n) = 0$ for all $n$. Hence,

$$\rho_1(u,n) = \rho(u,0) \prod_{m=1}^{n} \frac{F(u)}{m} = \rho(u,0) \frac{F(u)^n}{n!}, \tag{6.79}$$

and the corresponding normalized density is a Poisson process with rate $F(u)$

$$\rho_1(u,n) = \mathrm{e}^{-F(u)} \frac{F(u)^n}{n!}. \tag{6.80}$$

There are two time scales in the CK equation (6.75), the synaptic time constant $\tau$ and the time constant $\tau_a$, which characterizes the relaxation rate of population activity. In the limit $\tau \to 0$ for fixed $\tau_a$, Eq. (6.75) reduces to the neural master equation (6.67) with $\mathbf{u} = \mathbf{u}(\mathbf{n})$ such that $v_\alpha(\mathbf{u}(\mathbf{n}),\mathbf{n}) = 0$. On the other hand, if $\tau_a \to 0$ for fixed $\tau$, then we obtain deterministic voltage- or current-based mean-field equations

$$\begin{aligned} \tau \frac{du_\alpha}{dt} &= \langle v_\alpha \rangle(\mathbf{u}(t)) \equiv \sum_{\mathbf{n}} v_\alpha(\mathbf{u}(t),\mathbf{n}) \rho(\mathbf{u}(t),\mathbf{n}) \\ &= -u_\alpha(t) + \sum_{\beta=1}^{M} w_{\alpha\beta} \sum_{\mathbf{n}} n_\beta \rho(\mathbf{u}(t),\mathbf{n}). \end{aligned} \tag{6.81}$$

Since $\rho(\mathbf{u},\mathbf{n})$ is given by product of independent Poisson processes with rates $F(u_\alpha)$, consistent with the operating regime of the Buice *et. al.* master equation [91, 92], it follows that

$$\langle n_\beta \rangle = F(u_\beta), \tag{6.82}$$

and (6.81) reduces to the standard voltage- or current-based activity equation. Now suppose that the network operates in the regime $0 < \tau_a/\tau \equiv \varepsilon \ll 1$, for which 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 suggests 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. The resulting perturbations can be analyzed using a quasi-steady-state (QSS) diffusion or adiabatic approximation, in which the CK equation (6.75) is approximated by a Fokker–Planck equation [82]; see Sect. 1.6. However, when considering escape from a metastable state, it is necessary to use the WKB method outlined in Sect. 1.6, as will be illustrated below.

### 6.4.2 Metastability in a One-Population Model

One of the useful features of the master equation formulation of stochastic population dynamics is that one can apply methods previously used to analyze jump Markov processes at the cellular level. We will illustrate this by considering a first passage time (FPT) problem for a one-population version of the CK equation (6.75), which exhibits bistability in the deterministic limit [82]. This is analogous to the problem of spontaneous action potential generation presented in Sect. 1.6. In the case of a single homogeneous population of excitatory neurons, (6.75) becomes

$$\frac{\partial p}{\partial t} + \frac{\partial [v(u,n)p(u,n,t)]}{\partial u} = \frac{1}{\varepsilon}\sum_m A(n,m;u)p(u,m,t) \tag{6.83}$$

with drift term

$$v(u,n) = -u+n, \tag{6.84}$$

and tridiagonal transition matrix

$$A(n,n-1;u) = F(u), \quad A(n,n;u) = -F(u)-n, \quad A(n,n+1;u) = n+1. \tag{6.85}$$

As shown in Sect. 6.4.1, the steady-state density for $A$ is given by a Poisson process,

$$\rho(u,n) = \frac{[F(u)]^n \mathrm{e}^{-F(u)}}{n!}, \tag{6.86}$$

and the mean-field equation obtained in the $\varepsilon \to 0$ limit is

$$\frac{du}{dt} = \sum_{n=0}^{\infty} v(u,n)\rho(u,n) = -u+F(u). \tag{6.87}$$

If $F$ is given by the sigmoid (1.14), it is straightforward to show that (6.87) exhibits bistability for a range of thresholds and gains; see Fig. 6.6.

The CK equation (6.83) is identical in form to (1.143) under the change of variables $u \to v$, where $v$ is voltage, and $v(u,n) \to F(v,n)$. It follows that the general analysis developed in Sect. 1.6 can be applied directly to the stochastic neural population model. In particular, the escape rate $\lambda_0$ from the low activity state $u_-$ is given by (1.191). As in the case of stochastic ion channels, there are three basic steps needed in order to evaluate $\lambda_0$ using the particular form of the drift $v$ and transition matrix $A$ [72]:

(1) Find the unique nontrivial positive eigenfunction $\psi_n(u) = R^{(0)}(u,n)$ and associated eigenvalue $\mu(u) = -\Phi_0'(u)$. In the case of the neural population model, Eq. (1.171) takes the explicit form

$$F(u)\psi_{n-1}(u) - (F(u)+n)\psi_n(u) + (n+1)\psi_{n+1}(u) = \mu(-u+n)\psi_n(u). \tag{6.88}$$

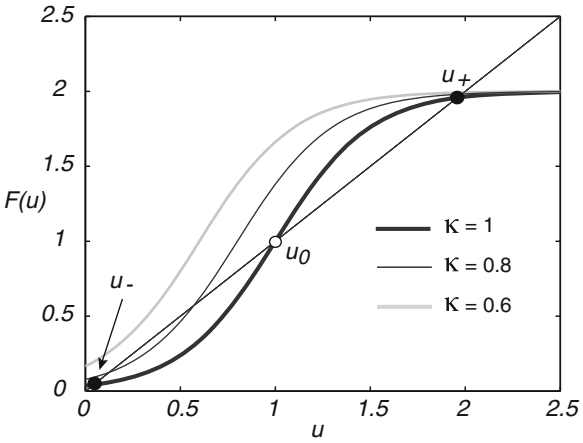


**Fig. 6.6** Bistability in the deterministic network satisfying $\dot{u} = -u + F(u)$ with $F$ given by the sigmoid (1.14) for gain $\eta = 4$ and threshold $\kappa = 1.0$, $F_0 = 2$. There exist two stable fixed points $u_\pm$ separated by an unstable fixed point $u_0$. As the threshold $\kappa$ is reduced the network switches to a monostable regime

