Chapter 15

Mean-Field Theory of Irregularly Spiking Neuronal Populations and Working Memory in Recurrent Cortical Networks

Alfonso Renart,[1] **Nicolas Brunel,**[2] **and Xiao-Jing Wang**[1]

[1]*Volen Center for Complex Systems, Brandeis University, Waltham, MA 02254, U.S.,* [2]*CNRS, Neurophysique et Physiologie du Système Moteur, Université Paris René Descartes, 45 rue des Saints Pères, 75270 Paris Cedex 06, France*

CONTENTS

1-58488-362-6/04/$0.00+$1.50

15.1 Introduction

In cortical neural circuits, the biophysics of neurons and synapses and the collective network dynamics produce spatiotemporal spike patterns that presumably are optimized for the functional specialization of the system, be it sensory, motor or memory. Therefore, different systems might use different codes. For example, the 'spike timing code' or 'correlation code' that relies on precise spike timing is critical for the computation of coincidence detection in the brainstem auditory pathways, and may also contribute to information processing in other neural systems. A 'burst code' is prevalent in central pattern generators of the motor systems, where rhythmicity is produced by oscillatory repetition of brief clusters of spikes (bursts). Neurons can also signal information using a 'rate code', by virtue of the frequency at which the spikes are discharged. The idea of rate coding originated from the work of [3], who discovered that a stimulus feature (such as intensity) could be accurately read out from the firing rate of a sensory neuron. Since then, many studies have shown that firing rates convey a large amount of stimulus-related information in neurons.

In a small neural network, such as the visual system of flies or the electrosensory system of electric fish, there are a few synaptic connections per cell and each spike has a large impact on the post-synaptic cell. Hence spike timing is expected to be important. Moreover, a reliable estimate of the firing rate of one or a few pre-synaptic inputs requires a long-time average of spike counts and is, hence, not adequate to subserve fast perceptual or motor behaviors in these systems at fast time scales ($\sim$ 100 milliseconds). The situation, however, is drastically different in a cortical circuit, where a huge number of neurons are available and organized into columns of functionally similar neurons [84]. A typical cortical neuron receives thousands of synapses, most of them from neighboring neurons [4, 76]; the impact of a single pre-synaptic spike onto a post-synaptic cell is relatively small. Moreover, spike trains of cortical neurons are highly stochastic and irregular (see e.g., [30, 108, 110], but see [59]), hence there is a lot of noise in spike timing. This fact raised the question of whether the observed spike train irregularity conveyed information or was rather a reflection of the various sources of noise present at the cellular and network levels [105]. Even if the spike times from single cells are noisy, information can still be conveyed in the average activity of pools of weakly correlated neurons. Suppose that a neuron receives connections from C_{cell} other neurons in a column. Being from the same column, the average activity of these inputs is similar, but since their spike trains are irregular the number $N_i(\Delta t)$ of spikes emitted by each cell i in the time interval $[t, t+\Delta t]$ is random. The total input to the post-synaptic neuron

$$f(t) \sim \sum_{i}^{C_{cell}} N_i(\Delta t)$$

provides an estimate of the average activity across the population. Since C_{cell} is large (100-1000) [17], and neurons are only weakly correlated [14, 31, 74, 130], noise can

be largely (though not completely) averaged out [105, 106], and the estimate of the average activity of the pre-synaptic pool can be quite accurate even with a small Δt. In other words, the population firing rate can be defined (almost) instantaneously in real time. Moreover, such a rate code can be readily decoded by a post-synaptic cell: the summation of thousands of synaptic inputs provides a means to readout the population firing rate at any time.

Population firing rate models were introduced in the early 1970s and have since become widely popular in theoretical neuroscience. These models are described as non-linear differential equations, to which tools of mathematical analysis are applicable. Thus, concepts like attractor dynamics, pattern formation, synchronous network oscillations, etc, have been introduced in the field of neurobiology (See [37] for a review and references). Early models, such as associative memory models, were formulated in terms of firing-rates [27, 64]. Broadly speaking, two different approaches can be used to construct a firing-rate model. A rate model can be built *heuristically*: for example, a unit is assumed to have a threshold-linear or sigmoid input-output relation [5, 125, 126]. This class of rate models is valuable for its simplicity; important insights can be gained by detailed analysis of such models. The drawback is that these models tend to be not detailed enough to be directly related to electrophysiology. For example, the baseline and range of firing rates are arbitrarily defined so they cannot be compared with those of real neurons. It is therefore difficult to use the available data to constrain the form of such models. On the other hand, a firing-rate model can also be *derived*, either rigorously or approximately, from a spiking neuron model. To do that, the dynamics of spiking neurons must be well understood. The analytical study of the dynamics of spiking neuron models was pioneered by [68], and has witnessed an exponential growth in recent years. Up to date, most of the work was done with the leaky-integrate-and-fire (LIF) neuron model [1, 8, 11, 18, 19, 21, 24, 51, 68, 77, 88, 114, 120]. The LIF model is a simple spiking model that incorporates basic electrophysiological properties of a neuron: a stable resting potential, sub-threshold integration, and spikes. A network model can be constructed with LIF neurons coupled by realistic synaptic interactions. Such models have been developed and studied for many problems, such as synchronization dynamics, sensory information processing, or working memory. In some instances, firing-rate dynamics can be derived from the underlying spiking neuron models [26, 36, 40, 109]. These firing rate models provide a more compact description that can be studied in a systematical way.

Analytical studies of networks of neurons are usually performed in the context of 'mean-field' theories. In such theories, the synaptic input of a neuron in the network is traditionally only described by its average: the 'mean-field'. This first class of models is applicable to networks in which neurons are weakly coupled and fire in a regular fashion. More recently, mean-field theories have been introduced in which the synaptic inputs are described not only by their mean, but also by the fluctuations of their synaptic inputs, which come potentially both from outside the network, and from the recurrent inputs. This second class of models is applicable to strongly coupled networks in which neurons fire irregularly [11, 118, 119].

The objective of this chapter is to provide a pedagogical summary of this latter

type of mean-field theory in its current state. We will introduce the theory in several steps, from single neurons, to self-consistent theory of a recurrent network with simple synapses, to realistic synaptic models. The chapter is organized into two parts, which address the two general ingredients of a mean-field theory for network models based on biophysics. First, a method is needed for the analytical description of a neuron's output in response to a large number of highly noisy pre-synaptic inputs, including realistic synaptic interactions (time course, voltage dependence) which are critical in determining the network behavior. This will be described in Section 15.2. Second, in a recurrent circuit, any neuron both receives inputs from, and sends output to, other neurons in the same network. Therefore, the pre-synaptic and post-synaptic firing rates are related to each other. The mean-field theory provides a procedure to calculate the neural firing rate in a self-consistent manner, in the steady-state. It also can also be extended to a description of the temporal dynamics of the neural firing rate. This will be discussed in Section 15.3. The self-consistent theory is then applied to a strongly recurrent network model of working memory which displays multi-stability between a resting state and memory-related persistent activity states.

15.2 Firing-rate and variability of a spiking neuron with noisy input

The first part of the present paper is devoted to the firing properties of a leaky integrate-and-fire (LIF) neuron in response to stochastic synaptic inputs. After the introduction of the LIF neuron, we proceed as follows: First, the statistical properties of the input current will be described, given certain assumptions about the stochastic activity of the pre-synaptic inputs to the neuron. Second, we will discuss the conditions under which the dynamics of the depolarization can be approximated by a diffusion equation. Third, we will show how to calculate the output mean firing rate and coefficient of variation (CV) of the cell given our assumptions. Next the effect of finite synaptic time constants will be explained. Finally, we provide a discussion on how realistic synaptic transmission, including voltage-dependent conductances and non-linear summation of inputs, can be incorporated into this framework.

15.2.1 The leaky integrate-and-fire neuron

In the LIF model, the voltage difference $V(t)$ across the membrane changes in response to an injected current $I(t)$ according to

$$C_m \frac{dV(t)}{dt} = -g_L(V(t) - V_L) + I(t), \tag{15.1}$$

where $C_m = 0.2$ nF is the total membrane capacitance, $g_L = 20$ nS is the leak conductance and $V_L = -70$ mV is the leak, or resting potential of the cell in the absence of

input (see e.g., [69]). According to this equation, the membrane is seen as a simple RC circuit, with a time constant τ_m given by

$$\tau_m = \frac{C_m}{g_L} = 10 \text{ ms}. \tag{15.2}$$

Spiking is implemented in the model by defining a threshold voltage V_{th} such that the neuron is said to emit a spike at time t_{spk} whenever $V(t = t_{spk}) = V_{th} = -50$ mV. Refractoriness is taken into account by clamping the voltage to a reset value $V_r = -60$ mV for a time $\tau_{ref} = 2$ ms after each spike, i.e., if $V(t = t_{spk}) = V_{th}$, then $V(t') = V_r$ for $t' \in (t^+_{spk}, t_{spk} + \tau_{ref})$. When the neuron is inserted in a network, $I(t)$ represents the total synaptic current, which is assumed to be a linear sum of the contributions from each individual pre-synaptic cell.

15.2.2 Temporal structure of the afferent synaptic current

We will start with the simplest description of the interaction between the pre- and post-synaptic neurons. It amounts to assuming that each pre-synaptic spike causes an instantaneous change in post-synaptic voltage which is independent of the current value of this voltage, and depends only on a parameter J measuring the strength of the synapse (more precisely, J is the amount of positive charge entering the membrane due to the spike). If C neurons synapse onto this cell, each with an efficacy $J_i \quad (i = 1, \ldots, C)$, then the current into the cell can be represented as

$$I(t) = \sum_{i=1}^{C} J_i \sum_j \delta(t - t^i_j), \tag{15.3}$$

where t^i_j is the time of the j^{th} spike from the i^{th} pre-synaptic neuron. If the neuron is initially at rest, and a pre-synaptic cell fires a single spike at time $t = 0$, then by integrating Equation (15.1) one obtains

$$V(t) = V_L + \frac{J_i}{C_m} \exp\left(-\frac{t}{\tau_m}\right) \Theta(t), \tag{15.4}$$

where $\Theta(t)$ is the Heaviside function, $\Theta(t) = 0$ if $t < 0$ and 1 if $t > 0$. Thus, the post-synaptic potential (PSP) produced by each pre-synaptic spike consists of an instantaneous "kick" of size $\bar{J}_i = J_i/C_m$ followed by an exponential decay with time constant τ_m. For example, if the unitary charge is $J_i = 0.04$ pC, and $C_m = 0.2$ nF, then the kick size is $\bar{J}_i = 0.04/0.2 = 0.2$ mV.

We consider a neuron receiving synaptic input from a large pool of C_E excitatory and C_I inhibitory cells. We make two important assumptions regarding the activity of these inputs: first, that each of them fires spikes according to a stationary Poisson process, i.e., with a constant probability of emitting a spike per unit time. Second, that these Poisson processes are independent from cell to cell, i.e., the occurrence of a spike from any given cell does not give any information about the firing probability of any other neuron. These assumptions will need to be verified at the network level for the theory to be self-consistent (see Section 15.3).

We denote the average firing rate of each excitatory (inhibitory) input $j = 1, \ldots, C_{E,I}$, as ν_{E_j} (ν_{I_j}), and the efficacy of the corresponding excitatory (inhibitory) synapse as J_{E_j} (J_{I_j}). For simplicity, we first assume that all the rates and synapse from each pre-synaptic population are identical, i.e., $\nu_{E_j} = \nu_E$ and $J_{E_j} = J_E$ for all j, and similarly for the inhibitory population.

In this simple situation, the temporal average of the total current is constant in time and given by

$$< I(t) > \equiv \mu_C = \sum_{j=1}^{C_E} J_{E_j}\nu_{E_j} - \sum_{i=1}^{C_I} J_{I_i}\nu_{I_i} = C_E J_E \nu_E - C_I J_I \nu_I. \tag{15.5}$$

For a Poisson process $s(t)$ of rate ν, $< (s(t)-\nu)(s(t')-\nu) >= \nu\delta(t-t')$. Thus, using the fact that the inputs are Poisson and independent, the connected two point correlation function of the total current is given by

$$\begin{aligned} < (I(t)- < I >)(I(t')- < I >) > &= \left[\sum_j^{C_E} J_{E_j}^2 \nu_{E_j} + \sum_i^{C_I} J_{I_i}^2 \nu_{I_i}\right] \delta(t-t') \\ &= (C_E J_E^2 \nu_E + C_I J_I^2 \nu_I)\delta(t-t') \\ &\equiv \sigma_C^2 \delta(t-t'). \end{aligned} \tag{15.6}$$

15.2.3 The diffusion approximation

In principle, the next step would be to solve the dynamics of the depolarization as described in Equation (15.1) in the presence of the stochastic current $I(t)$. As it is, this task is still too difficult, so we will make one further approximation, namely to replace the *point* process $I(t)$, by a process $\bar{I}(t)$ with the same mean and two-point correlation function as $I(t)$, such that the voltage response $V(t)$ to $\bar{I}(t)$ becomes continuous (instead of discrete as a result of the synaptic kicks) in time. The idea is to make the size of the voltage kicks $\bar{J}_{E,I} \equiv J_{E,I}/C_m$ small, while at the same time increasing their overall frequency by increasing $C_{E,I}$ (notice that since the sum of two Poisson processes is another Poisson process, $I(t)$ can be considered the difference of two Poisson processes of rates $C_E\nu_E$ and $C_I\nu_I$ respectively). For a cortical neuron, since it receives a large number of pre-synaptic contacts, each of which contributes only to a small fraction of the voltage distance between rest and threshold, one expects this approximation to be plausible and give accurate results.

Since the inputs to our cell are assumed to be stochastic, the temporal evolution of $V(t)$ is probabilistic. The fundamental object for the description of the dynamics of the membrane potential is the probability density $\rho(V,t|V_0,t_0)$ for $V(t) \in [V, V+dV]$ given that $V(t_0) = V_0$. If we consider our averages to be carried out over an *ensemble* of identical neurons, each with a different realization of the stochasticity, $\rho(V,t|V_0,t_0)$ can be considered a "population" density, so that $\rho(V,t|V_0,t_0)dV$ is the fraction of neurons among the ensemble with membrane potentials in $[V, V+dV]$ given that all neurons were at V_0 at $t = t_0$. In the Appendix, we present an intuitive derivation of a differential equation which governs the temporal evolution of

$\rho(V,t|V_0,t_0)$ in the presence of the stochastic input $I(t)$ (see e.g., [49, 54, 97, 99] for more details). Such equation reads

$$\frac{\partial}{\partial t}\rho(V,t|V_0,t_0) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial V^n}[A_n\, \rho(V,t|V_0,t_0)], \tag{15.7}$$

where

$$\begin{aligned} A_1(V) &= -\frac{(V-V_L)}{\tau_m} + (\bar{J_E} C_E \nu_E - \bar{J_I} C_I \nu_I) = \\ &= -\frac{(V-V_L)}{\tau_m} + \frac{\mu_C}{C_m} \equiv -\frac{(V-V_{ss})}{\tau_m} \\ A_2 &= (\bar{J_E}^2 C_E \nu_E + \bar{J_I}^2 C_I \nu_I) = (\frac{\sigma_C}{C_m})^2 \equiv \frac{\sigma_V^2}{\tau_m} \\ A_n &= (\bar{J_E}^n C_E \nu_E + (-1)^n \bar{J_I}^n C_I \nu_I) \qquad n = 3,4,\ldots \end{aligned} \tag{15.8}$$

where A_n are the infinitesimal moments of the stochastic process. The infinitesimal moments completely specify the dynamics of $\rho(V,t|V_0,t_0)$. The drift coefficient A_1 captures the deterministic component of the temporal evolution of $V(t)$; $V_{ss} = V_L + \mu_C/g_L$ is the steady-state voltage in the absence of stochasticity. The diffusion coefficient A_2 measures the fluctuations of $V(t)$. In the absence of threshold, the variance of the depolarization is $\sigma_V^2/2 = \sigma_C^2 \tau_m/(2C_m^2)$.

In what is referred to as the diffusion approximation, A_n for $n > 2$ are assumed to be negligible and set to zero [97, 117]. Looking at Equations (15.8), one can see under which conditions this will be a valid approximation. Since the infinitesimal moments depend on powers of the kick size times their overall rate, one expects the approximation to be appropriate if the kick size is very small but the overall rate is very large, in such a way that the size of all moments of order higher than two become negligible in comparison with the drift and diffusion coefficients. In particular, in the next sections we show how, in the limit of infinitely large networks, if the synaptic efficacies are scaled appropriately with the network size, the approximation can become exact.

We will for now take it for granted, and focus on the properties of Equation (15.7) when only the first two infinitesimal moments are non-zero. The resulting equation is called the Fokker-Planck equation for $\rho(V,t|V_0,t_0)$, and reads

$$\frac{\partial}{\partial t}\rho(V,t|V_0,t_0) = \frac{\partial}{\partial V}[\frac{(V-V_{ss})}{\tau_m}\rho(V,t|V_0,t_0)] + \frac{\sigma_V^2}{2\tau_m}\frac{\partial^2}{\partial V^2}[\rho(V,t|V_0,t_0)]. \tag{15.9}$$

The process described by this equation, characterized by a constant diffusion coefficient $D = \sigma_V^2/(2\tau_m)$ and a linear drift, is called the Ornstein-Uhlenbeck (O-U) process (see e.g., [123]). It describes the temporal evolution of $V(t)$ when the input to the neuron is no longer $I(t)$, but

$$\bar{I}(t) \equiv \mu_C + \sigma_C \eta(t), \tag{15.10}$$

where $\eta(t)$ is called a *white noise* process. It can be defined heuristically as a random variable taking values

$$\eta(t) = \lim_{dt \to 0} \mathbf{N}(0, \frac{1}{\sqrt{dt}}) \tag{15.11}$$

for all t independently, where we have defined $\mathbf{N}(\alpha, \beta)$ is a Gaussian random variable of mean α and variance β^2. The mean and two-point correlation function of the white noise process are therefore, $< \eta(t) >= 0$ and $< \eta(t)\eta(t') >= \delta(t-t')$ respectively. In effect, we are now replacing Equation (15.1) by

$$C_m \frac{dV(t)}{dt} = -g_L(V(t) - V_L) + \mu_C + \sigma_C \eta(t), \tag{15.12}$$

or

$$\tau_m \frac{dV(t)}{dt} = -(V(t) - V_{ss}) + \sigma_V \sqrt{\tau_m} \eta(t). \tag{15.13}$$

This is called the white-noise form of the Langevin equation of the process $V(t)$. It has the appeal that it is written as a conventional differential equation so that the dynamics of $V(t)$ is described in terms of its sample paths, rather than in terms of the temporal evolution of its probability distribution, as in the Fokker-Planck Equation (15.9). In general, the practical use of the Langevin equation is that it provides a recipe for the numerical simulation of the sample paths of the associated process. Developing Equation (15.13) to first order one obtains

$$V(t+dt) = (1 - \frac{dt}{\tau_m})V(t) + V_{ss}\frac{dt}{\tau_m} + \sigma_V \sqrt{\frac{dt}{\tau_m}} \mathbf{N}(0,1). \tag{15.14}$$

Assuming that dt/τ_m is small but finite, Equation (15.14) provides an iterative procedure which gives an *approximate* description of the temporal evolution of $V(t)$. This scheme is general and can be used for any diffusion process. For the O-U process in particular, in the absence of threshold Equation (15.9) can be solved exactly. The population density of this process is a Gaussian random variable with a time-dependent mean and variance [97, 123], so that

$$\begin{aligned} \rho(V,t|V_0,t_0) = \mathbf{N}\Big(V_{ss} + (V_0 - V_{ss})\exp(-\tfrac{t-t_0}{\tau_m}), \tfrac{\sigma_V}{\sqrt{2}} \\ \Big[1 - \exp(-\tfrac{2(t-t_0)}{\tau_m})\Big]^{1/2}\Big). \end{aligned} \tag{15.15}$$

Using this result one can find an exact iterative procedure for the numerical simulation of the process. Assuming V_0 is the value of the depolarization in the sample path at time t, e.g., $V_0 = V(t)$, the depolarization at a latter time $t + \Delta t$ will be

$$\begin{aligned} V(t+\Delta t) = V_{ss} + (V(t) - V_{ss})\exp(-\tfrac{\Delta t}{\tau_m}) \\ + \tfrac{\sigma_V}{\sqrt{2}}\Big[1 - \exp(-\tfrac{2\Delta t}{\tau_m})\Big]^{1/2} \mathbf{N}(0,1). \end{aligned} \tag{15.16}$$

This update rule is exact for all Δt [54].