Trying a solution for $\psi$ of the form

$$\psi_n(u) = \frac{\Lambda(u)^n}{n!} \tag{6.89}$$

yields the following equation relating $\Lambda$ and the corresponding eigenvalue $\mu$:

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

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

$$\mu = \left[\frac{F(u)}{\Lambda} - 1\right], \quad \Lambda = F(u) - \mu u.$$

We deduce that

$$\Lambda = u, \quad \mu = \left[\frac{F(u)}{u} - 1\right], \tag{6.90}$$

and the normalized eigenfunction is

$$\psi_n(u) = \frac{u^n}{n!} \mathrm{e}^{-u}. \tag{6.91}$$

Note that $\mu(u)$ vanishes at the fixed points $u_-, u_*$ of the mean-field equation (6.87) with $\mu(u) > 0$ for $0 < u < u_-$ and $\mu(u) < 0$ for $u_- < u < u_*$. Moreover, comparing Eq. (6.86) with (6.91) establishes that $\psi_n(u) = \rho(u,n)$ at the fixed points $u_*, u_\pm$. In conclusion $R^{(0)}(u,n) = \psi_n(u)$ and the effective potential $\Phi_0$ is given by

$$\Phi_0(u) = -\int_{u_-}^{u^*} \mu(y)dy. \tag{6.92}$$

The effective potential is defined up to an arbitrary constant, which has been fixed by setting $\Phi_0(u_-) = 0$.

(2) Determine the null eigenfunction $\eta_n(u) = S(u,n)$ of Eq. (1.174), which becomes

$$F(u)\eta_{m+1} - (F(u)+m)\eta_m + m\eta_{m-1} = \mu(u)[-u+m]\eta_m. \tag{6.93}$$

Trying a solution of the form $\eta_m = \Gamma^m$ yields

$$(F(u))\Gamma - (F(u)+m) + m\Gamma^{-1} = \mu(u)[-u+m]. \tag{6.94}$$

$\Gamma$ is then determined by canceling terms linear in $m$, which finally gives

$$\eta_n(u) = \left(\frac{u}{F(u)}\right)^n. \tag{6.95}$$

(3) Calculate the generalized eigenvector $\zeta_n = \hat{S}(u_*,n)$ of Eq. (1.186), which reduces to

$$F(u_*)\zeta_{n+1} + n\zeta_{n-1} - (F(u_*)+n)\zeta_n = u_* - n. \tag{6.96}$$

It is straightforward to show that this has the solution $\zeta_n = n$ (up to an arbitrary constant that does not contribute to the principal eigenvalue). It follows from Eq. (6.86) that the factor $B(u_*)$ defined by (1.190) is

$$\begin{aligned} B(u_*) &= \sum_{n=0}^{\infty} \rho(u_*,n)\left[-u_*n + n^2\right] \\ &= \left[-u_*\langle n\rangle + \langle n^2\rangle\right]. \end{aligned} \tag{6.97}$$

Recall that $\rho(u,n)$ is given by a Poisson density with rate $F(u)$, which implies that $\langle n^2\rangle = \langle n\rangle + \langle n\rangle^2$ with $\langle n\rangle = F(u)$. Therefore,

$$B(u_*) = F(u_*)\left[2F(u_*) - u_*\right], \tag{6.98}$$

which reduces to $B(u_*) = F(u_*)^2$ since $u^* = F(u^*)$ at a fixed point.

It is instructive to compare the effective potential $\Phi_0$ obtained using the WKB approximation with the potential obtained using the FP equation (1.146) based on the QSS approximation. First, substitute (6.86), (6.84), and (6.85) into Eqs. (1.147) and (1.148) under the change of variables $u \to v$ and $v(u,n) \to F(v,n)$. We find that $Z(u,n) = n\rho(u,n)$ so that

$$\mathscr{D}(u) = [-u\langle n\rangle + w\langle n^2\rangle] = B(u). \tag{6.99}$$

The steady-state solution of the FP equation (1.146) takes the form $C(u) \sim \exp^{-\hat{\Phi}_0(u)/\varepsilon}$ with stochastic potential

$$\hat{\Phi}_0(u) = -\int^u \frac{\mathscr{F}(y)}{\mathscr{D}(y)}dy = -\int^u \frac{-y+F(y)}{wF(y)[2F(y)-y]}dy. \tag{6.100}$$

Note that $\hat{\Phi}_0$ differs from the potential $\Phi_0$, Eq. (6.92), obtained using the more accurate WKB method. Equations (6.90) and (6.92) show that the latter has the integral form

$$\Phi_0(u) = -\int^u \frac{1}{w}\left[\frac{wF(y)}{y} - 1\right]dy. \tag{6.101}$$

Thus, there will be exponentially large differences between the steady states for small $\varepsilon$.

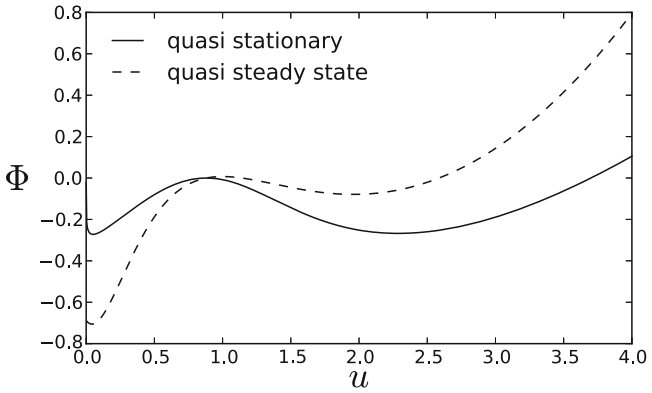


**Fig. 6.7** Comparison of the double-well potentials $\Phi_0(u)$ and $\hat{\Phi}_0(u)$ obtained using the quasistationary approximation and the QSS diffusion approximation, respectively [82]. Parameter values are chosen so that deterministic network is bistable: $F_0 = 2.3$, $\gamma = 4$, and $\kappa = 1$

In Fig. 6.7, we plot the potential function $\Phi_0$ of Eq. (6.101), which is obtained using the quasistationary approximation in a parameter regime for which the underlying deterministic network is bistable. We also plot the corresponding potential function $\hat{\Phi}_0$ of Eq. (6.100), under the QSS diffusion approximation. The differences between the two lead to exponentially large differences in estimates for the mean exit times when $\varepsilon$ is small. The mean exit time from the left and right well is shown in Fig. 6.8. Solid curves show the analytical approximation $T \sim 1/\lambda_0$, where $\lambda_0$ is given by (1.191), as a function of $1/\varepsilon$. For comparison, the mean exit time computed from averaged Monte Carlo simulations of the full stochastic system is shown as symbols. As expected, the log of the mean exit time is an asymptotically linear function of $1/\varepsilon$, and this is confirmed by Monte Carlo simulations. The slope is determined by the depth of the potential well, and the vertical shift is determined by the prefactor. Also shown is the corresponding MFPT calculated using the QSS diffusion approximation (dashed curves), which is typically several orders of magnitude out and validates the relative accuracy of the quasistationary approximation.

One of the applications of noise-induced transitions between metastable states in a stochastic population model is to the study of switching between up and down states during slow-wave sleep; see Sect. 5.1. For example, Holcman and Tsodyks

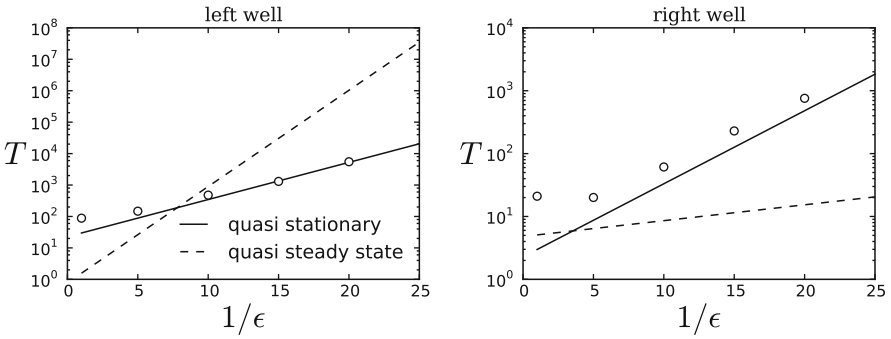


**Fig. 6.8** Mean exit time from the left and right well calculated using the quasistationary approximation (*solid line*) and the QSS diffusion approximation (*dashed line*) [82]. The *open circles* represent data points obtained by numerically solving the corresponding jump velocity Markov process using the Gillespie algorithm. Parameter values are the same as in Fig. 6.7

[282] consider an extension of the deterministic one-population model (6.87) by including the effects of synaptic depression and extrinsic noise. The voltage-based model equations take the form

$$\tau \frac{du}{dt} = -u + qw\chi F(u) + \sqrt{\tau}\sigma\xi(t) + I(t) \tag{6.102a}$$

$$\frac{dq}{dt} = \frac{1-q}{\tau_q} - \chi q F(V), \tag{6.102b}$$

where $q(t)$ is the depression variable, $\xi(t)$ is a white noise term, $I(t)$ is an external input, and the constant $\chi$ is known as a utilization parameter. The firing rate function is taken to be a linear threshold function $F(V) = \eta[u-\kappa]_+$. In the absence of noise ($\sigma = 0$) and external inputs ($I = 0$), the deterministic system undergoes a series of bifurcations as the weight $w$ is increased. For sufficiently small $w$, there exists a globally attracting stable fixed point which is a low activity or down state. At a critical value of $w$, a saddle and unstable node appear, with the latter undergoing a subcritical Hopf bifurcation at a second critical value of $w$—this generates a stable up state inside an unstable limit cycle. The system then exhibits bistability such that noise-induced transitions from the down to the up state have to cross both the separatrix of the saddle and the unstable limit cycle. The resulting fluctuations in activity are suggestive of slow-wave oscillations observed in cortical slices [549].

### *6.4.3 Metastability in a Two-Population Model*

It turns out that the analysis of metastability in a one-population model can be extended to higher-dimensional networks [72]. Consider, for example, the E–I network of Sect. 6.2. The corresponding CK equation (6.75) becomes