In Figure 15.1, sample paths of $V(t)$ in the presence of the original input current $I(t)$ obtained by numerical integration of Equation (15.1) are compared with sample paths in the presence of the effective input $\bar{I}(t)$, obtained using Equation (15.16). As illustrated in Figure 15.1, $\tau_{ref} = 2$ ms after emitting a spike, $V(t)$ begins to integrate its inputs again starting from V_r until it reaches V_{th}. The first time $V(t)$ reaches V_{th} is called the 'first-passage time' (denoted by T_{fp}). Taking the refractory period into account, the whole interval between consecutive spikes is called the inter-spike interval (ISI). Therefore, the statistics of ISIs can be analyzed using the theory of first-passage times of the Ornstein-Uhlenbeck process [97, 117].

15.2.4 Computation of the mean firing rate and CV

The Fokker-Planck Equation (15.9) can be rewritten as a continuity equation by defining

$$S(V,t|V_0,t_0) \equiv -\frac{(V-V_{ss})}{\tau_m}\rho(V,t|V_0,t_0) - \frac{\sigma_V^2}{2\tau_m}\frac{\partial}{\partial V}\rho(V,t|V_0,t_0)], \tag{15.17}$$

so that Equation (15.9) becomes

$$\frac{\partial}{\partial t}\rho(V,t|V_0,t_0) = -\frac{\partial}{\partial V}S(V,t|V_0,t_0). \tag{15.18}$$

Thus, $S(V,t|V_0,t_0)$ is the flux of probability (or probability current) crossing V at time t. To proceed, a set of boundary conditions on t and V has to be specified for $\rho(V,t|V_0,t_0)$. First one notices that, if a threshold exists, then the voltage can only be below threshold and can only cross it from below (the threshold is said to be an absorbing barrier). The probability current at threshold gives, by definition, the average firing rate of the cell. Since $\rho(V > V_{th},t|V_0,t_0) = 0$, the probability density must be zero at $V = V_{th}$, otherwise the derivative would be infinite at $V = V_{th}$ and so would be the firing rate according to Equation (15.17). Therefore, we have the following boundary conditions

$$\rho(V_{th},t|V_0,t_0) = 0 \qquad \text{and} \qquad \frac{\partial}{\partial V}\rho(V_{th},t|V_0,t_0) = -\frac{2\nu(t)\tau_m}{\sigma_V^2}, \tag{15.19}$$

for all t. The conditions at $V = -\infty$ ensure that the probability density vanishes fast enough to be integrable, i.e.,

$$\lim_{V\to-\infty}\rho(V,t|V_0,t_0) = 0 \qquad \text{and} \qquad \lim_{V\to-\infty}V\rho(V,t|V_0,t_0) = 0. \tag{15.20}$$

Since the threshold is an absorbing boundary, a finite probability mass is constantly leaving the interval $(-\infty,V_{th})$. Under this condition, there is no stationary distribution for the voltage, i.e., $\rho(V,t|V_0,t_0) \to 0$ as $t \to \infty$. In order to study the steady-state of the process, one can keep track of the probability mass leaving the integration interval at t, and re-inject it at the reset potential at $t + \tau_{ref}$. This injection

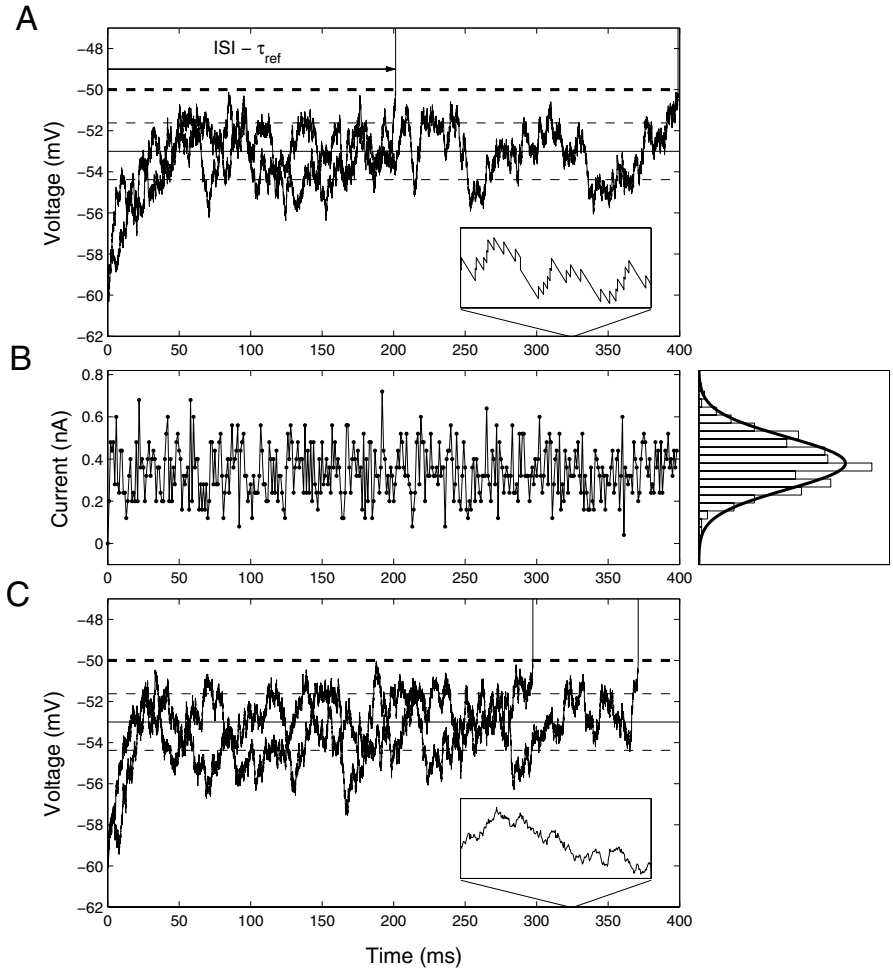

Figure 15.1

Leaky integrate-and-fire neuron model in response to stochastic inputs. **(A)**. Sample paths of the membrane potential $V(t)$ in response to a stochastic current $I(t)$ obeying Equation (15.3) with Poisson spike trains, with $\bar{J}_E = \bar{J}_I = 0.2$ mV, $C_E = C_I = 1000$, $\nu_E = 9$ Hz, $\nu_I = 0.5$ Hz. The resulting steady-state voltage is $V_{ss} = -53$ mV (thin solid line) with a standard deviation of $\sigma_V^2/2 = 1.38$ mV (thin dotted line). Three sample paths (different colors) are shown from the moment when the neuron starts to integrate its inputs ($\tau_{ref} = 2$ ms after the previous spike) until $V(t)$ reaches threshold for the first time. The time it takes for this to happen is called the first passage time and the total time in between two consecutive spikes is the inter-spike interval (ISI). Threshold ($V_{th} = -50$ mV) is shown as a thick dashed line (Continued).

Figure 15.1

Inset: Snapshot of the blue sample path from 323 to 326 ms shows the 0.2 mV discontinuities in $V(t)$ due to the synaptic inputs. **(B)** Current in top panel averaged over 1 ms time bins. Each point represents the average current into the neuron in the previous millisecond. For clarity of presentation, consecutive points are joined by lines. **Right**. Histogram of currents from left panel. The smooth blue line represents the distribution of $(1/\Delta t)\int_t^{t+\Delta t}\bar{I}(t')dt'$, where $\bar{I}(t)$ is the current into the cell in the diffusion approximation, Equation (15.10), and $\Delta t = 1$ ms. **(C)** Same as **A**, but with inputs now described by the diffusion approximation. The macroscopic structure of the sample paths is very similar. The differences between the Poisson input and the diffusion approximation can only be appreciated by looking at the inset.

of probability represents an extra probability current $S^{reset}(V,t)$, that adds to the current $S(V,t|V_0,t_0)$ associated to the sub-threshold dynamics of $V(t)$. Taking this into account, one can rewrite the Fokker-Planck equation like

$$\frac{\partial}{\partial t}\rho(V,t|V_0,t_0) = -\frac{\partial}{\partial V}[S(V,t|V_0,t_0) + S^{reset}(V,t)]. \tag{15.21}$$

Since this injection only results in a change of probability mass in $V = V_{reset}$, the new current is given by

$$S^{reset}(V,t) = \nu(t-\tau_{ref})\Theta(V - V_{reset}). \tag{15.22}$$

To find the solution for the steady-state distribution $\rho_{ss}(V)$, we insert expression (15.22) into the Fokker-Planck equation (15.21), and look for time independent solutions by setting the left hand side of this equation to zero,

$$\frac{\partial}{\partial V}[\frac{(V-V_{ss})}{\tau_m}\rho_{ss}(V)] + \frac{\sigma_V^2}{2\tau_m}\frac{\partial^2}{\partial V^2}\rho_{ss}(V) = -\nu\ \delta(V - V_{reset}). \tag{15.23}$$

Solving this equation with the boundary conditions (15.19-15.20), one obtains the following expression for the steady-state distribution [24]

$$\rho_{ss}(V) = \frac{2\nu\tau_m}{\sigma_V}\exp\left(-\frac{(V-V_{ss})^2}{\sigma_V^2}\right)\int_{\frac{V-V_{ss}}{\sigma_V}}^{\frac{V_{th}-V_{ss}}{\sigma_V}}\Theta(x - \frac{V_r - V_{ss}}{\sigma_V})e^{x^2}dx. \tag{15.24}$$

The function $\rho_{ss}(V)$ gives the fraction of cells in a non-refractory state with depolarizations in $(V, V+dV)$ in the steady state. Taking into account also the fraction $\nu\tau_{ref}$ of neurons in a refractory state, the steady state firing rate ν can be found by the normalization condition

$$\int_{-\infty}^{V_{th}}\rho_{ss}(V)dV\ +\ \nu\tau_{ref} = 1. \tag{15.25}$$

Plugging expression (15.24) into this equation and solving for ν one gets

$$\frac{1}{\nu} = \tau_{ref} + \tau_m\sqrt{\pi}\int_{\frac{V_r-V_{ss}}{\sigma_V}}^{\frac{V_{th}-V_{ss}}{\sigma_V}} e^{x^2}(1+\mathrm{erf}(x))dx, \tag{15.26}$$

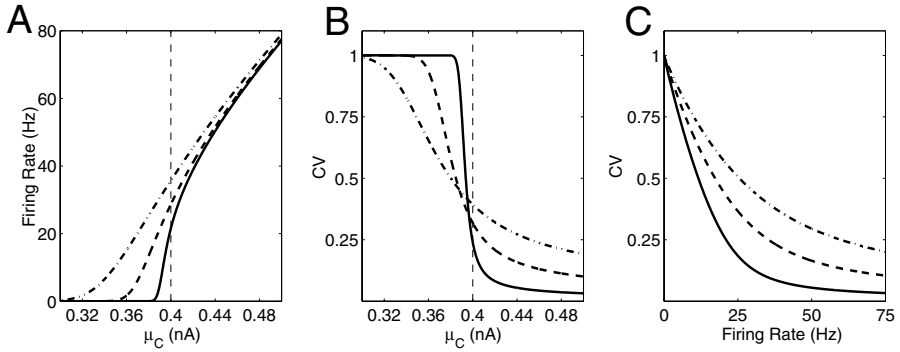

Figure 15.2

Firing rate **(A)** and CV **(B)** of the LIF neuron as a function of the mean current μ_C for three values of the effective standard deviation in the voltage $\sigma_V = 0.1$ mV (solid), 1 mV (dashed) and 4 mV (dot-dashed). **(C)** CV as a function of the mean firing rate when μ_C is varied as in the three curves in **A-B**. The parameters of the cell are $C_m = 0.2$ nF, $g_L = 20$ nS ($\tau_m = 10$ ms), $\tau_{ref} = 2$ ms, $V_L = -70$ mV, $V_{th} = -50$ mV and $V_r = -60$ mV.

where $\text{erf}(x) = 2/\sqrt{\pi}\int_0^x e^{u^2}du$.

Once the firing rate is known, the following recursive relationship between the moments of the first-passage time distribution of the Ornstein-Uhlenbeck process (see e.g., [117]) can be used to find $< T_{fp}^2 >$

$$\frac{\sigma_V^2}{2}\frac{d^2 < T_{pf}^k >}{dx^2} + (V_{ss} - x)\frac{d < T_{fp}^k >}{dx} = -k < T_{fp}^{k-1} >, \tag{15.27}$$

where $x = V_r$. Given that $< T_{fp} >= 1/\nu$ in Equation (15.26), the CV of the ISI is given by [19]

$$CV^2 \equiv \frac{< T_{fp}^2 > - < T_{fp} >^2}{< T_{fp} >^2} = 2\pi\nu^2 \int_{\frac{V_r - V_{ss}}{\sigma_V}}^{\frac{V_{th} - V_{ss}}{\sigma_V}} dx e^{x^2} \int_{-\infty}^{x} dy e^{y^2}(1 + \text{erf}(y)). \tag{15.28}$$

In Figures 15.2A-B we plot the mean firing rate and CV of the LIF neuron as given by Equations (15.26,15.28), as a function of the mean current μ_C for various values of the effective standard deviation in the voltage σ_V. The firing rate (Figure 15.2A) is a monotonic increasing function of the average current. Qualitatively, it starts to rise when the average current comes within a standard deviation of the current threshold, defined as $I_{th} \equiv g_L(V_{th} - V_L)$, and shown as a vertical dashed line in Figure 15.2A. It increases supra-linearly with μ_C for sub-threshold mean currents, and sub-linearly when the mean current is above threshold, eventually saturating at $1/\tau_{ref}$. Therefore, for a wide range of values of σ_V, $\mu_C \sim I_{th}$ is close to the point where the curvature of $\nu(\mu_C)$ changes sign. The deterministic current threshold I_{th} also marks

the transition between two different behaviors of the CV (Figure 15.2B). When the mean input is sub-threshold ($\mu_C < I_{th}$), spike discharge is triggered by fluctuations in the current, and spike trains are irregular. Therefore, in this regime, the CV is high, close to one. For supra-threshold mean current ($\mu_C > I_{th}$), the CV decays to zero and spiking becomes regular. The sharpness of this transition depends on σ_V. When the fluctuations are small, the transition is very sharp, so as soon as the neuron starts firing, it does so in a regular fashion. For large values of σ_V the transition is smooth (and if σ_V is large enough the CV can first increase noticeably from one for $\mu_C < I_{th}$, not shown), so the neuron fires initially irregularly and becomes progressively more regular as μ_C becomes much larger than I_{th}. In Figure 15.2C, the CV is plotted as a function of the firing rate for three values of σ_V as μ_C is increased gradually. In general, when the firing rate increases as a result of an increase in the mean current, the CV decreases. This decrease is faster with smaller fluctuations.

15.2.5 Effect of synaptic time constants

So far, we have assumed that the post-synaptic currents (PSPCs) are delta-functions, without any duration or temporal kinetics. In reality, synaptic currents rise and decay with time constants that range from 1 ms to several hundred ms. To incorporate a finite time constant of post-synaptic currents into the theoretical framework described in the previous section, we consider a variable $I(t)$ which, upon arrival of a spike at t_{spk}, evolves according to

$$\tau_{syn}\frac{dI(t)}{dt} = -I(t) + J\delta(t - t_{spk}). \tag{15.29}$$

For non-zero τ_{syn}, it is now $I(t)$ that has a discontinuous jump of size J/τ_{syn} when a pre-synaptic spike arrives. For $t > t_{spk}$, $I(t)$ decays exponentially back to zero with a time constant τ_{syn}. Importantly, the total area under this PSC is J independently of the value of τ_{syn}. The previous scheme can therefore be recovered continuously by letting $\tau_{syn} \to 0$. Now, instead of injecting an instantaneous charge J for every spike, we spread this same amount of charge over a time τ_{syn}. The effect of this on the voltage is to smoothen the rise of the PSP. Now the PSP is continuous and given by

$$V(t) = V_L + \left(\frac{J}{C_m}\right)\frac{\tau_m}{\tau_m - \tau_{syn}}\left[\exp(-\frac{t}{\tau_m}) - \exp(-\frac{t}{\tau_{syn}})\right]\Theta(t - t_{spk}), \tag{15.30}$$

with a rise time τ_{syn} and a decay time τ_m. If the membrane and the synaptic time constants are equal, the PSP has the shape of an α-function

$$V(t) = V_L + \left(\frac{J}{C_m}\right)\frac{t}{\tau_m}\exp(-\frac{t}{\tau_m})\Theta(t - t_{spk}). \tag{15.31}$$

The most important effect of a non-zero synaptic constant on the Fokker-Planck scheme presented above is the appearance of temporal correlations in the afferent current. The total synaptic input becomes a process with mean $< I >= \mu_C$, and an

exponential two-point correlation function $C_C(t,t') \equiv < (I(t) - < I >)(I(t') - < I >) >$ given by

$$C_C(t,t') = (\sigma_C^2/2\tau_{syn}) \exp(-|t-t'|/\tau_{syn}). \tag{15.32}$$

Using once again the diffusion approximation to replace the input to $I(t)$ by a Gaussian process with the same mean and correlation function, and defining $\delta I(t) = I(t) - \mu_C$, the Langevin equations of the process now read

$$\tau_m \frac{dV(t)}{dt} = -(V(t) - V_{ss}) + \frac{\delta I(t)}{g_L} \tag{15.33}$$

$$\tau_{syn} \frac{d}{dt} \delta I(t) = -\delta I(t) + \sigma_C \eta(t). \tag{15.34}$$

Although $V(t)$ is not Markovian anymore (knowledge of $I(t)$, in addition to $V(t)$, is needed to determine $V(t+dt)$ probabilistically), $V(t)$ and $I(t)$ *together* constitute a bi-variate Markov process [54, 99]. From Equations (15.33, 15.34) one can therefore derive a Fokker-Planck equation characterizing the evolution in time of the joint probability of V and I. However, the presence of temporal correlations in $I(t)$ makes the calculation of the firing rate much more involved than for the simple Ornstein-Uhlenbeck case and, indeed, the mean first-passage time can only be obtained in the case where $\tau_{syn} \ll \tau_m$, using perturbation theory on the parameter $k \equiv \sqrt{\tau_{syn}/\tau_m} \ll 1$ [23, 34, 40, 61, 67]. We present here only the final result: the firing rate is given by

$$\nu_{syn}(k) = \nu + k\alpha\sigma_V \left(\frac{\partial \nu}{\partial V_{th}} + \frac{\partial \nu}{\partial V_r} \right) + O(k^2), \tag{15.35}$$

where ν is the firing rate of the white noise case, Equation (15.26),

$$\alpha = -\zeta(1/2)/\sqrt{2} \sim 1.03$$

and ζ is the Riemann zeta function [2]. Note that the firing rate calculated in [23] does not include the term proportional to $\partial\nu/\partial V_r$, because of the approximation made in that paper, namely the neuron was assumed to be in the sub-threshold regime, in which the dependency of the mean firing rate on the reset potential is very weak.

Another way to write Equation (15.35) is to replace the threshold V_{th} and V_r in the expression for the mean first-passage time obtained for a white noise current (15.26), by the following effective k-dependent expressions

$$V_{th}^{eff} = V_{th} + \sigma_V \alpha k \tag{15.36}$$

$$V_r^{eff} = V_r + \sigma_V \alpha k. \tag{15.37}$$

This first order correction is in good agreement with the results from numerical simulations for $\tau_{syn} < 0.1\ \tau_m$. To extend the validity of the result to larger values of τ_{syn}, a second order correction can be added to the effective threshold, with coefficients determined by a fit to numerical simulations with values of τ_{syn} up to τ_m

[23]. A finite synaptic time constant leads to synaptic filtering of the pre-synaptic inputs which, in general, leads to a reduction in post-synaptic firing rates [23]. This effect is more pronounced for sub-threshold mean currents, since in this regime the neuronal firing results from the fluctuations in the current which can be filtered out by the synapse. Note that the effect of increasing τ_{syn} is to spread out in time the same amount of charge influx into the cell. Since charge is constantly leaking out of the cell membrane, the longer τ_{syn}, the lower the overall magnitude of the voltage fluctuations.