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(vp) - \frac{\partial}{\partial y}(\tilde{v}p) + \frac{1}{\varepsilon}\sum_{\mathbf{m}} A(\mathbf{n},\mathbf{m};\mathbf{x})p(\mathbf{x},\mathbf{n}) \tag{6.103}$$

where $\mathbf{x} = (x,y)$, $\mathbf{n} = (n_x,n_y)$, the drift terms are

$$v(x,n_x,n_y) = -x + [w_{EE}n_x - w_{EI}n_y], \tag{6.104}$$

$$\tilde{v}(y,n_x,n_y) = -y + [w_{IE}n_x - w_{II}n_y], \tag{6.105}$$

and $A$ has the nonzero components

$$A(n_x,n_y,n_x-1,n_y;\mathbf{x}) = F(x), \quad A(n_x,n_y,n_x,n_y-1;\mathbf{x}) = F(y), \tag{6.106}$$

$$A(n_x,n_y,n_x+1,n_y;\mathbf{x}) = n_x+1, \quad A(n_x,n_y,n_x,n_y+1;\mathbf{x}) = n_y+1, \tag{6.107}$$

and

$$A(n_x,n_y,n_x,n_y;\mathbf{x}) = -[F(x)+F(y)+n_x+n_y]. \tag{6.108}$$

Here $x$ and $y$ denote the excitatory and inhibitory population variables $a_E, a_I$, respectively. In the limit $\varepsilon \to 0$, we recover the deterministic equations (6.20). However, in contrast to Sect. 6.2, it is now assumed that the E–I network operates in a bistable regime as illustrated in Fig. 6.9.

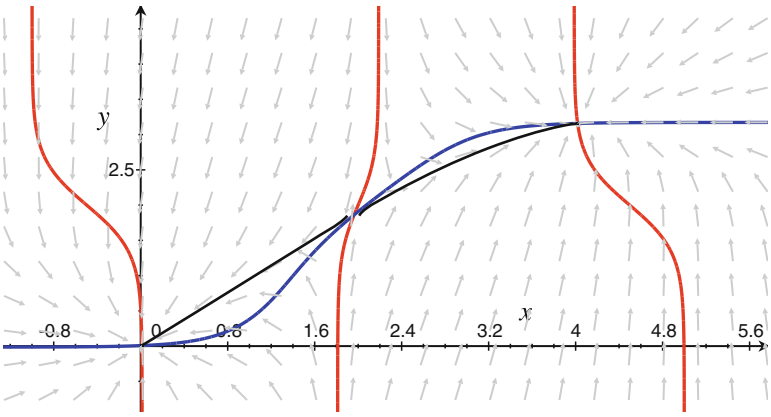


**Fig. 6.9** Bistability in an E–I network. The $x$-nullcline through the saddle is its stable manifold and acts as the separatrix $\Sigma$ between the two stable fixed points. Two deterministic trajectories are shown (directed *black curves*), starting from either side of the unstable saddle and ending at a stable fixed point. Parameter values are $F_0 = 1$, $\gamma = 3$, $\kappa = 2$, $w_{EE} = 5$, $w_{EI} = 1$, $w_{IE} = 9$, and $w_{II} = 6$

Again the MFPT can be identified as the inverse of the principal eigenvalue $\lambda_0$ of the associated linear operator on the right-hand side of (6.103). However, now the analysis is more complicated, since one has to consider stochastic trajectories crossing different points along the separatrix $\Sigma$ between the two metastable states.

Nevertheless, the various steps in the calculation of $\lambda_0$ proceed along similar lines to the one-population model. That is, a spectral projection method can be used to express $\lambda_0$ in terms of the inner product of a quasistationary density with a corresponding adjoint eigenfunction. The former can be calculated using a WKB approximation, except that now one has to determine the stochastic potential $\Phi_0$ by solving an equation of the form [82]

$$\mathscr{H} \equiv -x\mathscr{P}_x - y\mathscr{P}_y - F(x) - F(y) + \Lambda_x(x, \mathscr{P}_x, \mathscr{P}_y) + \Lambda_y(y, \mathscr{P}_x, \mathscr{P}_y) = 0, \quad (6.109)$$

where

$$\mathscr{P}_x = \frac{\partial \Phi_0}{\partial x} \quad \mathscr{P}_y = \frac{\partial \Phi_0}{\partial y}, \quad (6.110)$$

and

$$\Lambda_x = \frac{F(x)}{1 - w_{EE}\mathscr{P}_x - w_{IE}\mathscr{P}_y}, \quad \Lambda_y = \frac{F(y)}{1 + w_{EI}\mathscr{P}_x + w_{II}\mathscr{P}_y} \quad (6.111)$$

Mathematically speaking, Eq. (6.109) is identical to a stationary Hamilton–Jacobi equation for a classical particle with $\mathscr{H}$ identified as the Hamiltonian (see also Sect. 4.4). A trajectory of the particle is given by the solution of Hamilton's equations

$$\begin{aligned} \frac{dx}{dt} &= \frac{\partial \mathscr{H}}{\partial \mathscr{P}_x}, \quad \frac{dy}{dt} = \frac{\partial \mathscr{H}}{\partial \mathscr{P}_y}, \\ \frac{d\mathscr{P}_x}{dt} &= -\frac{\partial \mathscr{H}}{\partial x}, \quad \frac{d\mathscr{P}_y}{dt} = -\frac{\partial \mathscr{H}}{\partial y} \end{aligned} \quad (6.112)$$

Here $t$ is treated as a parameterization of trajectories rather than as a real-time variable. Given a solution curve $(x(t), y(t))$, known as a ray, the potential $\Phi_0$ can be determined along the ray by solving the equation

$$\frac{d\Phi_0}{dt} \equiv \frac{\partial \Phi_0}{\partial x}\frac{dx}{dt} + \frac{\partial \Phi_0}{\partial y}\frac{dy}{dt} = \mathscr{P}_x\frac{dx}{dt} + \mathscr{P}_y\frac{dy}{dt}. \quad (6.113)$$

Thus, $\Phi_0$ can be identified as the action along a zero-energy trajectory. One numerically solves for $\Phi_0$ by considering Cauchy data in a neighborhood of the stable fixed point $(x_-, y_-)$ [82].