Finally, one can also compute perturbatively the firing rate in the large synaptic time constant limit [83]. An interpolation between the two limits gives rather accurate results in the whole range of all synaptic time constants. A similar approach has been used to compute the firing rate of another simple spiking neuron, the quadratic neuron [22].

15.2.6 Approximate treatment of realistic synaptic dynamics

Real synaptic currents can depart in at least three ways from the currents considered until now: (i) individual post-synaptic currents can in some circumstances sum non-linearly, due to receptor saturation; (ii) post-synaptic currents are voltage dependent, because synaptic activation occurs as conductance change rather than current increase, and because the maximal conductance can itself be voltage-dependent; (iii) multiple synaptic time scales are present, due to the different kinetics of the AMPA, $GABA_A$, and NMDA receptors. We first describe the standard biophysical model for describing post-synaptic currents (see also e.g., [33, 120]), and then discuss separately how the three issues can be dealt with using approximate treatments.

15.2.6.1 Biophysical models of post-synaptic currents

Synaptic activation opens ion channels in the post-synaptic cell. The amount of current flowing through these channels is proportional to the product of the number of open channels times the driving force of the synaptic current:

$$I_{syn}(t) = g_{syn}(V)\, s(t)(V(t) - V_{syn}), \tag{15.38}$$

where $g_{syn}(V)$ is the (possibly voltage-dependent) maximal conductance, $s(t)$ is a gating variable measuring the fraction of open channels at the synapse and V_{syn} is the synaptic reversal potential. The term $V - V_{syn}$ is the driving force of the synapse, and it determines its polarity, i.e., whether a synaptic current is depolarizing ($V - V_{syn} < 0$) or hyper-polarizing ($V - V_{syn} > 0$). In the presence of a driving force term, all synaptic inputs are voltage-dependent.

We consider two types of kinetic schemes for the gating variable $s(t)$. If the underlying dynamics of the synaptic channels is fast compared with the typical firing rates of the spike trains at the synapse, the synapse is usually far from saturation and a linear kinetic scheme is appropriate. Additionally, in this situation the rise time of the post-synaptic currents (PSCs) is so fast that it can be considered instantaneous,

so that the kinetics can also be approximated by a first order system, i.e.,

$$\frac{ds(t)}{dt} = -\frac{s(t)}{\tau_s} + \sum_k \delta(t - t_k), \tag{15.39}$$

where $\sum_k \delta(t - t_k)$ represents the pre-synaptic spike train arriving at the synapse. The average fraction of open channels is linear in the firing rate ν across the synapse $\bar{s} = \tau_s \nu$. Since $s(t)$ is a fraction, i.e., necessarily less than one, this description is appropriate as long as $\nu \ll 1/\tau_s$. We will use it for the description of $GABA_A$R- and AMPAR-mediated transmission, which have synaptic time constants of $\tau_{GABA_A} = 10$ ms, and $\tau_{AMPA} = 2$ ms [13, 15, 58, 72, 127, 129]. This approximation is reasonable if $\nu < 1/\tau_{GABA_A} = 100$ Hz.

If the underlying channel dynamics is of the order of, or slower, than the typical inter-spike intervals of the spike trains crossing the synapse (which is the case for the NMDAR-mediated PSPcs, with a time constant of $50 - 100$ ms), the channels decay slowly in between spikes, and a few spikes in a train at hight frequencies can recruit a fraction of open channels close to unity. In this case, the effect of subsequent spikes is bounded by the saturation of all post-synaptic receptors, hence spikes sum non-linearly. Also, for slow channel dynamics, the PSC rise times are on the order of the fastest time-scales of the system (a few milliseconds), and can no longer be neglected. A non-linear, second order scheme, provides an accurate description of the kinetics of the gating variable $s(t)$ in these conditions:

$$\frac{ds(t)}{dt} = -\frac{s(t)}{\tau^{decay}} + \alpha x(t)(1 - s(t)) \tag{15.40}$$

$$\frac{dx(t)}{dt} = -\frac{x(t)}{\tau^{rise}} + \sum_k \delta(t - t_k). \tag{15.41}$$

For slow synaptic dynamics, the average gating variable is no longer a linear function of the pre-synaptic rate unless the firing rate is only a few Hz.

15.2.6.2 Average gating variable vs. rate in the non-linear model

An immediate consequence of this non-linear summation is that the average value of the gating variable becomes a non-linear function of the average firing rate of the spike train through the synapse. This function depends on the statistics of the spike train. If the spike train is regular, an approximation can be obtained by replacing $\sum_k \delta(t - t_k)$ by the mean firing rate ν of the spike train in Equation (15.41). In this case the average of the gating variable becomes

$$\bar{s} = \frac{\tau\nu}{1 + \tau\nu}, \tag{15.42}$$

where the effective time constant is equal to $\tau = \tau^{rise}\tau^{decay}\alpha$.

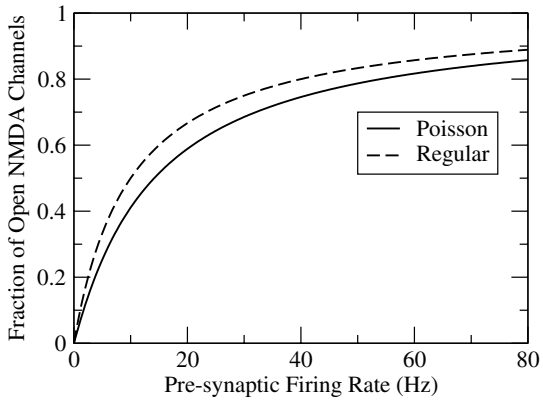

Figure 15.3

Average fraction of open NMDA channels as a function of the pre-synaptic firing rate. The solid (dashed) line, calculated with Equation (15.43) (Equation (15.42)), corresponds to the case when the spike train at the synapse is Poisson (periodic).

If the spike train is Poisson, the expression for $\bar{s}$ is [25]

$$\bar{s} = \frac{\nu\tau}{1+\nu\tau}\left(1+\frac{1}{1+\nu}\sum_{n=1}^{\infty}\frac{(-\alpha\tau^{rise})^n T_n}{(n+1)!}\right) \equiv \psi(\nu)$$
$$T_n = \sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{\tau^{rise}(1+\nu\tau)}{\tau^{rise}(1+\nu\tau)+k\tau^{decay}}. \qquad (15.43)$$

We will use this description for NMDAR-mediated transmission, with parameters $\tau^{decay}_{\text{NMDA}} = 100$ ms, $\tau^{rise}_{\text{NMDA}} = 2$ ms, and $\alpha = 0.5$ KHz. The effective NMDA time constant is thus $\tau_{\text{NMDA}} = 100$ ms. In Figure 15.3, the average gating variable of an NMDA synapse is plotted as a function of the average pre-synaptic firing rate ν, for the case of a regular and a Poisson input spike train. Note that, due to the saturation term on the right of Equation (15.40), the gating variable starts to saturate when the pre-synaptic rate becomes larger than $\sim 1/\tau_{\text{NMDA}} \sim 10$ Hz.

15.2.6.3 Voltage-dependence of the post-synaptic currents

A post-synaptic current is voltage-dependent because of its driving force; in addition the maximal conductance can also be voltage-dependent, as in the case of the NMDA channels [87].

In general, even if the maximal conductance does not depend on the voltage, the voltage dependence induced by the driving force term in the unitary synaptic current (15.38) modifies the previous framework for calculating the output firing rate of the cell in several ways.

Let us separate the time course of the gating variable into a deterministic component, associated to its temporal average (we assume stationary inputs, in which case

the average of the gating variable is constant), and a fluctuating component due to the stochastic nature of the spike trains in our model, i.e., $s_{syn}(t) = \bar{s}_{syn} + \delta s_{syn}(t)$. The unitary synaptic current, Equation (15.38), now becomes

$$I_{syn}(t) = g_{syn}\bar{s}_{syn}(V(t) - V_{syn}) + g_{syn}\delta s_{syn}(t)(V(t) - V_{syn}). \tag{15.44}$$

The main complication due to the driving force is that now the fluctuating component of the synaptic current (second term in the right-hand side of Equation (15.44)) becomes voltage dependent. This multiplicative dependence of the fluctuations on the membrane potential renders a rigorous treatment of the fluctuations in the current difficult. To avoid this complication, we replace the voltage by its average $\bar{V}$ in the driving force for the fluctuating component of the synaptic current, so that

$$I_{syn}(t) \sim g_{syn}\bar{s}_{syn}(V(t) - V_{syn}) + g_{syn}\delta s_{syn}(t)(\bar{V} - V_{syn}). \tag{15.45}$$

The deterministic part of the current $g_{syn}\bar{s}_{syn}(V(t) - V_{syn})$ can be dealt with easily by noting that $g_{syn}\bar{s}_{syn}$ can be absorbed in the leak conductance, and $g_{syn}\bar{s}_{syn}V_{syn}$ can be absorbed in the resting membrane potential V_L.

The resulting effect on neuronal properties is an increase in the total effective leak conductance of the cell

$$g_L \rightarrow g_L + g_{syn}\bar{s}_{syn}, \tag{15.46}$$

which is equivalent to a decrease of the membrane time constant from $\tau_m = C_m/g_L$, to $\tau_m^{eff} = C_m/(g_L + g_{syn}\bar{s}_{syn}) = \tau_m/\alpha_{\tau_m}$. Thus, the synaptic input makes the neuron leakier by a factor equal to the relative increase in conductance due to synaptic input ($\alpha_{\tau_m} = 1 + \bar{s}_{syn}g_{syn}/g_L$). The resting (or steady-state) membrane potential is also re-normalized

$$V_L \rightarrow \frac{g_L V_L + g_{syn}\bar{s}_{syn}V_{syn}}{g_L + g_{syn}\bar{s}_{syn}}, \tag{15.47}$$

and becomes a weighted average of the different reversal potentials of the various synaptic currents, where each current contributes proportionally to the relative amount of conductance it carries.

Voltage-dependence of NMDA channels. For NMDA channels to open, binding of neurotransmitter released by the pre-synaptic spike is not enough. The post-synaptic cell must also be sufficiently depolarized to remove their blockade by magnesium. It is conventional to model this using a voltage-dependent maximal conductance [66]:

$$g_{\text{NMDA}}(V) = \frac{g_{\text{NMDA}}}{(1 + ([\text{Mg}^{2+}]/\gamma)\exp(-\beta V(t))} \equiv g_{\text{NMDA}}\frac{1}{J(V(t))}, \tag{15.48}$$

with $[\text{Mg}^{2+}] = 1$ mM, $\gamma = 3.57$ and $\beta = 0.062$. To be able to incorporate this effect into the framework described in the previous sections, we linearize the voltage dependence of the NMDA current around the average voltage $\bar{V}$, obtaining

$$\frac{V(t) - V_E}{J(V(t))} \sim \frac{V(t) - V_E}{J(\bar{V})} + (V(t) - \bar{V})\frac{J(\bar{V}) - \beta(\bar{V} - V_E)(1 - J(\bar{V}))}{J^2(\bar{V})} + O((V(t) - \bar{V})^2) + \ldots \tag{15.49}$$

This linear approximation is very accurate for the range of values of $V(t)$ between reset and threshold [25]. Using it, the non-linear voltage-dependent NMDA current can be emulated by a linear current with a renormalized maximal conductance and reversal potential. Defining

$$I_{\text{NMDA}}(t) \equiv g^{eff}_{\text{NMDA}} s_{\text{NMDA}} (V(t) - V^{eff}_E), \tag{15.50}$$

the renormalized parameters read

$$\begin{aligned} g^{eff}_{\text{NMDA}} &= g_{\text{NMDA}} \frac{J(\bar{V}) - \beta(\bar{V} - V_E)(1 - J(\bar{V}))}{J^2(\bar{V})} \\ V^{eff}_E &= \bar{V} - \frac{g_{\text{NMDA}}}{g^{eff}_{\text{NMDA}}} \left(\frac{\bar{V} - V_E}{J(\bar{V})} \right). \end{aligned} \tag{15.51}$$

To give a qualitative idea of the properties of the linearized NMDA current, using $V_E = 0$ mV and $\bar{V} = -55$ mV, one obtains $g^{eff}_{\text{NMDA}} \sim -0.22 g_{\text{NMDA}}$ and $V^{eff}_E \sim -81.8$ mV. Since the slope of the $I - V$ plot for the original current is negative at voltages near the average depolarization of the neuron, the effective NMDA conductance is negative. However, since the effective reversal potential is lower than the cell's typical depolarization, the total effect of the effective NMDA current is depolarizing, as it should.

Calculation of the average voltage $\bar{V}$. To complete our discussion of the voltage-dependence, we need to compute the average voltage $\bar{V}$, that enters in Equation (15.45) and Equations (15.51). This can easily be done using Equation (15.24). The result is

$$\begin{aligned} \bar{V} &= \int_{-\infty}^{V_{th}} V[\rho_{ss}(V) + \nu \tau_{ref} \delta(V - V_r)] dV = \\ &= V_{ss} - (V_{th} - V_r) \nu \tau^{eff}_m - (V_{ss} - V_r) \nu \tau_{ref}. \end{aligned} \tag{15.52}$$

15.2.6.4 Fluctuations in the synaptic current in the case of multiple synaptic time scales

The results of Section 15.2.5 can be applied when a single time scale is present in synaptic currents. This is obviously not the case when fluctuations are due to AMPA, GABA_A and NMDA currents. In the absence of rigorous results for fluctuations with multiple time scales, one has to resort to an approximation. The approximation is based on the fact that the longer the synaptic time constant, the more the fluctuations of the gating variable will be filtered out (see Section 15.2.5). Therefore, we expect the fluctuations in the GABA_A and NMDA currents to be smaller in magnitude than those associated to the AMPA currents. We thus neglect their contribution and assume that $\delta s_{\text{NMDA}}(t) = \delta s_{\text{GABA}}(t) \sim 0$.

15.2.6.5 Summary: firing statistics of a neuron with realistic AMPA, GABA_A and NMDA synaptic inputs

Here, we summarize the description of a LIF neuron that receives C_E excitatory synaptic inputs and C_I inhibitory synaptic inputs (Figure 15.1), with synapses de-

scribed by individual conductances $g_{j_{\text{AMPA}}}$ and $g_{j_{\text{NMDA}}}$, $j = 1,2,...,C_E$; $g_{j_{\text{GABA}}}$, $j = 1,2,...C_I$). In the presence of these inputs, Equation (15.1) now reads

$$\begin{aligned} C_m \frac{dV(t)}{dt} = & - g_L(V(t) - V_L) - \\ & - \left[\sum_{j=1}^{C_E} g_{j_{\text{AMPA}}} s_{j_{\text{AMPA}}}(t) + \frac{g_{j_{\text{NMDA}}} s_{j_{\text{NMDA}}}(t)}{J(V(t))} \right] (V(t) - V_E) - \\ & - \left[\sum_{j=1}^{C_I} g_{j_{\text{GABA}}} s_{j_{\text{GABA}}}(t) \right] (V(t) - V_I). \end{aligned} \tag{15.53}$$

For simplicity, we again assume that the synaptic conductances and the firing rates of all pre-synaptic inputs from the same sub-population are identical. Using the approximations described in the previous sections, this equation becomes

$$\begin{aligned} C_m \frac{dV(t)}{dt} = & - g_L(V(t) - V_L) - \\ & - C_E \left[g_{\text{AMPA}} \bar{s}_{\text{AMPA}} \right] (V(t) - V_E) - \\ & - C_E \left[g^{eff}_{\text{NMDA}} \bar{s}_{\text{NMDA}} \right] (V(t) - V^{eff}_E) - \\ & - C_I \left[g_{\text{GABA}} \bar{s}_{\text{GABA}} \right] (V(t) - V_I) + \delta I(t), \end{aligned} \tag{15.54}$$

where $\bar{s}_{\text{AMPA}} = \nu_E \tau_{\text{AMPA}}$, $\bar{s}_{\text{GABA}_\text{A}} = \nu_I \tau_{\text{GABA}}$ and $\bar{s}_{\text{NMDA}} = \psi(\nu_E)$ where the function ψ is defined in Equation (15.43), and the fluctuations are described by

$$\tau_{\text{AMPA}} \frac{d}{dt} \delta I(t) = -\delta I(t) + \sigma_{eff}\, \eta(t) \tag{15.55}$$

$$\sigma^2_{eff} = g^2_{\text{AMPA}} (\bar{V} - V_E)^2 C_E \bar{s}_{\text{AMPA}} \tau_{\text{AMPA}}. \tag{15.56}$$

Since all the deterministic components of the current are now linear in the voltage, the equations describing the membrane potential dynamics can be expressed as

$$\tau^{eff}_m \frac{dV(t)}{dt} = -(V(t) - V_{ss}) + \frac{\delta I(t)}{g^{eff}_L} \tag{15.57}$$

$$\tau_{\text{AMPA}} \frac{d}{dt} \delta I(t) = -\delta I(t) + \sigma_{eff}\, \eta(t). \tag{15.58}$$

The effective membrane time constant is

$$\tau^{eff}_m = \frac{C_m}{g^{eff}_L} = \tau_m \frac{g_L}{g^{eff}_L}, \tag{15.59}$$

and the effective leak conductance of the cell is the sum of the passive leak conductance plus the increase in the conductances associated to all the synaptic inputs to the cell

$$g^{eff}_L = g_L + g_{\text{AMPA}} C_E \bar{s}_{\text{AMPA}} + g^{eff}_{\text{NMDA}} C_E \bar{s}_{\text{NMDA}} + g_{\text{GABA}} C_I \bar{s}_{\text{GABA}}. \tag{15.60}$$

In *in vivo* experiments, it was estimated that, even when neurons fire at low rates (a few hertz), g_L^{eff} is at least 3-5 times larger than g_L [32], therefore τ_m^{eff} is $3-5$ shorter than τ_m. For example, if $\tau_m = 10$ ms, then $\tau_m^{eff} \simeq 2-3$ ms. When neurons fire at higher rates (leading to larger synaptic conductances), the value of g_L^{eff} would be significantly larger and τ_m^{eff} would be even smaller.

The steady-state voltage V_{ss} now becomes

$$V_{ss} = [\, g_L V_L + (C_E g_{\text{AMPA}} \bar{s}_{\text{AMPA}}) V_E + (C_E g_{\text{NMDA}}^{eff} \bar{s}_{\text{NMDA}}) V_E^{eff} + (C_I g_{\text{GABA}} \bar{s}_{\text{GABA}}) V_I] / g_L^{eff}. \tag{15.61}$$

Note that the steady state potential V_{ss} is bounded between the highest and the lowest reversal potentials of the four currents to the neuron. In particular, it can never become lower than V_I. Thus, no matter how strong inhibition is, in this model the average membrane potential will fluctuate around a value not lower than the reversal potential of the inhibitory synaptic current, e.g., at approximately -70 mV.

Since Equations (15.57) and (15.58) can be mapped identically to Equations (15.33) and (15.34), one can now use equation (15.26) to compute the firing rate of a neuron,

$$\nu_{post} = \left[\tau_{ref} + \tau_m^{eff} \sqrt{\pi} \int_{\frac{V_r^{eff} - V_{ss}}{\sigma_{eff}}}^{\frac{V_{th}^{eff} - V_{ss}}{\sigma_{eff}}} e^{x^2} (1 + \text{erf}(x)) \right]^{-1}. \tag{15.62}$$

where τ_m^{eff} and V_{ss} are given by Equations (15.59-15.61); V_{th}^{eff} and V_r^{eff} are given by Equations (15.36-15.37). Note that now, the average voltage in the steady state $\bar{V}$ plays a role in determining the firing rate, through both V_{ss} and σ_{eff}. Since $\bar{V}$ is related linearly to the firing rate (Equation (15.52)), the firing rate is not an explicit function of the synaptic input. Even if the inputs are entirely external (feedforward), and all the synaptic conductances are fixed, $\bar{V}$ still depends on the post-synaptic firing rate ν itself. Therefore, ν must be determined self-consistently.