The rays $(x(t), y(t))$ (i.e., solutions to Hamilton's equations (6.112) in the $(x, y)$ plane) have an important physical meaning. The trajectory of the ray is the most likely trajectory or path leading away from a point in the neighborhood of a stable fixed point [160]. The rays shown in Fig. 6.10 are obtained by integrating Hamilton's equations (6.112) . These trajectories are only valid in one direction: away from the stable fixed points. For parameter values considered in Fig. 6.10, rays originating from the neighborhood of each stable fixed point cover separate regions, so that the most likely paths between points in each region are connected by deterministic trajectories starting at the boundary between the two regions. Note that this boundary is not the separatrix (gray curve). For example, a trajectory initially at the left fixed point which crosses the separatrix at the saddle would most likely follow

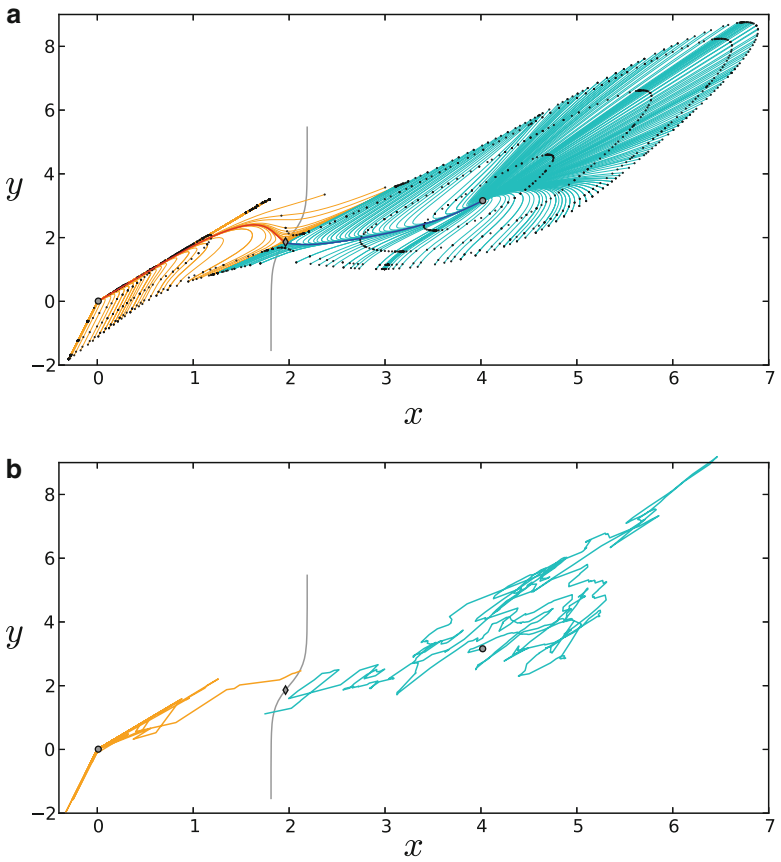


**Fig. 6.10** (**a**) Characteristic paths of maximum likelihood for the 2D model. Rays originating from the left (right) stable fixed point are shown in *orange* (*cyan*), with the ray connecting to the saddle shown in *red* (*blue*). The *gray curve* is the separatrix $\Sigma$. Level curves of constant $\Phi_0(x,y)$ are shown as *black dots*. Each ray has four dots for different values of $\Phi_0(x,y)$. Rays originating from the left fixed point have dots at $\Phi = 0.1, 0.2, \Phi_* + 0.01, \Phi_* + 0.02$, and rays originating from the right fixed point have dots at $\Phi_0 = 0.19, 0.23, 0.28, 0.30$, where $\Phi_* = \Phi(x_*, y_*) = 0.28$. All rays terminate at $\Phi_0 = \Phi_* + 0.02$. (**b**) Sample trajectories of the two-population velocity jump Markov process, whose associated probability density evolves according to (6.83), are computed using the Gillespie algorithm with $\varepsilon = 0.05$ and $N\Delta t = 1$. (The maximum likelihood paths are independent of $\varepsilon$.) Other parameter values are the same as in Fig. 6.9 [72]

a ray towards the saddle and then follow a deterministic trajectory to the right fixed point. If a trajectory crosses the separatrix away from the saddle, it is most likely to cross the separatrix above the saddle when starting from the left fixed point and below the saddle when starting from the right fixed point. In Fig. 6.11, the probability density function for the $y$ coordinate of the point on the separatrix reached by an exit trajectory is shown for each well (square symbols show the histogram for exit from the left well and likewise, 'o' symbols for the right well). Each density function is

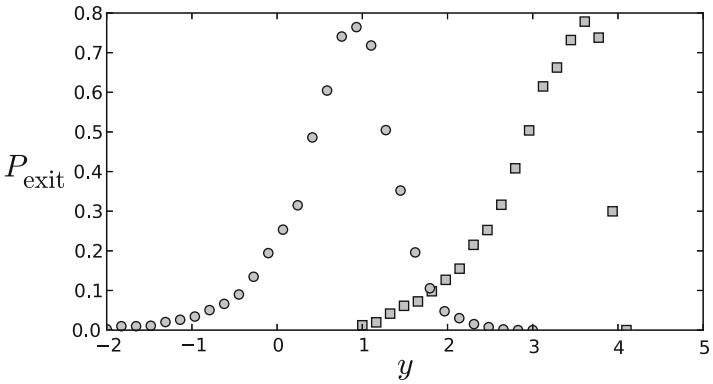


**Fig. 6.11** The probability density for the exit point ($y$ coordinate) where the separatrix is crossed by an exiting trajectory. Results are obtained by $10^2$ Monte Carlo simulation with the same parameters as used in Fig. 6.9, with $\varepsilon = 0.08$. The *square* symbols show trajectories from the left well, and 'o' symbols show trajectories from the right well

peaked away from the saddle point, showing a phenomena known as saddle point avoidance [398, 559]. As $\varepsilon \to 0$, the two peaks merge at the saddle point. Although one might expect the saddle point to be the most likely exit point, since it is the point on the separatrix where the potential $\Phi_0$ takes its minimum value, Fig. 6.11 shows that this is not necessarily true. Even though the most likely exit point is shifted from the saddle, the value of the potential around the saddle point still dominates the mean first exit time.