Equation (15.62) constitutes a non-linear input-output relationship between the firing rate of our post-synaptic neuron and the average firing rates ν_E and ν_I of the pre-synaptic excitatory and inhibitory neural populations. This input-output function is conceptually equivalent to the simple threshold-linear or sigmoid input-output functions routinely used in firing-rate models. What we have gained from all these efforts is a firing-rate model that captures many of the underlying biophysics of the real spiking neurons. This makes it possible to quantitatively compare the derived firing-rate model with detailed numerical simulations of the irregularly firing spiking neurons, an important step to relate the theory with neurophysiological data.

15.3 Self-consistent theory of recurrent cortical circuits

A cortical microcircuit receives afferent inputs and sends efferent outputs downstream, thereby information processing is carried out in a 'feedforward' fashion. At the same time, interesting computations may be accomplished by horizontal or recurrent synaptic connections within the local network. The relative importance of feedforward versus recurrent processing is likely to be different for each specific task, and vary from one cortical area to another. In the primary visual cortex (V1), recurrent synaptic connections are quite abundant [76]; their functional importance (such as to the generation of orientation selectivity) has been the subject of intense debate [39, 111]. Recently, there is growing interest in the recurrent networks of association cortical areas, such as the parietal cortex or prefrontal cortex. This interest was primarily motivated by the observation of 'working memory neurons' in these cortices. In experiments when an animal is required to remember a transient stimulus cue across a delay period of a few seconds, between the cue presentation and behavioral response, neurons in association areas display stimulus-selective, elevated persistent activity across the delay period [44, 55]. Since the elevated neural activity can be triggered by a brief input but outlast it for many seconds, persistent activity cannot be explained by a feedforward mechanism. It has been hypothesized that persistent activity can be self-sustained by synaptic 'reverberations' within a strongly recurrent local network (see for a review [6, 121]).

We will now discuss how a recurrent network of neurons can be described by mean-field theory. We will first consider how stationary states of such networks can be obtained in a self-consistent way. Next we discuss dynamical approaches which allow an assessment of the stability of the stationary states. Then, an example from an one-population network of excitatory cells is analyzed in detail, introducing the concept of bistability by means of a graphical analysis, and relating it to the phenomenon of persistent neural activity in working memory. A more detailed model of a network for object working memory is then described. Finally, we discuss the possibility of multi-stability in cortical networks in which both excitation and inhibition are strong, but roughly cancel each other out.

15.3.1 Self-consistent steady-state solutions in large unstructured networks

In a recurrent network, the post-synaptic neuron and its pre-synaptic inputs are part of the same network, and hence, if the activity in the network is not changing, their firing activity must, in a statistical sense, be the same. If, as we discussed in Section 15.2, the output firing rate only depends on the average rate of the inputs, then equalizing pre- and post-synaptic activity will yield an equation that determines the firing rates in the possible stationary states of the network. This is a very general necessary condition that has to be met in any steady-state solution of the network dynamics.

However, in order for us to be able use the input-output relationship found in Section 15.2, the dynamics of the network should be such that the network properties in this stationary states are consistent with the assumptions we made in Section 15.2. Thus, these assumptions impose several additional conditions that the steady-states should obey to be truly self-consistent:

- The fluctuations in the inputs must be approximately independent from neuron to neuron. This condition will be trivially satisfied when the major part of the noise comes from external independent sources. It will also be satisfied when the network is sparsely connected, i.e., when the connection probability between any pair of neurons is weak. In this case, the 'noise' term coming from the recurrent network itself becomes uncorrelated from neuron to neuron [19, 24, 118, 119].

- The probability of a spike being emitted in the network at any moment must be constant in time. Thus, the steady state must be stable with respect to any instability that leads to non-stationary global network activity, such as synchronized oscillations.

- The neurons must emit approximately as Poisson processes for the input-output relationship to be valid. This is in general expected to be true when the average total input to the neurons is sub-threshold, which will be the situation of interest in our discussion.

Several types of local network connectivity and synaptic structure are conceivable. They differ mainly in the source of the fluctuations in the synaptic current to the neurons. One approach is to investigate the behavior of the network as a function of its size N and of the number of connections per cell C. The strategy is to scale the PSP size $\bar{J} \equiv J/C_m$ (a measure of the synaptic strength) with C, and study the behavior of the network as $C \to \infty$. An advantage of this procedure is that **a)** the behavior of the network is much simpler and easier to analyze in the $C = \infty$ limit, and **b)** network behaviors which are only quantitatively different for finite C, become qualitatively different as C becomes infinite. Additionally several of the technical assumptions we had to make in Section 15.2 become exact in this limit.

Alternatively, one can assume that N and C are large but finite. In this case one does not assume any specific scaling of the PSP size with C, but rather uses the formulas for arbitrary values of these parameters (as in the previous sections) and studies the behavior of the resulting equations when they take realistic values informed by the available experimental data. Some of the hypothesis made in the calculations will only be verified approximately, but the theory will be more directly comparable to experiments, where C and J/C_m are finite.

In the following subsections we describe the self-consistency equations obtained in each of these scenarios. For ease of exposition, our discussion will be carried out in the simplest case of neurons connected through instantaneous synapses, unless specified otherwise.

15.3.1.1 Fully connected networks; External noise

In a fully connected network $C = N$. In such a network, all neurons see essentially the same recurrent input. In order to obtain a finite mean synaptic input (the 'mean-field') the PSPs are usually assumed to scale as $1/N$. In this case, the total synaptic input has a mean of order 1, and noise of order $1/\sqrt{N}$. This is apparent in the equations for the moments of the diffusion process (15.8), where it is clear that in this case only the first moment remains non-zero as $N \to \infty$. Thus, the recurrent component of the synaptic current becomes deterministic. In this framework, noise is assumed to come from unspecified external sources, and is assumed to be independent for each neuron.

Let us consider the simple case of a single neural population. We express the input-output relationship of the cell, Equation (15.62), as $\nu_{\text{post}} = \phi(\mu(\nu_{\text{pre}}), \sigma)$, making explicit the dependency of the rate of the post-synaptic cell on the rate of its pre-synaptic inputs through the mean μ and standard deviation σ of the fluctuations in the total afferent current. In the steady-state $\nu_{\text{post}} = \nu_{\text{pre}} \equiv \nu$, so the self-consistent relationship can be written as

$$\nu = \phi(\mu(\nu), \sigma), \tag{15.63}$$

where the mean input current $\mu(\nu) = \mu_{ext} + \mu_{rec}$ is the sum of an external tonic current μ_{ext} and of a mean recurrent synaptic current $\mu_{rec}(\nu)$. The noise component of the current comes exclusively from outside the network, i.e., $\sigma = \sigma_{ext}$. The solution of Equation (15.63) can be obtained graphically by plotting $\phi(\mu(\nu), \sigma)$ vs. ν and by looking at its intersections with the diagonal line [8, 11, 120]. Alternatively, one can plot the f-I curve $\nu = \phi(\mu, \sigma)$ vs. μ and look at its intersections with $\nu = \mu^{-1}(\mu)$ vs. μ [20]. An example of this kind of analysis is given below. When several populations are present, the framework is extended by adding one self-consistency equation per population (again see specific example below).

A general feature of an all-to-all network is that the level of noise is independent of the activity in the recurrent network. Thus the activity of the neurons is modulated by changes in the mean current they receive. As we shall see below, this has consequences on the statistics of their spike trains, a consequence that can be tested experimentally.

15.3.1.2 The balanced state

The all-to-all network architecture with $1/N$ couplings, though simple, is not very realistic. In the cortex, synaptic couplings are much stronger than $1/N$ and neurons are not fully connected. This motivates the study of networks which are sparsely connected, and with stronger coupling. Let us consider a network in which each neuron receives C random connections from a total of N neurons. If the network is very sparse, i.e., if $N \gg C$, the probability that two neurons receive a substantial fraction of common inputs becomes very small, so the recurrent inputs in this network will be effectively uncorrelated between any pair of post-synaptic cells. Since the second infinitesimal moment of the diffusion process, which measures the fluctuations in the

synaptic current, scales as $C\bar{J}^2$ (see Equation (15.8)), to keep the fluctuations finite as $C \to \infty$, one should scale the synaptic couplings as $\bar{J} \sim 1/\sqrt{C}$. On the other hand, the mean synaptic input scales as $\bar{J}C \sim \sqrt{C}$ and diverges to plus or minus infinity (if the coupling is excitatory or inhibitory, respectively). Thus, one immediately sees that if the network is composed of a single excitatory population, the neuron will either be at saturation or totally silent for large C. To obtain plausible levels of activity in this framework, one needs, therefore, to introduce an inhibitory population.

Let us write $C_E = c_E C$, $C_I = c_I C$, where c_E and c_I are finite. As already anticipated, in order to keep the diffusion coefficient A_2 finite, the scaling $\bar{J}_{E,I} = j_{E,I}/\sqrt{C}$ must be used. The infinitesimal moments of the stochastic process become

$$\begin{aligned} A_1(V) &= -\frac{(V - V_{ss})}{\tau_m} + \sqrt{C}\left[j_E c_E \nu_E - j_I c_I \nu_I\right] \\ A_2 &= j_E^2 c_E \nu_E + j_I^2 c_I \nu_I \\ A_n &= C^{1-\frac{n}{2}}\left[j_E^n c_E \nu_E + (-1)^n j_I^n c_I \nu_I\right] \qquad n = 3, \ldots \end{aligned} \tag{15.64}$$

In this case, as $C \to \infty$, all terms of order $n > 2$ vanish, and Equation (15.7) becomes identical to the Fokker-Planck Equation (15.9). Therefore, in this case the diffusion approximation becomes exact. The second term in the drift coefficient A_1, which gives the mean current into the cell, diverges as $\sqrt{C}$. Thus, unless the excitatory and the inhibitory drives into the cell *balance* each other to within $1/\sqrt{C}$, the resulting massive excitatory or inhibitory drive will drive the neuron towards saturation, or total silence. [118] showed, in a recurrent network of binary neurons, that this balanced state can arise as a dynamically stable state in a very robust way. Using that neuronal model, a complete description of the temporal fluctuations of activity, beyond the Poisson assumption that we have been using, can be performed. As we shall now see, the equations which determine the average firing rate of the excitatory and inhibitory populations in the balanced state are very general, and applicable to any single neuron model that assumes that the different synaptic inputs to the cell are summed linearly.

Let us consider the two population network in which each population receives $C_{E,I}^{ext} = c_{E,I}^{ext} C$ excitatory Poisson inputs of rate ν^{ext} through synapses of strength $\bar{J}_{E,I}^{ext} = j_{E,I}^{ext}/\sqrt{C}$ from outside the network. We know that the fluctuating component of the current will be of order one and the mean inputs will be

$$\begin{aligned} \frac{\mu_E}{C_m} &= \sqrt{C}\left[j_{EE} c_{EE} \nu_E - j_{EI} c_{EI} \nu_I + j_E^{ext} c_E^{ext} \nu^{ext}\right] \\ \frac{\mu_I}{C_m} &= \sqrt{C}\left[j_{IE} c_{IE} \nu_E - j_{II} c_{II} \nu_I + j_I^{ext} c_I^{ext} \nu^{ext}\right]. \end{aligned} \tag{15.65}$$

Following the arguments presented in the previous sections, one can now *impose* that $\mu_{E,I}$ be order one, and see whether there is a stable self-consistent solution arising from this constraint. The simplest case is when $C = \infty$, which one expects to be also qualitatively correct for large but finite networks. In this case, a finite mean current can only be obtained if the balance is perfect, i.e., if the total excitation and inhibition

cancel each other precisely. This means that the terms in square brackets in Equation (15.65) have to vanish identically, leading to a set of two coupled linear equations [118]

$$\begin{aligned} j_{EE}c_{EE}\nu_E - j_{EI}c_{EI}\nu_I + j_E^{ext}c_E^{ext}\nu^{ext} &= 0 \\ j_{IE}c_{IE}\nu_E - j_{II}c_{II}\nu_I + j_I^{ext}c_I^{ext}\nu^{ext} &= 0, \end{aligned} \tag{15.66}$$

which implies that the self-consistent rates of the two populations become a linear function of the external input

$$\nu_E = k_E\ \nu^{ext} + O\left(\frac{1}{\sqrt{C}}\right); \qquad \nu_I = k_I\ \nu^{ext} + O\left(\frac{1}{\sqrt{C}}\right). \tag{15.67}$$

In contrast to a fully connected network, a balanced network dynamics is an intrinsic source of noise. In fact, even if the network is purely deterministic (with constant inputs instead of stochastic Poisson trains), and the external afferents are assumed to be regular, the balanced network can give rise to chaotic network dynamics and highly irregular neural activities [118, 119]. The firing rates in this network can, therefore, be determined self-consistently without making any assumptions about the specific single neuron model as long as the synaptic currents from different inputs are summed linearly. Although this is a quite remarkable result, the linearity of the self-consistent rates on the external input raises a fundamental problem from a computational perspective: does this mean that a balanced network cannot subserve non-linear behaviors such as bistability? Put it differently, can a network be both bistable and generate its own noise? We will come back to this issue below.

The arguments presented above are valid for synapses modeled as voltage-independent synaptic currents. With conductance-based synaptic currents, the situation is quite different, since in the large C limit, the total synaptic conductance diverges to infinity, hence effective neuronal time constant tends to zero. In this limit, the membrane potential is slaved to an effective 'steady-state' potential that stays finite in that limit [107]. Thus, there is no more 'balance' condition to be fulfilled in this situation, unless additional hypothesis are used. In any case, the simple balanced network picture is useful as a metaphor for networks with strong coupling and highly irregular firing of its constituent neurons. We will use the term 'balanced network' in this loose sense in the following.

15.3.1.3 Large but finite sparse networks

We again assume a large sparse network ($N \gg C$) so that the recurrent inputs to different neurons can still be assumed independent, but we now assume that the number of connections per neuron C is large ($C \gg 1$) but finite, and the coupling strength to be small (the unitary PSP size $J/C_m \ll (V_{th} - V_L)$) but finite [11]. For example, $N \sim 10{,}000$ and $C \sim 1{,}000$; $J/C_m \sim 0.1 - 0.3$ mV whereas $(V_{th} - V_L) \sim 10 - 15$ mV. Thus, in this case, both the mean synaptic input and the fluctuations around it depend on the firing rate of the pre-synaptic neurons. For the case of a single population the self-consistent equation becomes

$$\nu = \phi(\mu(\nu), \sigma(\nu)) \tag{15.68}$$

where the mean and variance of the current are given respectively by

$$\mu(\nu) = \mu_{ext} + CJ\nu \tag{15.69}$$
$$\sigma^2(\nu) = \sigma^2_{ext} + CJ^2\nu. \tag{15.70}$$

In this 'extended' mean-field theory, not only the mean inputs are included in the description, but also the fluctuations around the 'mean-field' are relevant. As emphasized above, this approach is only applicable to network states in which neurons fire in an approximately Poissonian way, and when the low connection probability makes the emission processes of neurons essentially uncorrelated. Moreover, since J and C are finite, this approach is only approximate. However, simulations show it gives very accurate results when $C\nu\tau_m$ is large (several hundreds) and $J/(C_m V_{th})$ is small (less than several percent), as seem to be the case in cortex [10, 24, 19]. Equation (15.68) can again be solved graphically to obtain the self-consistent, steady-state firing rates in the network (see below).

It is straightforward to extend this description to a two population network of excitatory and inhibitory neurons. The equations are, for finite C_E, C_I, (E-to-E) J_{EE}, (I-to-E) J_{EI}, (E-to-I) J_{IE}, and (I-to-I) J_{II}:

$$\nu_E = \phi(\mu_E, \sigma_E) \tag{15.71}$$
$$\nu_I = \phi(\mu_I, \sigma_I) \tag{15.72}$$
$$\begin{aligned}
\mu_E &= \mu_{extE} + [C_E J_{EE}\nu_E - C_I J_{EI}\nu_I] \\
\mu_I &= \mu_{extI} + [C_E J_{IE}\nu_E - C_I J_{II}\nu_I] \\
\sigma^2_E &= \sigma^2_{extE} + \left[C_E J^2_{EE}\nu_E + C_I J^2_{EI}\nu_I\right] \\
\sigma^2_I &= \sigma^2_{extI} + \left[C_E J^2_{IE}\nu_E + C_I J^2_{II}\nu_I\right].
\end{aligned} \tag{15.73}$$

The stationary states of these two population networks and their stability properties have been studied extensively [11, 19]. Since the number of connections per neuron in these networks is large, they behave qualitatively like the balanced networks discussed in the previous section.

15.3.1.4 Spatial distribution of activity in finite heterogeneous networks

Mean-field equations have been derived for heterogeneous networks of binary neurons [119] and for heterogeneous networks of noisy LIF neurons [10]. Consider a network of N neurons in which the probability that two neurons are connected is $C/N \ll 1$. Each neuron will receive C connections on average, and the cell-to-cell fluctuations in the number of afferents will be order $\sqrt{C}$. In principle, when C is large, the fluctuations in the number of connections are small compared to the mean. However, since networks of excitatory and inhibitory cells settle down in a balanced state in which excitation and inhibition cancel each other out to within $1/\sqrt{C}$ (see above), the effective average input to the neurons becomes of the same order as its fluctuations, and this is reflected in wide distributions of firing rates in the steady states. To calculate this distributions self-consistently one proceeds as follows: The

temporal average currents can be written as

$$\mu_E = J_{EE}\sum_{j=1}^{N_E} c_j \nu_j^E - J_{EI}\sum_{j=1}^{N_I} c_j \nu_j^I$$
$$\mu_I = J_{IE}\sum_{j=1}^{N_E} c_j \nu_j^E - J_{II}\sum_{j=1}^{N_I} c_j \nu_j^I, \tag{15.74}$$

where c_j is a binary random variable such that $\text{Prob}(c_j = 1) = C_{E,I}/N_{E,I} \equiv \varepsilon$, and where we have assumed for simplicity that the excitatory (inhibitory) synaptic efficacies are uniform for each type of connections (J_{EE}, J_{EI}, J_{IE}, J_{II}). The temporal averages of the current $\mu_{E,I}$ are now random variables due to the randomness in the connectivity. Their *spatial* averages are equal to

$$\bar{\mu_E} = \varepsilon[J_{EE}N_E\bar{\nu}_E - J_{EI}N_I\bar{\nu}_I]$$
$$\bar{\mu_I} = \varepsilon[J_{IE}N_E\bar{\nu}_E - J_{II}N_I\bar{\nu}_I], \tag{15.75}$$

where $\bar{\nu}_{E,I}$ are the average excitatory and inhibitory rates across the population. The variance of $\mu_{E,I}$ across the population is

$$\sigma^2_{\mu_E} = \varepsilon\left[(1-\varepsilon)\left(J^2_{EE}N_E\bar{\nu}^2_E + J^2_{EI}N_I\bar{\nu}^2_I\right) + J^2_{EE}N_E\sigma^2_{\nu_E} + J^2_{EI}N_I\sigma^2_{\nu_I}\right]$$
$$\sigma^2_{\mu_I} = \varepsilon\left[(1-\varepsilon)\left(J^2_{IE}N_E\bar{\nu}^2_E + J^2_{II}N_I\bar{\nu}^2_I\right) + J^2_{IE}N_E\sigma^2_{\nu_E} + J^2_{II}N_I\sigma^2_{\nu_I}\right], \tag{15.76}$$

with $\sigma^2_{\nu_{E,I}}$ equal to the variance of the spatial distribution of rates across the network. Since $\mu_{E,I}$ are the sum of many independent contributions, their distribution will be approximately Gaussian, so we can write

$$\mu_{E,I}(z) = \bar{\mu}_{E,I} + \sigma_{\mu_{E,I}}\, z, \tag{15.77}$$

where $z = \mathbf{N}(0,1)$. Rigorously speaking, the randomness in the connectivity will also induce cell-to-cell variability in the temporal fluctuations in the synaptic current. However, this effect is weak compared to the effect on the mean, so one can neglect it and still get very accurate results [10]. We will therefore assume that they are constant across the population, and close the self-consistency loop by writing the rates as a function of the mean and variance of the synaptic current

$$\nu_{E,I}(z) = \phi(\mu_{E,I}(z), \sigma_{E,I}). \tag{15.78}$$

To estimate the spatial distribution of rates across the network $\rho_{E,I}(\nu)$, we thus write

$$\rho_{E,I}(\nu) = \int \rho(z)\rho_{E,I}(\nu|z)dz = \int \rho(z)\delta(\nu - \phi(\mu_{E,I}(z), \sigma_{E,I}))dz. \tag{15.79}$$

The firing rate distributions obtained in this way agree very well with the results from numerical simulations [10] and are also qualitatively similar to the wide distributions of firing rates seen in cortex [70].