## 6.5 Spatially Structured Networks and Neural Fields

So far we have not made any assumptions about the topology of the underlying neural network, that is, the structure of the weight matrix $\mathbf{W}$ with components $w_{ij}$. If one looks at a region of cortex such as primary visual cortex (V1), one finds that it has a characteristic spatial structure, in which a high density of neurons ($10^5$ per $mm^3$ in primates) are distributed according to an approximately two-dimensional (2D) architecture. That is, the physical location of a vertical column of neurons within the two-dimensional cortical sheet often reflects the specific information processing role of that population of neurons. For example, in V1 there is an orderly retinotopic mapping of the visual field onto the cortical surface, with left and right halves of the visual field mapped onto right and left visual cortices, respectively. Superimposed upon this are additional two-dimensional maps reflecting the fact that neurons respond preferentially to stimuli with particular features such as local orientation [613]. (A more detailed description of the functional architecture of V1 is given in Sect. 8.1.) This suggests labeling neurons according to their spatial location in cortex. We now give a heuristic argument for how such labeling leads to a continuum neural field model of cortex, following along similar lines to Gerstner and Kistler [214].

For simplicity, consider a population of neurons distributed along a one-dimensional axis. (Extensions to higher dimensions proceed in a similar fashion.) Suppose that we partition space into segments of length $d$ such that the number of neurons in segment $[nd,(n+1)d]$ is $N=\rho d$ where $\rho$ is the cell density. We treat neurons in that interval as a homogeneous population of cells (cortical column) labeled by the integer $n$ and assume that synaptic interactions between the $n$th and $m$th populations only depend on the discrete locations of the populations on the line. Writing $\Phi_{nm}(t)=\rho d\Phi(nd,md,t)$ and $u_n(t)=u(nd,t)$, Eq. (6.4) becomes

$$u(nd,t)=\rho d\sum_m\int_{-\infty}^{t}\Phi(nd,md,t-t')F(u(md,t'))dt'.$$

Taking the limit $d\to 0$, the summation on the right-hand side can be replaced by an integral to give

$$u(x,t)=\int_{-\infty}^{\infty}\int_{-\infty}^{t}\Phi(x,y,t-t')F(u(y,t'))dt'dy, \tag{6.114}$$

where we have absorbed the factor $\rho$ into $\Phi$. Following our derivation of the discrete voltage-based model (6.6), suppose that we can decompose the integral kernel as

$$\Phi(x,y,t)=w(x,y)\Phi(t),\quad \Phi(t)=\mathrm{e}^{-t/\tau}H(t).$$

That is, we assume that there is only one type of neuron so that the temporal kernel $\Phi(t)$ is independent of the presynaptic label $y$ and the postsynaptic label $x$. Applying the differential operator $\mathbb{L}_t=\partial_t+\tau^{-1}$ to the integral equation for $u$ then leads to the scalar neural field equation

$$\frac{\partial}{\partial t}u(x,t)=-\frac{u(x,t)}{\tau}+\int_{-\infty}^{\infty}w(x,y)F(u(y,t))dy. \tag{6.115}$$

Alternatively, we could have applied the differential operator $\mathbb{L}_t$ to the corresponding synaptic drive $z(x,t)=\int_{-\infty}^{t}\Phi(t-t')F(u(x,t'))dt'$ to obtain the activity-based neural field model

$$\frac{\partial}{\partial t}z(x,t)=-\frac{z(x,t)}{\tau}+F\left(\int_{-\infty}^{\infty}w(x,y)z(y,t)dy\right). \tag{6.116}$$

Following the same basic procedure, it is straightforward to incorporate into the neural field equation (6.115) or (6.116) additional features such as synaptic depression [329–331], adaptive thresholds [135, 329], and axonal propagation delays [16, 139, 296, 307, 382, 528, 603]. For example, a voltage-based neural field equation with synaptic depression takes the form

$$\begin{aligned}\frac{\partial}{\partial t}u(x,t)&=-\frac{u(x,t)}{\tau}+\int_{-\infty}^{\infty}w(x,y)q(y,t)F(u(y,t))dy,\\ \frac{\partial}{\partial t}q(x,t)&=\frac{1-q(x,t)}{\tau_q}-\beta q(x,t)F(u(x,t)),\end{aligned} \tag{6.117}$$

with $\beta = 1-\gamma$. In the case of axonal delays, $\tau_{ij} \to \tau(x,y)$ in the continuum limit. Assuming that an action potential propagates with constant speed $v$ along the axon, then $\tau(x,y) = |x-y|/v$ so that the voltage-based equation (6.115) becomes

$$\frac{\partial}{\partial t}u(x,t) = -\frac{u(x,t)}{\tau} + \int_{-\infty}^{\infty} w(x,y)F(u(y,t-|x-y|/v))dy. \tag{6.118}$$

Two-dimensional versions of these various models are obtained by taking $x \to \mathbf{x} = (x_1,x_2)$ and $y \to \mathbf{y} = (y_1,y_2)$ with $d\mathbf{y} = dy_1 dy_2$.

Now suppose that there are $M$ classes of neuron distributed along the line labeled by the population index $a = 1,\ldots M$. Equation (6.114) then generalizes to the multi-population integral equation

$$u_a(x,t) = \int_{-\infty}^{\infty}\int_{-\infty}^{t}\sum_{b=1}^{M} \Phi_{ab}(x,y,t-t')F_b(u_b(y,t'-|x-y|/v_{ab}))dt'dy. \tag{6.119}$$

We have included axonal delays with $v_{ab}$ the conduction velocity along axonal projections from neurons of type $b$ to neurons of type $a$. Assuming that $\Phi_{ab}(x,y,t) = w_{ab}(x,y)\Phi(t)$ with $\Phi(t) = \mathrm{e}^{-t/\tau}H(t)$, we obtain multi-population neural field equations:

$$\frac{\partial u_a}{\partial t} = -\frac{u_a(x,t)}{\tau} + \sum_{b=1}^{M}\int_{-\infty}^{\infty} w_{ab}(x,y)F_b(u_b(y,t-|x-y|/v_{ab}))dy, \tag{6.120}$$

and

$$\frac{\partial z_a}{\partial t} = -\frac{z_a(x,t)}{\tau} + F_b\left(\sum_{b=1}^{M}\int_{-\infty}^{\infty} w_{ab}(x,y)z_b(y,t-|x-y|/v_{ab})dy\right) \tag{6.121}$$

for $a = 1,\ldots,M$. The latter is a version of the *Wilson–Cowan equations* for cortical dynamics [675, 676]. Note that all synapses innervated by a particular type of neuron have the same sign. That is, if type $b$ neurons are excitatory (inhibitory), then $w_{ab}(x,y) \geq 0$ ($w_{ab}(x,y) \leq 0$) for all $a = 1,\ldots,M$ and $(x,y)$. Thus, one of the major reasons for considering more than one class of neuron is to incorporate both excitatory and inhibitory synapses. It can be argued that since excitatory and inhibitory synapses tend to have different time courses in response to action potentials, one should take $\Phi_{ab}(x,y,t) = w_{ab}(x,y)\Phi_b(t)$, suggesting that the activity-based model (6.121) with $\tau \to \tau_a$ is more biologically realistic than the voltage-based model, at least in the case of excitatory–inhibitory networks [173]. However, in practice, both versions of the neural field equations are used to model cortical dynamics. Since both versions exhibit very similar types of solution, and since most analytical results have been obtained for voltage-based neural fields, we will mainly focus on the latter.

Under certain additional simplifying assumptions, it is possible to incorporate inhibition into the scalar neural field equations (6.115) or (6.116) [505]. For example, consider a two-population model ($M = 2$) of excitatory ($a = E$) and

inhibitory ($a = I$) neurons evolving according to the pair of continuum voltage-based equations

$$\frac{\partial u_E}{\partial t} = -\frac{u_E(x,t)}{\tau_E} + \int_{-\infty}^{\infty} w_{EE}(x,y)F_E(u_E(y,t))dy + \int_{-\infty}^{\infty} w_{EI}(x,y)F_I(u_I(y,t))dy \tag{6.122a}$$

$$\frac{\partial u_I}{\partial t} = -\frac{u_I(x,t)}{\tau_I} + \int_{-\infty}^{\infty} w_{IE}(x,y)F_E(u_E(y,t))dy + \int_{-\infty}^{\infty} w_{II}(x,y)F_I(u_I(y,t))dy, \tag{6.122b}$$

with $w_{EE}, w_{IE} \geq 0$ and $w_{EI}, w_{II} \leq 0$. Now suppose that $w_{II} \equiv 0$, $F_I(u_I) = u_I/\tau_I$ and $\tau_I \ll \tau_E$. It follows that we can eliminate $u_I$ by setting

$$u_I(x) \sim \tau_I \int_{-\infty}^{\infty} w_{IE}(x,y)F_E(u_E(y,t))dy,$$

which leads to a scalar equation for $u_E$ of the form (6.115) with effective weight distribution

$$w(x,y) = w_{EE}(x,y) + \int_{-\infty}^{\infty} w_{EI}(x,y')w_{IE}(y',y)dy'. \tag{6.123}$$

It is then possible for $w(x,y)$ to change sign as a function of $x,y$. (Often $w$ is modeled as a difference of Gaussians or exponentials—the so-called Mexican hat weight distribution.) The reduced model can be used to investigate the effects of inhibition on stationary solutions and propagating waves. However, in contrast to the full two-population model, it does not support oscillatory solutions (in the absence of axonal delays, higher-order synapses, or some form of adaptation such as synaptic depression).

It is important to emphasize that there does not currently exist a multi-scale analysis of conductance-based neural networks that provides a rigorous derivation of neural field equations, although some progress has been made in this direction [100, 147, 184, 307, 675, 676]. One crucial step in the derivation of neural field equations presented here was the assumption of slowly varying synaptic currents, which is related to the assumption that there is not significant coherent activity at the level of individual spikes. This allowed us to treat the output of a neuron (or population of neurons) as an instantaneous firing rate. A more rigorous derivation would need to incorporate the mean-field analysis of local populations of stochastic spiking neurons into a larger-scale cortical model and to carry out a systematic form of coarse graining or homogenization in order to generate a continuum neural field model. Nevertheless, the heuristic approach does provide a framework for relating parameters of neural field equations to biophysical parameters such as membrane/synaptic time constants and axonal delays and also prescribes how to incorporate additional physiological processes such as synaptic depression and spike frequency adaptation. Moreover, neural field models make it possible to explore the dependence of cortical dynamics on the detailed anatomy of local and long-range synaptic connections. It is often assumed that $w$ depends on the Euclidean distance between interacting cells within the 2D cortical sheet so that

$w(\mathbf{x},\mathbf{y}) = w(|\mathbf{x}-\mathbf{y}|)$. However, this is an oversimplification of the detailed architecture of cortex [64–66, 308, 526]; see Sect. 8.1. A related simplification is to take axonal delays to depend on Euclidean distance according to $|\mathbf{x}-\mathbf{y}|/v$, where $v$ is the speed of propagation.