15.3.2 Stability and dynamics

Are the states in which rates are solutions of equations (15.63), (15.68) or (15.71) and (15.72) stable? In order to answer rigorously this question, one must come back to the Fokker-Planck approach of Section 15.2.3, and write down the corresponding equation for the distribution of membrane potentials of neurons in the network, $\rho(V,t|V_0,t_0)$, coupled to the average firing rate $\nu(t)$ through the boundary conditions [1, 24]. A brief sketch of this approach is provided in the Appendix. Although this is the rigorous way to assess the stability of the steady-state solutions within the strong noise framework, analytical investigations of the Fokker-Planck equations are rather involved, and their numerical resolutions are also complicated [88]. Furthermore, their generalization to noisy situations with realistic synaptic dynamics is even more involved [21, 40]. Thus, it is of interest to investigate the possibility of approximating the dynamics by simpler dynamical equations, such as the Wilson-Cowan-type equations [126]. It is important to emphasize, however, that each dynamical description is suitable only for certain types of instabilities. For instance, an approximate dynamics in terms of firing rates cannot predict the instabilities of the network to a state where neurons are synchronized 'spike-to-spike'. Once a particular dynamical description is selected, the stability of the steady states against perturbations which comply with the assumptions of the chosen dynamical picture can be assessed.

Approximate firing rate dynamics have been found in some situations. For example, in weakly coupled networks with long synaptic time constants, one can derive an equation for the synaptic gating variable $s(t)$ [36, 37]. Let us write Equation (15.62) as $\nu = \phi(\mu_V, \sigma_V)$, where $\mu_V = V_{ss} - V_L = \mu_V^{ext} + C\bar{J}s(t)$. We consider a single population, fully connected network, with $\sigma_V = \sigma_C^{ext}\sqrt{\tau_m}/C_m = 5$ mV. The dynamics for $s(t)$ reads

$$\begin{aligned} \frac{ds(t)}{dt} &= -\frac{s(t)}{\tau_{syn}} + \nu(t) \\ \nu(t) &= \phi(\mu_V^{ext} + C\bar{J}s(t), \sigma_V), \end{aligned} \tag{15.80}$$

where $\mu_V^{ext} = \bar{J}_{ext}C_{ext}s_{ext}$ is the contribution of the external input to the steady state voltage. According to this scheme, the firing rate is always at its steady state value given the input, whereas the synaptic gating variable only follows the rate with its characteristic time constant τ_{syn}. This approximation is justified for the LIF model, since it has been shown that the population firing rate follows the synaptic input instantaneously [21, 40], provided that there is sufficient input noise, and that the synaptic time constant is comparable to, or longer than, the effective membrane time constant τ_m^{eff}. To what extent this approximation holds true for the Hodgkin-Huxley neuron model, or for real neurons, remains to be established. Qualitatively, the reason is that when there is enough input noise, there is always a significant fraction of the neurons across the population close enough to threshold and 'ready' to respond immediately to a change in current. Thus, it is appropriate to use the steady-state relationship $\nu = \phi(\mu_V, \sigma_V)$ even if the inputs are not stationary, e.g., $\nu(t) = \phi(\mu_V(t), \sigma_V) = \phi(\mu_V^{ext} + C\bar{J}s(t), \sigma_V)$, as in Equation (15.80).

The fixed point of this system is

$$s_{ss} = \tau_{syn}\phi(\mu_V^{ext} + C\bar{J}s_{ss}, \sigma_V). \tag{15.81}$$

To check the stability of the fixed points of this network, the standard procedure is to consider a small perturbation of a steady state

$$s = s_{ss} + \delta s \exp(\lambda t), \tag{15.82}$$

where δs is a small perturbation that grows at a rate λ. Stability of the steady state implies $Re(\lambda) < 0$ for all possible perturbations. Inserting Equation (15.82) in Equation (15.80), we get

$$\lambda = -\frac{1}{\tau_{syn}} + \left.\frac{d\phi(\mu_V(s), \sigma_V)}{ds}\right|_{s=s_{ss}}. \tag{15.83}$$

Thus, the stability condition $\lambda < 0$ is

$$\left.\frac{d\phi(\mu_V(s), \sigma_V)}{ds}\right|_{s=s_{ss}} < \frac{1}{\tau_{syn}}. \tag{15.84}$$

Equation (15.84) is a condition on the slope of the input-output function ϕ at the value of the input current given by s_{ss}. Since it is more intuitive to work with firing rates, we can express it as a condition on the slope of ϕ as a funcion of ν if we note that we only need the value of this slope at the steady state. In general

$$\frac{d\phi}{d\nu} = \left(\frac{d\phi}{ds}\right)\left(\frac{ds}{d\nu}\right). \tag{15.85}$$

Since the output rate is an instantaneous function of s, we can calculate the first term on the right hand side for all $s(t)$. The second term we do not know in principle, but on the steady state $s_{ss} = \tau_{syn}\nu_{ss}$. Thus

$$\left.\frac{d\phi}{ds}\right|_{s=s_{ss}} = \frac{1}{\tau_{syn}} \left.\frac{d\phi}{d\nu}\right|_{\nu=\nu_{ss}}, \tag{15.86}$$

so that the stability condition becomes

$$\left.\frac{d\phi(\mu_V(\nu), \sigma_V)}{d\nu}\right|_{\nu=\nu_{ss}} < 1. \tag{15.87}$$

Thus, for a fixed point to be stable, the slope of the output rate as a function of the input rate, should be less than one at the fixed point. This is shown graphically in Figure 15.4, where $\phi(\mu_V(\nu), \sigma_V)$ is plotted as a function of ν, both for an excitatory network ($\bar{J} = 0.5$ mV) and an inhibitory network ($\bar{J} = -0.5$ mV). The external inputs are adjusted so that there is an intersection with the diagonal at 1 Hz in both cases. When the network is excitatory (Figure 15.4A; in this case we use $\mu_V^{ext} = 0$ mV), the function $\phi(\mu_V(\nu), \sigma_V)$ raises very fast from zero rates to saturation. Thus, the slope

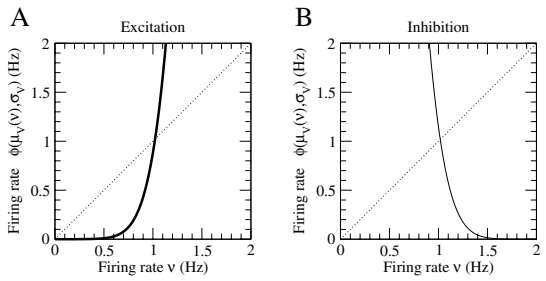

Figure 15.4

Self-consistent solution of Equation (15.80) and its stability properties. The output firing rate $\phi(\mu_V(\nu), \sigma_V)$ is plotted versus ν in two situations. **(A)** Excitatory network with $\mu_V^{ext} = 0$ mV, $\bar{J} = 0.5$ mV. **(B)** Inhibitory network with $\mu_V^{ext} = 20.4$ mV, $\bar{J} = -0.5$ mV. Other parameters are: $C = 1000$, $\tau_m = \tau_{syn} = 20$ ms ($C_m = 0.5$ nF, $g_L = 25$ nS), $\tau_{ref} = 2$ ms, $V_{th} = -50$ mV, $V_r = -60$ mV, $\sigma_V = 5$ mV. For both types of networks, there is a self-consistent solution around 1Hz. In the excitatory network, this self-consistent solution is highly unstable, because the slope of $\phi(\mu_V(\nu), \sigma_V)$ vs. ν is much larger than one; in the inhibitory network, the self-consistent solution is highly stable because of the large negative slope. In a balanced network where inhibition strongly dominates the recurrent circuit, the slope becomes infinite negative. Note that in the excitatory network, there are two other solutions: one at zero rate, and one close to saturation rates (about 500 Hz).

of $\phi(\mu_V(\nu), \sigma_V)$ is much larger than one at the self-consistent rate ν_{ss} and this steady state is, thus, unstable. The conclusion here is that low firing rates are expected to be hard to achieve in purely excitatory networks unless they are weakly coupled.

On the other hand, when the network is inhibitory (Figure 15.4B; $\mu_V^{ext} = 20.4$ mV), a supra-threshold external input is required to obtain an active network. The function ϕ now decreases as a function of ν (due to the fact that the mean decreases with ν). Equation (15.87) is now trivially satisfied when the coupling is predominantly inhibitory, $\bar{J} < 0$. Hence, a network state at this rate is stable. In the balanced network of Section 15.3.1.2, inhibition strongly dominates recurrence because it has to compensate for the external inputs. In this limit, the slope becomes infinitely negative. Note, however, that this strong stability of the purely inhibitory network is peculiar to synaptic couplings without latency. In presence of a latency, oscillatory instabilities appear even in strongly noisy networks [19, 24, 26].

Another simplified rate dynamics which has been frequently used is

$$\begin{aligned} \tau \frac{d\nu(t)}{dt} &= -\nu(t) + \phi(\mu_V^{ext} + C\bar{J}s(t), \sigma_V) \\ s(t) &= \tau_{syn}\nu(t), \end{aligned} \tag{15.88}$$

where τ remains unspecified. Although the fixed points of the systems described by Equations (15.80) and (15.88) are the same, this latter scheme neglects the dynamics for the synaptic variable, and instead uses an arbitrary time constant for the process by which the firing rate attains its steady state value. In conditions of high noise, Equations (15.80) seem, therefore, better suited to describe the time course of network activity than Equations (15.88).

We can extend Equations (15.80) to allow the description of a network with two (excitatory and inhibitory) neural populations, with firing rates ν_E and ν_I, and with synaptic latencies. If synaptic activation has a latency of τ_{lE} for excitation and τ_{lI} for inhibition, then we have

$$\begin{aligned} \frac{ds_E(t)}{dt} &= -s_E(t)/\tau_{syn,E} + \nu_E(t - \tau_{lE}) \\ \nu_E(t) &= \phi_E(\mu_V^{extE} + C_E\bar{J}_{EE}s_E(t) - \bar{J}_{EI}C_I s_I(t), \sigma_V^E) \\ \frac{ds_I(t)}{dt} &= -s_I(t)/\tau_{syn,I} + \nu_I(t - \tau_{lI}) \\ \nu_I(t) &= \phi_I(\mu_V^{extI} + C_E\bar{J}_{IE}s_E(t) - \bar{J}_{II}C_I s_I(t), \sigma_V^I). \end{aligned} \tag{15.89}$$

Equations (15.89) are Wilson-Cowan type dynamical mean-field equations which are 'derived' from an underlying biophysical description of neurons and synapses. In principle, the behavior of this model can be compared quantitatively (albeit approximately) with that of the original large-scale network of irregularly spiking LIF neurons. One should bear in mind, however, that when time delays are included, the dynamics become significantly richer, and the analysis more complicated. In fact, rigurously speaking, in the presence of temporal delays the system becomes infinite-dimensional, even if one deals with a single population (a function evaluated at $t + \tau$,

i.e., displaced in time, can be expressed as an infinite series of the time-derivatives of the function evaluated at t). Still, simplified descriptions of this complicated dynamical system can be used which produce results in good agreement with those from simulations [21, 26, 40]. In general, the Wilson-Cowan type equations do a good job of predicting the mean rate instabilities. They predict oscillatory instabilities only in the above mentioned conditions (large amplitude noise filtered by synapses with time constants comparable to membrane time constants, see [21, 26, 40]). For small noise, instabilities not predicted by the rate equations occur, in which neurons are synchronized 'spike-to-spike' (see e.g., [1, 62, 122] and refs therein). For other discussions about reductions to firing rate equations, see [7, 52, 109]. In a network in which a significant part of the fluctuations is generated by the network itself, Equations (15.80-15.88) can also be generalized to include the variance as a dynamical variable [11]. This approach gives rather accurately the mean rate instabilities in such networks, but not the oscillatory instabilities induced by the interplay between the dynamics of the variance and the mean, where the full Fokker-Planck approach must be used [19].

15.3.3 Bistability in a single population network

We now come back to the fully connected network of N_E excitatory cells. As we have seen in Section 15.3.2 (Figure 15.4), such networks can have more than one steady-state for the firing rates, provided the excitatory coupling is strong enough. Thus, an excitatory network can be *bistable*. In absence of external inputs, and with linear synapses, bistability typically occurs between one steady state at zero rate and another one at rates close to saturation. How is this picture affected by realistic synaptic dynamics? A situation of interest is when synaptic currents are mediated by saturating NMDA receptors [120]. By the arguments presented above, the fluctuations in these currents are negligible, because of the long decay time constant. In addition to N_E recurrent contacts, each neuron also receives C_{ext} AMPAR-mediated excitatory synaptic inputs from outside the network. Each of these external inputs provides spikes according to an independent Poisson process, and the firing rate of each input is random from a distribution with mean ν_{ext}. Thus, in this particular example, all the noise is generated outside the network, and is independent from cell to cell by construction.

We consider a slightly different dynamical picture from the one introduced in the previous section in Equation (15.80). The main difference is that, since we want to include the realistic synaptic dynamics, Equation (15.40), appropriate for the slow NMDA channel dynamics, the synaptic currents depend now non-linearly on the firing rates. Following the arguments presented in Section 15.3.2, we use the mean level of synaptic activity at the recurrent synapses $\bar{s}_{\mathrm{NMDA}}$ as the dynamical variable, since the time constant of the NMDA-mediated currents is the slower time scale of the system. However, the synaptic activity depends now in a non-linear way on the input rate. Equations (15.40) and (15.80) suggest seeking a dynamical equation for

the variable s of the form

$$\frac{d\bar{s}_{\text{NMDA}}}{dt} = -\frac{\bar{s}_{\text{NMDA}}}{\tau_{\text{NMDA}}} + (1 - \bar{s}_{\text{NMDA}})F(\nu), \tag{15.90}$$

where the function $F(\nu)$ is determined in a self-consistent way by the steady state dependency of $\bar{s}_{\text{NMDA}}$ on ν (note that we are going directly from the firing rate ν to the $\bar{s}_{\text{NMDA}}$ variable. Thus we are neglecting the dynamics associated to the variable $x(t)$ in Equation (15.41)).

Imposing $d\bar{s}_{\text{NMDA}}/dt = 0$ we obtain

$$\bar{s}^{ss}_{\text{NMDA}} = \frac{F(\nu)}{\frac{1}{\tau_{\text{NMDA}}} + F(\nu)} \equiv \psi(\nu), \text{ so that } F(\nu) = \frac{\psi(\nu)}{\tau_{\text{NMDA}}(1 - \psi(\nu))}. \tag{15.91}$$

Inserting this expression back into Equation (15.90), we obtain

$$\frac{d\bar{s}_{\text{NMDA}}}{dt} = -\frac{1}{\tau^{eff}_{\text{NMDA}}}\left[\bar{s}_{\text{NMDA}} - \psi(\nu)\right], \tag{15.92}$$

where $\tau^{eff}_{\text{NMDA}} = \tau_{\text{NMDA}}(1 - \psi(\nu))$, and where ν is given by

$$\nu = \left[\tau_{ref} + \tau^{eff}_m \sqrt{\pi} \int_{\frac{V_r - V_{ss}}{\sigma_V}}^{\frac{V_{th} - V_{ss}}{\sigma_V}} e^{x^2}(1 + \text{erf}(x))dx\right]^{-1}, \tag{15.93}$$

which depends on $\bar{s}_{\text{NMDA}}$ through τ^{eff}_m and V_{ss} (see Equations (15.61) and (15.59)). Thus, due to the saturation implicit in Equation (15.90), the effective time constant of this dynamics depends on the firing rate, and becomes faster at higher pre-synaptic activity. When the firing rates change slowly enough, this dynamics produces quantitative agreement with the results from simulations of the full spiking network [95].

Looking at these expressions, one notices that the dependence of the firing rate on the mean recurrent synaptic activity is always through the product $N_E \bar{g}_{\text{NMDA}} \bar{s}_{\text{NMDA}} \equiv g_{\text{tot}} \bar{s}_{\text{NMDA}} \equiv \tilde{s}$. The steady states of our dynamics are thus given by the solutions of the following equation

$$\frac{\tilde{s}}{g_{\text{tot}}} = \psi(\nu(\tilde{s})), \tag{15.94}$$

which correspond to the intersections of the curves given by each side of this equation plotted as a function of $\tilde{s}$. This equation generalizes Equation (15.63) to the situation of non-linear synapses. Qualitatively, these intersections correspond to the points in which the activity at the synapse (the right-hand side of Equation (15.94)) is equal to the feedback provided by the network (the left-hand side of the same equation), which is a necessary condition for the network to be at a steady state. The advantage of using $\tilde{s}$ as our variable, is that now the right-hand side of the self-consistency equation is no longer dependent on the total synaptic conductance g_{tot}, which measures the gain, or amplification, of the mean activity at a single synapse by the network. Thus, as g_{tot} is varied, the self-consistent solutions of the dynamics move like the intersections of the two curves in Equation (15.94) as the slope of

the straight line measuring the network feedback is changed. In Figure 15.5A, the function $\psi(\nu(\tilde{s}))$ and the line $\tilde{s}/g_{\rm tot}$ have been plotted for three values of $g_{\rm tot}$. This figure shows that, depending on the value of the gain, the two curves can intersect either once or three times, allowing for the possibility of several coexisting steady state solutions.

The next step is to look at the stability of these solutions. It can be done along the lines of Section 15.3.2. Let us, for brevity use $\bar{s} \equiv \bar{s}_{\rm NMDA}$. We rewrite the dynamical equation as

$$\frac{d\tilde{s}}{dt} = \frac{1}{\tau^{eff}_{\rm NMDA}}(g_{\rm tot}\psi(\nu(\tilde{s})) - \tilde{s}) \equiv G(\tilde{s}), \tag{15.95}$$

The stability of a steady-state $\tilde{s}_{ss}$ is given by the slope of $G(\tilde{s})$ evaluated at $\tilde{s}_{ss}$. If

$$\left.\frac{dG(\tilde{s})}{d\tilde{s}}\right|_{\tilde{s}=\tilde{s}_{ss}} < 0 \qquad \text{or} \qquad \left.\frac{d\psi(\nu(\tilde{s}))}{d\tilde{s}}\right|_{\tilde{s}=\tilde{s}_{ss}} < \frac{1}{g_{\rm tot}} \tag{15.96}$$

then $\tilde{s}_{ss}$ is stable. Recall that $G(\tilde{s_{ss}}) = 0$. If $G(\tilde{s})$ has a negative slope at $\tilde{s}_{ss}$, then it is positive for $\tilde{s}$ slightly less than $\tilde{s}_{ss}$, therefore $\tilde{s}$ will increase in time according to Equation (15.95), converging towards $\tilde{s}_{ss}$. Similarly, $G(\tilde{s})$ is negative for $\tilde{s}$ slightly larger than $\tilde{s}_{ss}$, again $\tilde{s}$ will converge back to $\tilde{s}_{ss}$. Therefore, $\tilde{s}_{ss}$ is stable. Conversely, if the derivative of $G(\tilde{s})$ is positive at a steady state, the latter is unstable.

Equation (15.96) implies that to assess the stability of a steady state solution graphically by looking at the intersections of the two functions in Equation (15.94), the stable fixed points will be those in which the sigmoid function has a lower slope than the straight line at the intersection.

Figure 15.5A shows that if the recurrent gain $g_{\rm tot}$ is lower than the dashed line marked by $g^{Low}_{\rm tot}$, the slope $1/g_{\rm tot}$ is too high and there is only one fixed point with low, but non-zero activity, which is always stable. On the other hand, when the recurrent gain is higher than the dashed line marked by $g^{High}_{\rm tot}$, the only fixed point, which is also always stable, corresponds to a state of high activity. In between, there is a range of values of $g_{\rm tot}$ in which three fixed points coexist. The ones with higher and lower activity are stable (marked with a filled circle), and the intermediate one (open circle) is unstable. When $g_{\rm tot}$ lies within this range, the network is said to be bistable. The intermediate unstable point corresponds to the steady-state which we showed in Figure 15.4 with excitatory connections.

As shown in Figure 15.5C, when the network is bistable, transient inputs can switch the state of the network between its two stable states. The network can, in this sense, be used as a working memory device (see next section), as the presence or absence of an input to the network can be read out from its activity *after* the stimulus is no longer physically present. One atractive feature of this encoding scheme is its robustness. Indeed, the activity state of the network does not reflect the occurrence of a single stimulus, but rather of a *class* of stimuli. In our example, quite a different range of amplitudes of the applied current will lead to the same steady-state. The network is said to use 'attractor dynamics', since each fixed point attracts the state of the network from a wide range of initial conditions. This concept can be understood by imagining that the state of the network, in our example the gating variable $\bar{s}$, slides

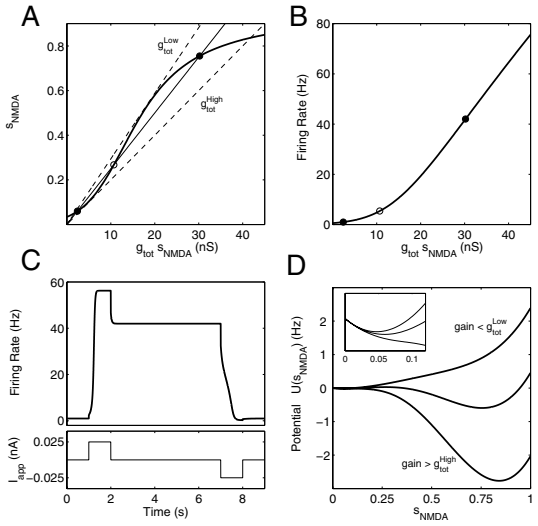

Figure 15.5

Bistability in a simple, one-population recurrent network. **(A)**. Mean synaptic activity as a function of the total recurrent synaptic input $\tilde{s} = g_{\rm tot} s_{\rm NMDA}$ (sigmoid; thick solid line) and same quantity when $\tilde{s}$ is interpreted as the total recurrent network feedback (thin dashed lines). $g_{\rm tot}^{High} = 50$ nS and $g_{\rm tot}^{Low} = 34$ nS correspond to the highest and lowest values of the network gain in which the network is bistable for this network parameters. The crossings of a straight line corresponding to an intermediate value of $g_{\rm tot} = 40$ nS with the sigmoid, correspond to the 3 steady state solutions in the bistable regime. Stable (unstable) solutions are marked with a filled (open) circle. **(B)**. Mean firing rate of the neurons as a function of the total recurrent synaptic input. The fixed point solutions in **A** are shown. Note the low rate (~ 40 Hz) of the stable high activity state, and the non-zero rate (~ 1 Hz) of the stable low activity state. Network parameters are: $C_{ext}\nu_{ext} = 0.7$ KHz, $g_{\rm AMPA}^{ext} = 3.1$ nS. Parameters for the single cells are like in Figure 15.1. Synaptic parameters are given in Section 15.2.6.2. **(C)**. Time course of activity (top) in the network when $g_{\rm tot} = 40$ nS. Brief current pulses (bottom) switch the network between its two stable states. **(D)**. Potential function associated to the dynamics of Equation (15.92) for three different values of $g_{\rm tot}$, corresponding to the situations where there is only one low activity steady-state (upper curve), one high activity steady-state (lower curve) and for a bistable network (middle curve). The values of $s_{\rm NMDA}$ at the steady states are given by the minima of the potential function. **Inset**: blown-up version of the low activity region of U shows the disappearance of the low activity minimum of the potential as $g_{\rm tot}$ increases.

down a hilly landscape. The valleys (or minima) correspond to the steady-states, and the class of stimuli which are attracted to each steady-state (called its basin of attraction) are the set of all locations in the landscape which roll down to the same minima. Indeed, the dynamics (15.92) can be re-written as

$$\frac{d\bar{s}}{dt} = -\frac{dU(\bar{s})}{d\bar{s}}. \tag{15.97}$$

This dynamics describes the movement of a point particle at location $\bar{s}$ sliding down the landscape defined by the function $U(\bar{s})$ in the presence of high friction. $U(\bar{s})$ is also called the potential of the dynamics, and it is such that the speed of the particle at location $\bar{s}$ is equal to minus the slope of U at that location. In Figure 15.5D we show three examples of the $U(\bar{s})$ for values of the gain at the recurrent connections g_{tot} such that the network has either a single high ($g_{\text{tot}} > g_{\text{tot}}^{High}$) or low ($g_{\text{tot}} < g_{\text{tot}}^{Low}$) activity steady-state, and for an intermediate value of g_{tot}, where the network is bistable. For low enough gain, $U(\bar{s})$ has a single minimum, and as the gain increases, a second minimum at a higher value of $\bar{s}$ appears. These two minima coexist for a range of values of the gain, but if the gain is high enough, the low activity minimum disappears (see inset in Figure 15.5D).

Several features deserve comments: First, in contrast to the network of linear synapses of Section 15.3.2, the firing rate in the high activity fixed point is about 40 Hz (Figure 15.5B), much less than saturation rates, even in the absence of inhibition. This rate is in the upper range of the available physiological data for persistent activity(20-50 Hz). This relatively low rate is due to the saturation properties of the NMDA receptor. Second, the low activity state has a low firing rate of about 1Hz. This is due to the presence of noise in the system. In the absence of noise, i.e., when the synaptic current is constant in time, the input-output function of the neuron becomes a sigmoid with a 'hard' threshold. For currents below this threshold the output rate is identically zero (see trend for decreasing noise levels in Figure 15.2) and in the supra-threshold regime it increases as a sub-linear function of the input current until saturation at $1/\tau_{ref}$ is reached. In these conditions, when the network is bistable, the low activity state is necessarily zero. When noise is included in the description, the firing rate can be non-zero even in the sub-threshold regime: the membrane potential, which hovers around its steady state below threshold, crosses this threshold once in a while as a result of the random fluctuations in the input current [8, 118]. Such a state of low, fluctuation-driven activity has been suggested to correspond to the background or spontaneous activity state found in the cortex [11]. The fact that the low activity state is stable in Figure 15.5 is due to the fact that the excitatory feedback is unrealistically weak (see caption of Figure 15.5). When excitatory feedback is stronger, the non-zero low rate state becomes unstable, and inhibition becomes necessary to achieve stability at low rates [11]. Hence, a more detailed model is required.

15.3.4 Persistent neural activity in an object working memory model

As we have just demonstrated, the self-consistency equations whose solutions provide the firing rates of the different neural populations in the steady-states of the recurrent network can, in some cases, have multiple solutions for the same set of parameters and external inputs to the network. When this is the case, transient external inputs can switch the state of the network among its possible stationary solutions. Conceptually, the network can now function as a short-term or working memory system, as its state of activity is no longer uniquely specified by the 'static' variables of the system (cellular or network parameters, unspecific external inputs, etc) but also carries information about the recent history of transient inputs to the network. More generally, a recurrent network can display multi-stability, whereby a resting state of spontaneous activity coexists with multiple attractor states (stable neural firing patterns), each of which encodes a different sensory stimulus. Therefore, the identity of a transient input is encoded and stored in the level of spiking activity of a distinct neural assembly in the recurrent circuit. Such stimulus-selective persistent activity has been documented during electrophysiological experiments on behaving monkeys during working memory tasks [35, 41, 42, 43, 44, 45, 46, 47, 48, 55, 56, 73, 79, 81, 85, 86, 89, 100, 124].

We now describe in some detail an object working memory model that has been analyzed both at the mean-field level and with numerical simulations. For the model to comply with the basic phenomenology of the data from object working memory experiments, the simple bistable network presented in the previous section has to be considerably enlarged. First, local networks in association cortices are likely to be endowed with much more than two attractors. The experiments in the temporal lobe with a large number of stimuli (up to 100) [81, 85, 86, 100] suggest the following picture:

- In the absence of external stimulation, networks in the temporal lobe are in a spontaneous activity state, in which all neurons fire at low levels of several Hz;
- Upon presentation of a particular *familiar* stimulus, a small sub-population of neurons in localized areas of the temporal lobe exhibit persistent activity; this fraction of neurons can be estimated to be around 1% or a few % [81]. Thus, the representation of *familiar* stimuli in these areas is sparse;
- Representations of different stimuli have very small overlaps, since neurons typically respond to only one or a few images in the set of shown images [81].

A model for object working memory based on these observations has been built in several stages [9, 11, 12, 21]. The model of [11] is a network of randomly connected excitatory and inhibitory neurons much as the one discussed in Section 15.3.1.3. In addition, the excitatory population is divided in sub-populations, where a given sub-population is assumed to have a strong visual response to a particular stimulus.

A schematic representation of the architecture of the model is shown in Figure 15.6A. The network consists of two large pools of interacting pyramidal cells and interneurons. Both populations are fully connected with themselves and with

each other. The pyramidal cell population is itself divided in several sub-populations. Since the experiments show that single cells only respond to a very small fraction of the stimulus set (approximately 1%), the model assumes that each sub-population shows selective responses to a single stimulus, and that the different sub-populations are non-overlapping. In addition to this set of 'selective' sub-populations, there is a sub-population of neurons which are not selective to any particular object. All neurons also receive unspecific excitation from outside the network. In the model of [11], synaptic transmission was assumed to be instantaneous. [25] proposed a model with more biologically plausible synapses (a full description of the spiking neuron model, as well as the complete set of mean-field equations, can be found in that paper). In the model, excitatory transmission is both AMPAR- and NMDAR-mediated, though with a dominant contribution of the NMDA component at the recurrent synapses and a dominant AMPA component on the external inputs, while inhibitory transmission is mediated by $GABA_A$ receptors. Neurons belonging to the same selective sub-population are assumed to be frequently co-activated by the visual input which drives them effectively, and, therefore, the excitatory synapses connecting them are assumed to have undergone Hebbian synaptic potentiation, so that the synaptic strength at these recurrent connections is supposed to be larger than average.

As shown in Figure 15.6B, when the strength of these synapses is increased to be approximately twice the average (baseline) excitatory coupling strength, there is a sudden 'bifurcation' at which bistability emerges in each selective sub-population. Therefore, a graded difference in the coupling strength could lead to qualitatively different network behaviors (e.g., with or without persistent activity). Again the firing rate in the elevated persistent activity state is fairly low, in agreement with the data. Also, although the firing rates predicted by the mean-field model (solid curve) slightly overestimate the results from the direct simulations of the original spiking neural network (filled dots), the agreement between the two is reasonably good.

An example of the behavior of the network when the sub-populations are bistable is shown in Figure 15.6C. The upper part of the figure shows rastergrams from selected neurons from each of the sub-populations, and the lower part shows the population firing rate for each of the sub-populations. During the time interval marked as sample, the firing rate of the external inputs to the sub-population marked as number 1 is transiently elevated. The resulting increase in activity in this sub-population persists during a delay of several seconds, self-sustained by recurrent synaptic reverberations. At the end of the delay, an excitatory input to the whole network, signaling the behavioral response [29], switches the network back to the resting state. The persistent delay activity slightly increases the excitatory drive to the interneuron population, which also increases its activity. As a result, other sub-populations not selected by the transient input are more strongly inhibited during the delay.

15.3.5 Stability of the persistent activity state

The stability of the persistent state in a single population network is easily read out from a graph such as the one of Figure 15.5. In networks with several populations,

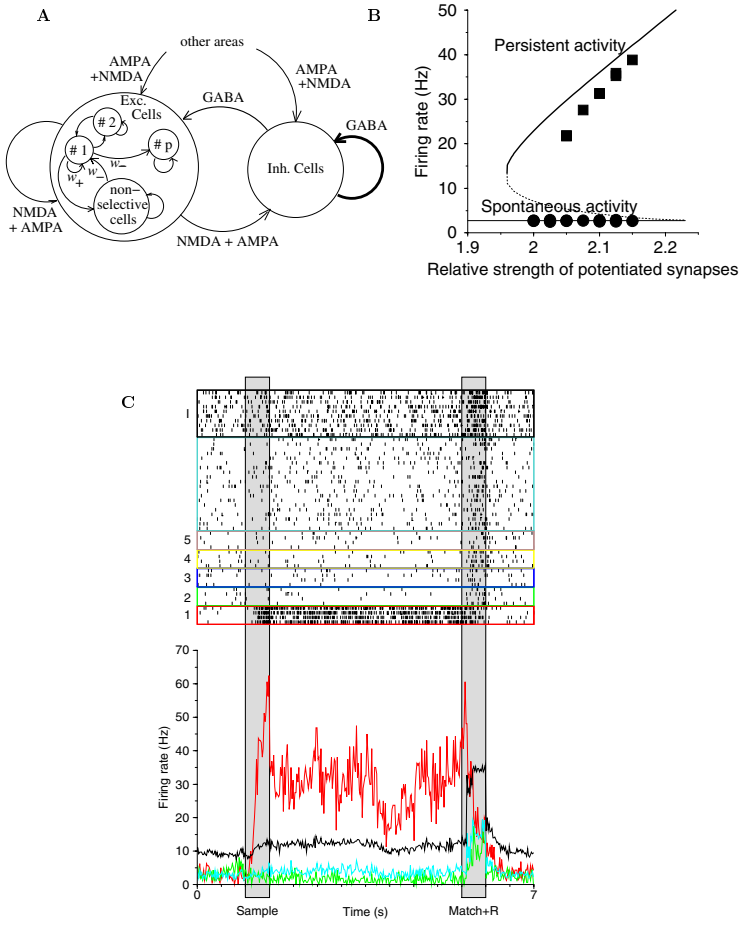

Figure 15.6

Behavior of an object working memory network. **(A)**. Schematic representation of the network. Circles represent the different sub-populations. Labels on the arrows indicate the type of synaptic connection between them. The width of the arrows qualitatively represents the strength of the corresponding synaptic connections. **(B)**. Bifurcation diagram showing the onset of bistability as a function of the strength of the connections within a selective sub-population relative to a baseline. Lines are the prediction from the mean-field version of the model, with solid (dashed) lines representing stable (unstable) steady states. Squares (spontaneous activity) and circles (elevated persistent activity state) are results from simulations of the spiking network (continued).

Figure 15.6

(C). Time course of the activity of the different sub-populations of the network in a delayed match-to-sample protocol. The network is initially in a resting state with low and uniform spontaneous firing activity. A brief stimulus to one of the neural sub-populations (indicated in red) triggers persistent activity that is self-sustained by recurrent reverberations and that is confined to that neural sub-population. This memory is erased, and the network is turned off, by another transient input during the match+response time epoch. Top: rastergrams. Bottom: firing rate histogram across the corresponding sub-populations. Red: selective sub-population receiving a transient external excitation during the sample period. Green, yellow, blue and brown: sub-populations selective to other stimuli. Cyan: sub-population of non-selective excitatory cells. Black: inhibitory interneurons. (See color insert.)

such as the excitatory-inhibitory networks, the analysis becomes more complicated, and the stationary state with the highest firing rate can destabilize through an oscillatory instability. Oscillatory synchrony can then disrupt bistability, or multi-stability. Conditions for stability of a persistent activity state in presence of synchrony can be understood from the simple following intuitive argument. Consider a network of neurons which, upon the arrival of a transient excitatory input, switch their activity from a few Hz in spontaneous activity to an elevated activity state of 20-40 Hz. Let us consider a single neuron firing at 25 Hz in this persistent activity state. This neuron fires, on average, every 40 ms. Since the tonic external input has not changed, in order for the neuron to maintain this firing rate, the recurrent network must provide enough current during the next few tens of milliseconds after each spike so that the cell will spike again. How can this be achieved? A possibility is that the network operates in an asynchronous state, where the assumed statistical independence of the firing times in different neurons is satisfied. In such a state, the fraction of neurons firing a spike across the network is, on average, constant. This property implies that, if the number of cells is large enough, the network will generate a tonic input, constant in time on average, which can sustain the firing of the single cells in a stable manner.

On the other hand, when there is some degree of synchrony, the fraction of cells firing at any given time starts to fluctuate on average, even for a very large network size. Consider the extreme case in which, at some instant of time, the whole network is perfectly synchronized, i.e., all neurons fire at the same time. Unless some mechanism exists which can keep a memory of this burst of activity and somehow delay it, turning it into an input to the cells at a *later* time, the activity of the network would decay to the resting spontaneous state [60]. A mechanism working in that direction could be implemented by the long time constant of NMDAR-mediated synaptic transmission [120]. Intuitively, according to Equation (15.29), if a synapse has time constant τ_{syn} the current into the cell resulting from a single spike at t_{spk} is still 37% of its maximum at $t_{spk} + \tau_{syn}$. This delay between the occurrence of an excitatory event, and its effect on the post-synaptic cell could, therefore, be beneficial for the stability of persistent activity in the presence of synchrony.

In addition, a long synaptic time constant for excitation might help to stabilize the asynchronous state itself. Oscillations easily occur in networks of interacting excitatory and inhibitory neurons, if the time constant of inhibition is longer than that of excitation (see e.g., [26, 120]). The intuitive reason is that, if such a network is perturbed from its steady state, the excitation will build up before the inhibition has time to suppress it. This excess of excitation will result in an increased inhibition which eventually overcomes the excitation, resulting in an overall suppresion of the excitation in the network. A decreased drive to the inhibitory cells leads to a decay of their activity, releasing the excitatory population and the rhythmic cycle starts again. When the excitation is slower, this type of oscillatory instability is prevented, as any excitatory perturbation results in an increased inhibition before the excitation has time to build up.

An example of the effect of changing the effective time constant of excitation on the stability of persistent activity is shown in Figure 15.7, for the object-working memory network described above. Remember that in this network, $GABA_A$R- mediated inhibition has a time constant of 10 ms, and AMPAR- and NMDAR-mediated excitation have time constants of 2 ms and 100 ms respectively. In the Figure, the temporal course of the average activity of a sub-population of selective cells after the application of a transient excitatory input is shown, as the relative contributions of AMPA and NMDA receptors at each excitatory synapse is varied systematically, thus taking the network from a scenario in which excitation is slower than inhibition to one in which it is faster. When the AMPA:NMDA ratio of the charge entry per unitary EPSC is 0.1 (measured at $V = -55$ mV, near threshold), the average activity is fairly constant in time and shows only small amplitude fluctuations which do not destabilize the persistent state. As the ratio is progressively increased, excitation becomes faster, and the amplitude of the fluctuations grows. However, due to the delay effect mentioned above, the state of persistent elevated activity is still stable for fairly large amplitude fluctuations in the average activity (see Figure 15.7c). Indeed, in these conditions, the power spectrum of the average activity shows a clear peak near 40 Hz (see inset in Figure 15.7e). Although the issue of whether the persistent activity observed in the cortex is indeed oscillatory is controversial, a similar spectral structure has been recorded in local field potentials of area LIP [90] (but see [30]). As expected, when excitation becomes too fast, the amplitude of the oscillations becomes too large, and NMDAR-mediated excitation is unable to bridge the gap between activity bursts, with the resulting destabilization of the persistent activity state, and the network's working memory behavior is lost (Figure 15.7D).