It is also possible to construct continuum neural field equations for stochastic population models. For the sake of illustration, consider the Langevin equation (6.71), except we simplify the multiplicative noise by additive noise $\sigma dW_\alpha$ for constant noise strength $\sigma$. The continuum limit of Eq. (6.71) proceeds as follows. First, set $A_\alpha(t) = A(\alpha\Delta d,t)$ and $w_{\alpha\beta} = \rho\Delta d w(\alpha\Delta d,\beta\Delta d)$ where $\rho$ is a synaptic density and $\Delta d$ is an infinitesimal length scale. We also assume that the noise strength $\sigma_\alpha = \sigma/\sqrt{\Delta d}$ and define $W_\alpha(t)/\sqrt{\Delta d} = W(\alpha\Delta d,t)$. Taking the limit $\Delta d \to 0$ with $\tau_\alpha = \hat{\tau}$ for all $\alpha$ gives

$$\hat{\tau} dA(x,t) = \left[-A(x,t) + F\Big(\int w(x-y)A(y,t)dy\Big)\right]dt + \sigma dW(x,t) \tag{6.124}$$

with

$$\langle dW(x,t)\rangle = 0, \quad \langle dW(x,t)dW(y,t)\rangle = \delta(x-y)dt. \tag{6.125}$$

Similarly, we can write down a stochastic version of a voltage-based neural field equation, namely,

$$\tau dU(x,t) = [-U(x,t) + \int w(x-y)F(U(y,t))dy]dt + \sigma dW(x,t) \tag{6.126}$$

From a numerical perspective, any computer simulation would involve rediscretizing space and then solving a time-discretized version of the resulting stochastic differential equation. On the other hand, in order to investigate analytically the effects of noise on spatiotemporal dynamics, 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 [546], as illustrated in Sect. 7.4. Recently, Buice and Cowan [91] have used path integral methods and renormalization group theory to establish that a stochastic neural field model based on a continuum version of a birth–death master equation belongs to the universality class of directed percolation and consequently exhibits power law behavior, suggestive of many measurements of spontaneous cortical activity in vitro and in vivo [30, 508]; see Sect. 9.5. Although the existence of power law behavior is still controversial [28], the application of path integral methods provides yet another example of how analytical techniques familiar in the study of PDEs are being adapted to studies of continuum neural fields. (For reviews on path integral methods for stochastic differential equations see [115, 618, 698].)

Finally, note that several groups have constructed equivalent PDE models for neural fields with axonal propagation delays, which take the form of damped inhomogeneous wave equations [307, 382, 467, 527, 528, 603]. The basic idea is to assume a particular form for the synaptic weight distribution and to use Fourier transforms. Consider, for example, a 2D version of the multi-population integral

equation (6.119). Suppose that $\Phi_{ab}(\mathbf{x},\mathbf{y},t) = w_{ab}(|\mathbf{x}-\mathbf{y}|)\Phi(t)$ and introduce the auxiliary field

$$\Psi_{ab}(\mathbf{x},t) = \int_{\mathbb{R}^2} w_{ab}(|\mathbf{x}-\mathbf{y}|)F_b(\mathbf{y},t-|\mathbf{x}-\mathbf{y}|/v_{ab})d\mathbf{y},$$

where we have set $F_b(\mathbf{y},t) = F_b(u_b(\mathbf{y},t))$. Fourier transforming this equation with

$$\hat{\Psi}_{ab}(k,\omega) = \int_{\mathbb{R}^2}\int_{-\infty}^{\infty} \mathrm{e}^{-i(\mathbf{k}\cdot\mathbf{r}+\omega t)}\Psi_{ab}(\mathbf{x},t)dt\,d\mathbf{x},$$

and

$$w_{ab}(r) = \frac{w_{ab}^0}{2\pi}\mathrm{e}^{-r/\sigma_{ab}},$$

we find that

$$\hat{\Psi}_{ab}(k,\omega) = w_{ab}^0\frac{\mathscr{A}_{ab}(\omega)}{(\mathscr{A}_{ab}(\omega)^2+k^2)^{3/2}}\hat{F}_b(k,\omega)$$

with $\mathscr{A}_{ab}(\omega) = 1/\sigma_{ab} + i\omega/v_{ab}$. If one now makes a long-wavelength approximation by Taylor expanding the denominator of the above equation about $k=0$ and rearranging to give $(\mathscr{A}_{ab}(\omega)^2 + 3k^2/2)\hat{\Psi}_{ab}(k,\omega) = \hat{F}_b(k,\omega)$, one can then apply the inverse Fourier transform to derive the damped wave equation

$$\left[\left(\frac{1}{\sigma_{ab}}+\frac{1}{v_{ab}}\partial_t\right)^2 - \frac{3}{2}\nabla^2\right]\Psi_{ab}(\mathbf{x},t) = w_{ab}^0F_b(u_b(\mathbf{x},t)).$$

The current $u_a(\mathbf{x},t)$ is then related to the field $\Psi_{ab}(\mathbf{x},t)$ according to

$$u_a(\mathbf{x},t) = \int_{-\infty}^{t}\Phi(t-t')\sum_b\Psi_{ab}(\mathbf{x},t')dt',$$

which can itself be converted to a PDE by applying the inverse operator $\mathbb{L}_t$. There have also been various extensions of the PDE theory including improvements upon the long-wavelength approximation [142] and incorporation of network inhomogeneities [526]. The damped wave equation (6.127) and its generalizations have been used extensively to study large-scale phenomena such as EEG rhythms, where axonal propagation delays are important [60, 467]; see also Sect. 9.4. PDE models have also been used to study single and multi-bump stationary solutions of scalar neural field equations, in which the time-independent equations reduce to fourth-order differential equations with a Hamiltonian structure [355, 356].