These arguments raise the possibility that NMDAR-mediated excitation, or more generally, slow synaptic or cellular recurrent excitation could help to prevent oscillatory instability resulting from the excitation-inhibition loop, thereby contributing to the stability of persistent activity generated in recurrent cortical networks. Other factors can of course affect this stability. For instance, mutual inhibition between interneurons can in some conditions reduce the propensity of instability in an excitation-inhibition loop [62, 116]. This effect can be understood using a two-population rate model like Equations (15.89). When there is no synaptic latency, it can be shown by the linear stability analysis of a steady state that the I-to-I coupling

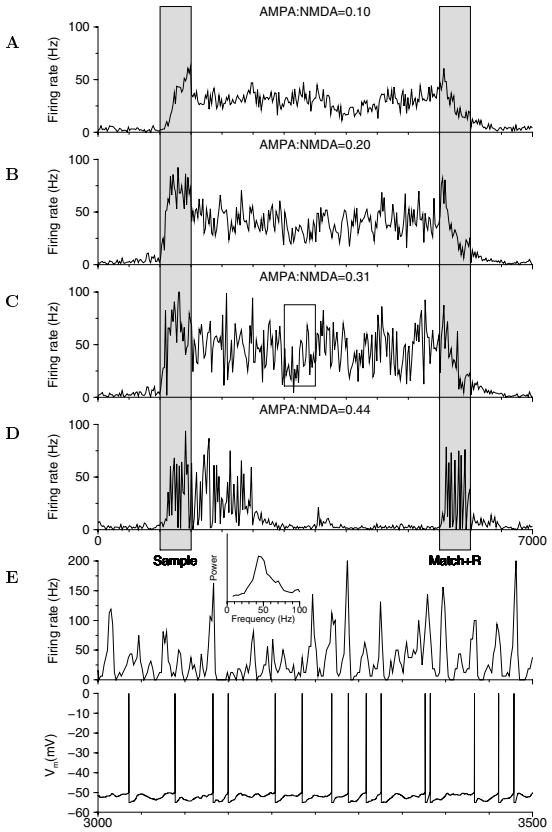

Figure 15.7

Stability of pesistent activity as a function of the AMPA:NMDA ratio. **(A-D)**. Temporal course of the average firing rate across a sub-population of selective cells in the network of Figure 15.6 after transient excitatory input, for different levels of the AMPA:NMDA ratio. This ratio is defined as that of charge entry through a unitary post-synaptic current at $V = -55$ mV (near threshold). As the ratio is increased, oscillations of a progressively larger amplitude develop, which eventually destabilize the persistent activity state. **(E)**. Snapshot of the activity of the network in **C** between 3 and 3.5 seconds. Top: Average network activity. Bottom: Intracellular voltage trace of a single neuron. **Inset**. Power spectrum of the average activity of the network, showing a peak in the gamma (~ 40 Hz) frequency range. Persistent activity is stable even in the presence of synchronous oscillations.

effectively reduces the time constant of the inhibitory population dynamics; faster inhibition thus reduces the likelihood for this type of oscillatory instability [112, 116]. However, it is important to emphasize that the oscillatory instability with reduced NMDA:AMPA ratio in our working memory models (Figure 6 and Figure 6 in [29]) was observed in the presence of strong I-to-I synaptic connections. The same result was also obtained with a spatial working memory model of Hogkin-Huxley-type conductance-based neurons [112]. Note that the models of [116] and [62] did not include synaptic latency, which has been shown to favor fast synchronous oscillations in a purely inhibitory network [21, 19, 26]. In general, stronger NMDA:AMPA ratio promotes asynchrony. Stronger I-to-I coupling without latency contributes to network asynchrony, but with synaptic/cellular latency could lead to oscillatory instability.

Other factors which might also contribute to the stability of persistent activity states include intrinsic ionic currents with long time constants [112], or bistability at the single cell level [28, 75]. Finally, heterogeneities, both in cellular and in connectivity properties, and noise, tend to desynchronize the network.

15.3.6 Multistability in balanced networks

Can bistability, or multi-stability, occur in a balanced network? Furthermore, can multistability occur between several states in which all neurons fire in a Poissonian fashion? These questions are interesting from a theoretical point of view, but are also raised directly by available data suggesting that the irregularity in the output spiking activity of cortical neurons is as high in high-rate persistent activity states as in the low-rate spontaneous activity state [30].

The balanced model of Section 15.3.1.2 seems incompatible with multistability, since the rates depend linearly on the external inputs through Equation (15.66). Indeed, unless the matrix of gain coefficients in equations (15.66) is singular (which requires a biologically unrealistic fine-tuning of parameters) these equations have a single solution for a fixed external input and are, consequently, incompatible with bistability. This is the manifestation of a general problem: any non-linear behavior in balanced networks requires a significant amount of fine-tuning in network parameters.

Several partial solutions to this problem can be suggested. First, in networks with finite connectivity such as the one of Section 15.3.4, multistability can be found even though the background state of the network is qualitatively a 'balanced state' (strong coupling, irregularly firing neurons). In fact, the multistability properties of such networks can be understood in the limit $C \to \infty$, $J \sim 1/\sqrt{C}$, if the relative size f of the selective sub-populations are taken to be small compared to the rest of the network, $f \sim 1/\sqrt{C}$: in this case, the leading order to the mean input in both cue and non-cue populations vanishes due to the 'balance' condition, but the corrections to the leading order are finite and different between the cue and non-cue populations, leading to a non-linear equation for the rates in the cue population (see [20] for a discussion of the sparse coding limit $f \to 0$). Thus, multistability is relatively easily achieved. The intuitive picture is that, while the global activity of the network is set

by the 'balance' condition set by the strong global inhibition, the activity in the small selective excitatory sub-population becomes essentially uncoupled from the rest of the network. Consequently, this sub-population behaves essentially as the weakly coupled excitatory network of Section 15.3.3. However, a direct consequence of this scenario is that the CV in persistent activity must be lower than in spontaneous activity, because the mean inputs are larger in the cue population, while the variance remains unchanged (see Figure 2). Thus, there is no multistability between several 'balanced states'.

It is therefore possible for small sub-populations within a larger balanced network to be bistable in a robust way, but at the price that the small sub-populations themselves do not remain balanced in both steady-states. Is there an alternative? The idea would be to find a scenario in which the variance in the cue population also increases in a significant way from spontaneous to persistent activity, so that the increase in CV induced by the increase in variance counterbalances the decrease induced by the increase in the mean. One can even imagine a scenario in which the mean does not change, but the variance does. Such a scenario was introduced in [93, 94] for a network with finite connectivity C. It is a generalization of the model of [11] (see Figure 15.6A) in which the interneurons are also subdivided in selective sub-populations. Such a network is divided functionally in 'columns' or 'micro-columns' composed both of excitatory and inhibitory populations. Both populations are activated in a selective way when their preferred stimulus is shown. Consequently, in a persistent state, both excitatory and inhibitory populations raise their firing rates. A similar phenomenon was observed in experiments monitoring the activity of neurons in the prefrontal cortex of primates during working memory tasks. Recordings of nearby putative pyramidal cells and interneurons showed that the two sub-populations increase their firing rate during the delay period [92]. This has lead to the postulate of a micro-columnar organization of the pre-frontal cortex [92].

The 'micro-columnar' network has been studied at the mean-field level [93, 94]. In order to do a systematic investigation of the spiking variability resulting from different types of network organizations, the mean-field theory has been extended to be self-consistent both at the level of rates and CVs. In the previously discussed models, Poisson spiking statistics was an assumption, so the irregularity in the spiking activity in the pre-synaptic spike trains was 'fixed'. In [93] this assumption was relaxed by assuming that the neuronal spike trains can be described as renewal processes characterized by their mean rate and CV (a renewal point process is characterized by independent ISI intervals from an arbitrary distribution). When the statistics of the renewal spike trains are close to the Poisson case, the output rate and CV of the post-synaptic neuron can be calculated as a function of the rate and CV of its inputs, leading to steady state solutions in which both the rate and the CV are calculated self-consistently. Using this framework, multistability in the micro-columnar network described above has been studied using simple heuristic firing rate dynamics similar to Equation (15.88). The synaptic interactions between neurons depend on whether they belong to the same or to different micro-columns, and again, selective micro-columns are characterized by stronger excitatory recurrent interactions. In Figure 15.8, we show the time course of activity of a bistable micro-column in

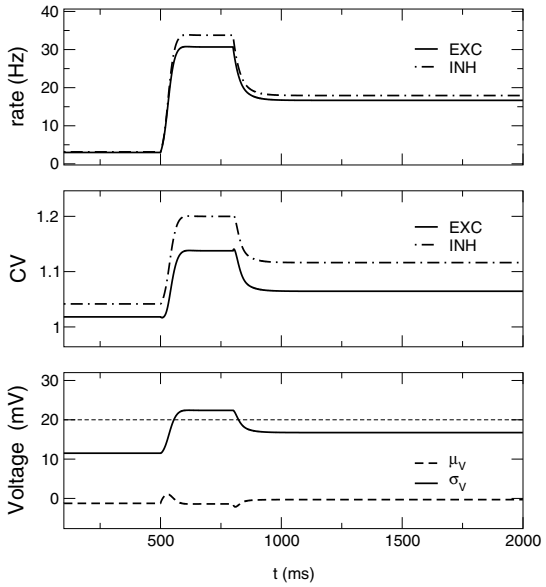

Figure 15.8

Bistability in a balanced multi-columnar cortical circuit. **(A)**. Temporal evolution of the firing rate from the excitatory and inhibitory sub-populations of a column. At $t = 500$ ms, a transient excitatory input was applied to both sub-populations. The elevated activity state in response to this input outlasts the stimulus offset. Note the elevated firing rate of the inhibitory sub-population also. **(B)**. Same as above for the CV of the two sub-populations. The CV increases with the firing rate. **(C)**. The figure shows the quantities $\mu_V = \mu_C/g_L$ and $\sigma_V = \sqrt{\tau_m}\sigma_C/C_m$ of the neurons in the excitatory sub-population. They correspond to the mean and standard deviation of the current, but are expressed in mV to facilitate comparison with the distance between V_{th} and V_L, equal to 20 mV (dashed line). The mean input current remains essentially the same for both the resting state and persistent state, regardless of their very different firing rates. The increase in firing rate in response to the stimulus is due to an increase in the amplitude of the fluctuating component of the current, hence the increase in CV above. In this network, both stable states are in the balanced regime.

response to a transient input. Panels A and B show the average rate and CV of the excitatory and inhibitory sub-populations in the micro-column. In this network, since the micro-column remains balanced in the elevated rate state, the CV of both sub-populations remains close to one in the delay period. Both the similar courses of activity of the excitatory and inhibitory populations, and the high CV during elevated persistent activity are consistent with measurements from prefrontal neurons in working memory tasks [30, 92]. The reason for this behavior is that the mean current to both sub-populations (see the lower panel for the case of the excitatory sub-population) remains approximately constant as the network switches between its two stable states. The increase in firing rate is due to an increase in the fluctuations in the current. Indeed, as a result of this, the CV actually increases in the elevated firing rate state.

This increase in CV is in contrast to the decrease in CV in models in which the network is not balanced in the elevated activity state, like the networks described in the previous two sections. This qualitative difference between relative change in spiking variability in these scenarios should, in principle, be experimentally testable, although the small difference in CV observed in Figure 15.8 would be hard to detect in experimentally recorded spike trains, due to limited sampling problems. Further experimental data are needed to resolve this issue.

As suggested at the beginning of this section, this scenario still suffers from a fine tuning problem. The range of multistability in the network with balanced persistent state is extremely small for realistic numbers of inputs per cell [93]. In fact, such multistability vanishes in the large C limit, even if the sub-populations are taken to scale as $1/\sqrt{C}$, because in that limit the difference in the fluctuations between spontaneous and persistent activity vanishes.

The fundamental problem which precludes robust bistability in balanced networks is the different scaling of the first two moments of the input current with the number of inputs and with the connection strength. While the mean scales as JC, the variance scales as J^2C. It is thus impossible to find a scaling relationship between J and C that keeps both moments finite when $C \to \infty$.

It is possible that cross-correlations in the activity of different neurons might provide a solution to the 'linearity' problem of balanced networks. The different scaling of the mean and the variance is a direct consequence of the fact that we have assumed the different inputs to the cell to be independent, so that the variance of their linear sum is the sum of their variances. If the inputs to the cell showed significant correlations, the variance would now scale as $(JC)^2$, in which case any scaling relationship between J and C would have the same effect on the mean and on the variance. It would therefore be of great interest to incorporate cross-correlations in a self-consistent manner into the picture we have been describing in this chapter.

15.4 Summary and future directions

In this chapter, we have presented analytical mean-field techniques that can be used to study the collective properties of large networks of spiking neurons. In analyzing the self-consistent steady-states of these networks, we observed that the self-consistency equations have sometimes multiple stable states. This leads quite naturally to the the interpretation of these networks as models of working memory systems. The methods discussed here help to understand in detail in which conditions multistability can be achieved in large networks of spiking neurons. The results that have been discussed are of course only the current status of a rapidly growing field. Extensions of both the mean-field techniques and of network architectures for working memory are either already done, under way, or should be done in the near future. We discuss here several of these possible extensions.

- **More realistic single neuron models.** The LIF lacks several features of real neurons. First, it lacks any sub-threshold resonance phenomena [65]. Generalizations of LIFs with several variables have been introduced that possess such sub-threshold resonance properties and can be studied analytically in stochastic contexts along the lines of Section 15.2.3 [98]. Second, it lacks an intrinsic spiking mechanism. The firing rate of neurons with intrinsic spike generation mechanism can be studied in the context of the 'quadratic integrate-and-fire' neuron [22], and even more realistic neurons can be studied analytically (Fourcaud et al. SFN 2002 abstract). Furthermore, mean-field theory can be extended to a recurrent network of Hodgkin-Huxley-type conductance-based single neurons [109]. This generalization may be important, e.g., the network stability may be different depending on whether single neurons are described by Hogkin-Huxley-type models or leaky integrate-and-fire models [26, 50].

- **More realistic synaptic dynamics.** The mean-field description of realistic synaptic interactions can be improved in at least two ways. First, synaptic fluctuations act through conductance changes, which are multiplied with the driving force $(V - E_{syn})$ to yield synaptic current. Therefore the noise is multiplicative. We have sidepassed this difficulty by replacing V with its average, so that the noise term becomes additive to the voltage equation. It would be desirable to be able to deal analytically with multiplicative noise. Second, synapses display short-term depression and facilitation [113, 131]. Mean-field models that incorporate synaptic depression have been investigated [115, 120], but the implications of short-term plasticity to recurrent networks, especially to working memory models, still await to be fully explored.

- **Extension to correlations between neurons.** In this chapter we have always assumed that the spiking activity of different cells was independent. Although the experimentally observed cross-correlations are relatively weak [14, 31, 74, 130], they might have a large impact on the input-output relation-

ship of a neuron, since when the correlation coefficient of the inputs to a cell is not zero, the fluctuations in its total afferent synaptic are proportional to the number of inputs to the neuron, instead of to its square root, as in the models we have described. Although the analytic treatment of cross-correlations is technically complicated, a systematic characterization of their effect on the rate and variability of simple spiking neuron models is becoming available [38, 82, 101, 102]. The real challenge is to extend the framework here presented in such a way to include cross-correlations in a self-consistent way. A first step in this direction has been taken by [78], where cross-correlation functions in a recurrent fully connected recurrent network of spike response neurons [53] have been calculated.

Finally, let us end with a note on recurrent networks that display a *continuum* of stable neural firing patterns. Some working memory systems are believed to encode features of sensory stimuli that are *analog quantities* (such as spatial location, direction, eye position, frequency, etc). Such systems have been hypothesized to be implemented in the brain by recurrent neural networks endowed with a continuous family of attractor states. Heuristic firing-rate models of this kind have been widely investigated [5, 16, 28, 63, 91, 103, 125, 128]. More recently, more biophysical continuous attractor models of spiking neurons have been developed for spatial working memory [29], parametric working memory [80] and short-term memory in the oculomotor system [104]. Theoreticians have also begun to analyze mean-field models that are derived from these spiking neural network models [80, 104, 109]. Further progress in this direction will considerably advance our theoretical understanding of recurrent cortical networks, and shed insights into the cellular and network mechanisms of working memory.

Acknowledgments. A.R. thanks Ruben Moreno for helpful discussions. A.R. and X-J.W. are supported by NIMH, the Alfred Sloan Foundation and the Swartz Foundation. N.B. is supported by the C.N.R.S.

Appendix 1: The diffusion approximation

We will follow the exposition by [97]. We consider the case of a single post-synaptic neuron which receives C_E excitatory and C_I inhibitory independent Poisson inputs of rates ν_E and ν_I respectively, each delivering a charge J_E and J_I per spike though an "instantaneous" synaptic current (see above). For this discussion, we measure voltages with respect to V_L, i.e., $V_L = 0$. We will also measure the effect of each pre-synaptic spike by the size of the resulting instantaneous jump in the membrane potential $\bar{J}_{E,I} = J_{E,I}/C_m$. Since the process is Markov, it satisfies

$$\rho(V,t+\Delta t|V_0,t_0) = \int_{-\infty}^{\infty} dV' \rho(V,t+\Delta t|V',t)\rho(V',t|V_0,t_0). \tag{15.98}$$

If Δt is sufficiently small, so that $\Delta t \ll \tau_m$, and so that the probability of receiving more than one spike in Δt is negligible, and since the pre-synaptic spikes produce discrete jumps, if follows that

$$\rho(V,t+\Delta t|V',t) = [1-(C_E\nu_E + C_I\nu_I)\Delta t]\delta(V'_0 - V) + \\ + C_E\nu_E\Delta t\,\delta(V'_1 - V) + C_I\nu_I\Delta t\,\delta(V'_2 - V), \qquad (15.99)$$

where the first, second and third terms correspond to the probabilities of receiving no spikes, an excitatory spike or an inhibitory spike in Δt respectively, and $V'_{0,1,2}$ are the values of the depolarization at $t+\Delta t$ in these three cases, given that the depolarization was V' at t. In order to calculate $V'_{0,1,2}$, we use the fact that, since Δt is small enough, the exponential time course of V in between spikes can be approximated by a linear decay. Thus

$$V'_0 = V'(1-\frac{\Delta t}{\tau_m}) \\ V'_1 = V'(1-\frac{\Delta t_1}{\tau_m}) + \bar{J}_E + [V'(1-\frac{\Delta t_1}{\tau_m}) + \bar{J}_E]\frac{\Delta t_2}{\tau_m} \\ V'_2 = V'(1-\frac{\Delta t_1}{\tau_m}) - \bar{J}_I + [V'(1-\frac{\Delta t_1}{\tau_m}) - \bar{J}_I]\frac{\Delta t_2}{\tau_m}, \qquad (15.100)$$

where $\Delta t_1 + \Delta t_2 = \Delta t$. Using the property $\delta(f(x)) = \delta(x-x')/|\partial_x f(x')|$ with x' such that $f(x') = 0$, and expanding to first order in $\Delta t/\tau_m$, equation (15.98) can be expressed as

$$\rho(V,t+\Delta t|V_0,t_0) = (1+\frac{\Delta t}{\tau_m})\Big[(1-(C_E\nu_E + C_I\nu_I)\Delta t)\,\rho(V(1+\frac{\Delta t}{\tau_m}),t|V_0,t_0) + \\ + C_E\nu_E\Delta t\,\rho([V-\bar{J}_E](1+\frac{\Delta t}{\tau_m}),t|V_0,t_0) + \\ + C_I\nu_I\Delta t\,\rho([V+\bar{J}_I](1+\frac{\Delta t}{\tau_m}),t|V_0,t_0)\Big], \qquad (15.101)$$

which, upon taking the limit $\Delta t \to 0$ becomes

$$\frac{\partial}{\partial t}\rho(V,t|V_0,t_0) = \frac{\partial}{\partial V}[(\frac{V}{\tau_m})\rho(V,t|V_0,t_0)] + C_E\nu_E[\rho(V-\bar{J}_E,t|V_0,t_0) \\ - \rho(V,t|V_0,t_0)] + +C_I\nu_I[\rho(V+\bar{J}_I,t|V_0,t_0) - \rho(V,t|V_0,t_0)]. \qquad (15.102)$$

Finally, expressing the terms in square brackets as a Taylor series expansion around V, one obtains

$$\frac{\partial}{\partial t}\rho(V,t|V_0,t_0) = \sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\,\frac{\partial^n}{\partial V^n}[A_n\,\rho(V,t|V_0,t_0)]. \qquad (15.103)$$

where

$$A_1(V) = -\frac{V}{\tau_m} + \bar{J}_E C_E\nu_E - \bar{J}_I C_I\nu_I \qquad (15.104)$$

$$A_n = \bar{J}_E^n C_E\nu_E + (-1)^n\bar{J}_I^n C_I\nu_I \qquad n = 2,3,\ldots \qquad (15.105)$$

are called the infinitesimal moments of the process. The intuitive nature of the diffusion approximation becomes now clear: the smaller $\bar{J}_{E,I}$, the fewer the terms needed to express $\rho(V \mp \bar{J}_{E,I}, t|V_0, t_0)$ as a Taylor series expansion arround V, and the fewer the terms one has to maintain in the infinite-order Equation (15.103) to give an accurate description of the process.

Appendix 2: Stability of the steady-state solutions for $\rho_{ss}(V)$

The function $\rho_{ss}(V)$ is the solution of the stationary Fokker-Planck Equation (15.23) with the appropriate boundary conditions. To assess the dynamical stability of this solution, one has to use the general Fokker-Planck Equation (15.9) to test the effect of small perturbations on the steady-state distribution. We briefly mention the logic of this procedure.

For the sake of simplicity, we only discuss here a simple situation in which synapses are instantaneous, with a latency τ_l ms after the pre-synaptic spike time. In this case, the dynamical counterparts to Equations (15.63, 15.68) are

$$\tau_m \frac{\partial \rho}{\partial t} = \frac{\sigma_{ext}^2}{2} \frac{\partial^2 \rho}{\partial V^2} - \frac{\partial}{\partial V} \left[(V - \mu_{ext} - \bar{J} \tau_m \nu(t - \tau_l)) \rho \right], \tag{15.106}$$

in the weakly coupled, fully connected case (no noise in recurrent inputs), and

$$\begin{aligned} \tau_m \frac{\partial \rho}{\partial t} = & \frac{\left(\sigma_{ext}^2 + C\bar{J}^2 \tau_m \nu(t - \delta)\right)}{2} \frac{\partial^2 \rho}{\partial V^2} \\ & - \frac{\partial}{\partial V} \left[(V - \mu_{ext} - C\bar{J} \tau_m \nu(t - \tau_l)) \rho \right], \end{aligned} \tag{15.107}$$

in the strongly coupled, sparsely connected case (noise in recurrent inputs). In both cases the boundary conditions are given by Equations (15.19,15.20).

The stationary solution to Equation (15.106) (resp. 15.107) is Equation (15.63) (resp. 15.68). To study their stability, a linear stability analysis must be performed. It consists in looking for solutions to Equations (15.106, 15.107) of the form

$$\rho(V, t|V_0, t_0) = \rho_{ss}(V) + \delta\rho(V, \lambda) \exp(\lambda t) \tag{15.108}$$

$$\nu(t) = \nu_{ss} + \delta\nu(\lambda) \exp(\lambda t), \tag{15.109}$$

where ρ_{ss}, ν_{ss} correspond to the stationary solution, $\delta\rho$ and $\delta\nu$ are small deviations around the stationary solution that evolves in time with the (complex) growth rate λ. Upon inserting Equation (15.109) in Equations (15.106) or (15.107), and keeping the term first order in $\delta\rho$ and $\delta\nu$, an equation results for the possible growth rates λ. Solutions with $\mathrm{Re}(\lambda) > 0$ indicate that the stationary state is unstable. Instabilities

might come about with a positive real eigenvalue: this is a mean rate instability, which typically occur in a network with strong recurrent excitation. Alternatively, an instability associated with a positive real part of a complex eigenvalue signals a Hopf bifurcation. If the bifurcation is supercritical, the network exhibits a synchronized oscillation with a frequency close to $\mathrm{Im}(\lambda)$. For more details on this approach, see [1] for the scenario with a simplified model with purely external noise, and [24] for a model with recurrent noise.

References

[1] L. F. Abbott and C. van Vreeswijk (1993), Asynchronous states in a network of pulse-coupled oscillators, *Phys. Rev. E,* **48**: 1483-1490.

[2] M. Abramowitz and I. A. Stegun (1970), *Tables of Mathematical Functions,* Dover Publications, NY.

[3] E. D. Adrian (1928), *The Basis of Sensation: The Action of the Sense Organs*, W. W. Norton: NY.

[4] B. Ahmed and J. C. Anderson and R. J. Douglas and K. A. Martin and J. C. Nelson, Polyneuronal innervation of spiny stellate neurons in cat visual cortex, *J. Comp. Neurol.,* **341**: 39-49.

[5] S. Amari (1977), Dynamics of pattern formation in lateral-inhibition type neural fields, *Biol. Cybern.,* **27**: 77-87.

[6] D. J. Amit (1995), The Hebbian paradigm reintegrated: local reverberations as internal representations, *Behav. Brain Sci.,* **18**: 617.

[7] D. J. Amit and M. V. Tsodyks (1992), Effective neurons and attractor neural networks in cortical environment, *Network*, **3**, 121-137.

[8] D. J. Amit and M. V. Tsodyks (1991), Quantitative study of attractor neural network retrieving at low spike rates I: Substrate – spikes, rates and neuronal gain, *Network*, **2**: 259-274.

[9] D. J. Amit and M. V. Tsodyks (1991), Quantitative study of attractor neural network retrieving at low spike rates II: Low-rate retrieval in symmetric networks, *Network*, **2**: 275.

[10] D. J. Amit and N. Brunel (1997), Dynamics of a recurrent network of spiking neurons before and following learning, *Network,* **8**: 373-404.

[11] D. J. Amit and N. Brunel (1997), Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex, *Cerebral Cortex*, **7**: 237-252.

[12] D. J. Amit and N. Brunel and M. V. Tsodyks (1994), Correlations of cortical

Hebbian reverberations: experiment *vs* theory, *J. Neurosci.,* **14**: 6435-6445.

[13] M. C. Angulo and J. Rossier and E. Audinat (1999), Postsynaptic glutamate receptors and integrative properties of fast-spiking interneurons in the rat neocortex, *J. Neurophysiol.,* **82**: 1295-1302.

[14] W. Bair and E. Zohary and W. T. Newsome (2001), Correlated firing in macaque visual area MT: time scales and relationship to behavior, *J. Neurosci,* **21**: 1676-1697.

[15] M. Bartos and I. Vida and M. Frotscher and J. R. P. Geiger and P. Jonas (2001), Rapid signaling at inhibitory synapses in a dendate gyrus interneuron network, *J. Neurosci.,* **21**: 2687-2698.

[16] R. Ben-Yishai and R. Lev Bar-Or and H. Sompolinsky (1995), Theory of orientation tuning in visual cortex, *Proc. Natl. Acad. Sci. USA,* **92**: 3844-3848.

[17] V. Braitenberg and A. Schutz (1991), *Anatomy of the Cortex*, Springer-Verlag.

[18] P. C. Bressloff and S. Coombes (2000), Dynamics of strongly coupled spiking neurons, *Neural Computation,* **12**: 91-129.

[19] N. Brunel (2000), Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons, *J. Comput. Neurosci.,* **8**: 183-208.

[20] N. Brunel (2000), Persistent activity and the single cell f-I curve in a cortical network model, *Network*, **11**: 261-280.

[21] N. Brunel and F. Chance and N. Fourcaud and L. Abbott (2001), Effects of synaptic noise and filtering on the frequency response of spiking neurons, *Phys. Rev. Lett.,* **86**: 2186-2189.

[22] N. Brunel and P. Latham (2003), Firing rate of noisy quadratic integrate-and-fire neurons, submitted manuscript.

[23] N. Brunel and S. Sergi (1998), Firing frequency of integrate-and-fire neurons with finite synaptic time constants, *J. Theor. Biol.,* **195**: 87-95.

[24] N. Brunel and V. Hakim (1999), Fast global oscillations in networks of integrate-and-fire neurons with low firing rates, *Neural Computation,* **11**: 1621-1671.

[25] N. Brunel and X. J. Wang (2001), Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition, *J. Comput. Neurosci.,* **11**: 63-85.

[26] N. Brunel and X.-J. Wang (2003), What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance, *J. Neurophysiol. in press*.

[27] M. A. Cohen and S. Grossberg (1983), Absolute stability of global pattern formation and parallel memory storage by competitive neural networks, *Transactions IEEE*, **SMC-13**: 815-826.

[28] M. Camperi and X.-J. Wang (1998), A model of visuospatial short-term memory in prefrontal cortex: recurrent network and cellular bistability, *J. Comput. Neurosci.*, **5**: 383-405.

[29] A. Compte and N. Brunel and P. S. Goldman-Rakic and X.-J. Wang (2000), Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model, *Cerebral Cortex,* **10**: 910-923.

[30] A. Compte and C. Constantinidis and J. Tegnér and S. Raghavachari and M. Chafee and P. S. Goldman-Rakic and X.-J. Wang (2002), Spectral properties of mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task, (Submitted to *J. Neurophysiol.*).

[31] C. Constantinidis and P. S. Goldman-Rakic (2002) Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex, *J. Neurophysiol.,* **88**: 3487-3497.

[32] A. Destexhe and D. Paré (1999), Impact of network activity on the integrative properties of neocortical pyramidal neurons *in vivo*, *J. Neurophysiol.,* **81**: 1531-1547.

[33] A. Destexhe and Z. F. Mainen and T. J. Sejnowski (1998), Kinetic models of synaptic transmission, in *Methods in Neuronal Modeling*, C. Koch and I. Segev (eds.), MIT Press, Cambridge, MA, 1-25.

[34] C. R. Doering and P. S. Hagan and C. D. Levermore(1987), Bistability driven by weakly colored gaussian-noise: the fokker-planck boundary layer and mean 1st-passage times, *Phys. Rev. Lett,* **59**: 2129-2132.

[35] C. A. Erickson and R. Desimone (1999), Responses of macaque perirhinal neurons during and after visual stimulus association learning, *J. Neurosci.*, **19**: 10404-10416.

[36] G. B. Ermentrout (1994), Reduction of conductance based models with slow synapses to neural nets, *Neural Computation,* **6**: 679-695.

[37] G. B. Ermentrout (1998), Neural networks as spatio-temporal pattern-forming systems, *Rep. Prog. Phys.*, **61**: 353-430.

[38] J. Feng and D. Brown (2000), Impact of correlated inputs on the output of the ingrate-and-fire model, *Neural Computation,* **12**: 671-92.

[39] D. Ferster and K. D. Miller (2000), Neural mechanisms of orientation selectivity in the visual cortex, *Annu. Rev. Neurosci.,* **23**: 441-471.

[40] N. Fourcaud and N. Brunel (2002), Dynamics of firing probability of noisy integrate-and-fire neurons, *Neural Computation,* **14**: 2057-2110.

[41] S. Funahashi and C. J. Bruce and P. S. Goldman-Rakic (1989), Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex, *J. Neurophysiol.*, **61**: 331-349.

[42] S. Funahashi and C. J. Bruce and P. S. Goldman-Rakic (1991), Neuronal activ-

ity related to saccadic eye movements in the monkey's dorsolateral prefrontal cortex, *J. Neurophysiol.*, **65**: 1464-1483.

[43] S. Funahashi and C. J. Bruce and P. S. Goldman-Rakic (1990), Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms, *J. Neurophysiol.*, **63**: 814-831.

[44] J. M. Fuster (1995), *Memory in the Cerebral Cortex,* MIT Press, Cambridge MA.

[45] J. M. Fuster (1973), Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory, *J. Neurophysiol.,* **36**: 61-78.

[46] J. M. Fuster and G. Alexander (1971), Neuron activity related to short-term memory, *Science*, *173*: 652-654.

[47] J. M. Fuster and J. P. Jervey (1981), Inferotemporal neurons distinguish and retain behaviourally relevant features of visual stimuli, *Science*, **212**: 952-955.

[48] J. M. Fuster and R. H. Bauer and J. P. Jervey (1982), Cellular discharge in the dorsolateral prefrontal cortex of the monkey in cognitive tasks, *Exp. Neurol.,* **77**: 679-694.

[49] C. W. Gardiner (1986), *Handbook of Stochastic Methods for Physics Chemistry and the Natural Sciences,* Springer-Verlag.

[50] C. Geisler and N. Brunel and X.-J. Wang (2003), What determines the frequency of fast network oscillations with irregular neural discharges? II. Contributions of single cell membrane dynamics, in preparation.

[51] W. Gerstner (2000), Population dynamics of spiking neurons: fast transients, asynchronous states, and locking, *Neural Computation* **12**: 43-89.

[52] W. Gerstner (1995), Time structure of the activity in neural network models, *Phys. Rev. E ,* **51**: 738-758.

[53] W. Gerstner and W. M. Kistler (2002), *Spiking Neuron Models: Single Neurons, Populations, Plasticity,* Cambridge University Press.

[54] D. T. Gillespie (1992), *Markov Processes, an Introduction for Physical Scientists,* Academic Press.

[55] P. S. Goldman-Rakic (1995), Cellular basis of working memory, *Neuron,* **14**: 477-485.

[56] P. S. Goldman-Rakic (1987), Circuitry of primate prefrontal cortex and regulation of behavior by representational memory, in *Handbook of Physiology – The Nervous System V*, F. Plum and V. Mountcastle (eds.), Bethesda, Maryland: American Physiological Society, 373-417.

[57] J. Guckenheimer and P. Holmes (1983), *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Vields,* Springer Verlag.

[58] A. Gupta and Y. Wang and H. Markram (2000), Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex, *Science,* **287**: 273-278.

[59] M. Gur and A. Beylin and D. M. Snodderly (1997), Response variability of neurons in primary visual cortex (V1) of alert monkeys, *J. Neurosci.,* **17**: 2914-2920.

[60] B. S. Gutkin and C. R. Laing and C. L. Colby and C. C. Chow and G. B. Ermentrout (2001), Turning on and off with excitation: the role of spike time asynchrony and synchrony in sustained neural activity, *J. Comput. Neurosci.*, **11**: 121-134.

[61] P. S. Hagan and C. R. Doering and C. D. Levermore (1989), Mean exit times for particles driven by weakly colored noise, *SIAM J. Appl. Math.,* **49:** 1480-1513.

[62] D. Hansel and G. Mato (2003), Asynchronous states and the emergence of synchrony in large networks of interacting excitatory and inhibitory neurons, *Neural Comp.,* **15**: 1-56.

[63] D. Hansel and H. Sompolinsky (1998), Modeling feature selectivity in local cortical circuits, in *Methods in Neuronal Modeling,* C. Koch and I. Segev (eds.), MIT Press, Cambridge, MA.

[64] J. J. Hopfield (1984), Neurons with graded response have collective computational properties like those of two-state neurons, *Proc. Natl. Acad. Sci. U.S.A.*, **81**: 3088-3092.

[65] B. Hutcheon and Y. Yarom (2000), Resonance, oscillation and the intrinsic frequency preferences of neurons, *TINS,* **23**: 216-222.

[66] C. E. Jahr and C. F. Stevens (1990), Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics, *J. Neurosci.,* **10**: 3178-3182.

[67] M. M. Klosek and P. S. Hagan (1998), Colored noise and a characteristic level crossing problem, *J. Math. Phys.,* **39**: 931-953.

[68] B. W. Knight (1972), Dynamics of encoding in a population of neurons, *J. Gen. Physiol.,* **59**: 734-766.

[69] C. Koch (1999), *Biophysics of Computation: Information Processing in Single Neurons*, Oxford University Press.

[70] K. W. Koch and J. M. Fuster (1989), Unit activity in monkey parietal cortex related to haptic perception and temporary memory, *Exp. Brain Res.,* **76**: 292-306.

[71] A. A. Koulakov and S. Raghavachari and A. Kepecs and J. E. Lisman, Model for a robust neural integrator, *Nat. Neurosci.*, **5**: 775-782.

[72] U. Kraushaar and P. Jonas (2000), Efficacy and stability of quantal GABA

release at a hippocampal interneuron-principal neuron synapse, *J. Neurosci.*, **20**: 5594-5607.

[73] K. Kubota and H. Niki (1971), Prefrontal cortical unit activity and delayed alternation performance in monkeys, *J. Neurophysiol.*, **34:** 337-347.

[74] D. Lee and N. L. Port and W. Kruse and A. P. Georgopoulos (1998), Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex, *J. Neurosci.,* **18**: 1161-1170.

[75] J. E. Lisman and J.-M. Fellous and X.-J. Wang (1998), A role for NMDA-receptor channels in working memory, *Nat. Neurosci.*, **1**: 273-275.

[76] K. A. Martin (2002), Microcircuits in the visual cortex, *Curr. Opin. Neurobiol.* **12**: 418-425.

[77] M. Mattia and P. Del Giudice (2002), Population dynamics of interacting spiking neurons, *Phys. Rev. E*, **66**: 051917.

[78] C. Meyer and C. van Vreeswijk (2002), Temporal correlations in stochastic networks of spiking neurons, *Neural Computation,* **14**: 369-404.

[79] E. K. Miller and C. A. Erickson and R. Desimone (1996), Neural mechanisms of visual working memory in prefrontal cortex of the macaque, *J. Neurosci.*, **16**: 5154-5167.

[80] P. Miller and C. Brody and R. Romo and X.-J. Wang (2003), A network model of parametric working memory, Submitted to *Cerebral Cortex*.

[81] Y. Miyashita (1988), Neuronal correlate of visual associative long-term memory in the primate temporal cortex, *Nature*, **335**: 817-820.

[82] R. Moreno and J. de la Rocha and A. Renart and N. Parga (2002), Response of spiking neurons to correlated inputs, *Phys. Rev. Lett.*, **89**: 288101.

[83] R. Moreno and N. Parga (2003), Response of a leaky integrate and fire neuron to white noise input filtered by synapses with an arbitraty time constant, in preparation.

[84] V. B. Mountcastle (1997), The cortical organization of the neocortex, *Brain*, **120**: 701-722.

[85] K. Nakamura and K. Kubota (1995), Mnemonic firing of neurons in the monkey temporal pole during a visual recognition memory task, *J. Neurophysiol.*, **74**: 162-178.

[86] Y. Naya and K. Sakai and Y. Miyashita (1996), Activity of primate inferotemporal neurons related to a sought target in pair-association task, *Proc. Natl. Acad. Sci. USA*, **93**: 2664-2669.

[87] L. Nowak and P. Bregestovski and P. Ascher and A. Herbet and A. Prochiantz (1984), Magnesium gates glutamate-activated channels in mouse central neurones, *Nature*, **307**: 462-5.

[88] D. Q. Nykamp and D. Tranchina (2000), A population density approach that facilitates large-scale modeling of neural networks: analysis and an application to orientation tuning, *J. Comp. Neurosci.,* **8:** 19-30.

[89] S. P. Ó Scalaidhe and F. A. W. Wilson and P. S. Goldman-Rakic (1997), Areal segregation of face-processing neurons in prefrontal cortex, *Science*, **278**: 1135-1138.

[90] B. Pesaran and J. S. Pezaris and M. Sahani and P. P. Mitra and R. A. Andersen (2002), Temporal structure in neuronal activity during working memory in macaque parietal cortex, *Nat. Neurosci.*, **5**: 805-811.

[91] A. D. Redish and A. N. Elga and D. S. Touretzky (1996), A coupled attractor model of the rodent head direction system, *Network,* **7**: 671-685.

[92] S. G. Rao and G. V. Williams and P. S. Goldman-Rakic (1999), Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in PFC, *J. Neurophysiol.*, **81**: 1903-1916.

[93] A. Renart (2000), *Multi-Modular Memory Systems*, Universidad Autónoma de Madrid.

[94] A. Renart, J. de la Rocha, N. Parga, and E.T. Rolls (2001). A model of the IT-PF network in object working memory which includes balanced persistent activity and tuned inhibition. *Neurocomputing* **38-40**, 1525-1531.

[95] A. Renart and P. Song and X. J. Wang (2003), Homeostatic synaptic plasticity leads to robust spatial working memory without fine-tuning of cellular or network properties, Submitted to *Neuron*.

[96] A. Renart and R. Moreno and X. J. Wang and N. Parga (2003), Bistability in balanced recurrent networks, submitted.

[97] L. M. Ricciardi (1977), *Diffusion Processes and Related Topics on Biology* Springer-Verlag, Berlin.

[98] M. Richardson and N. Brunel and V. Hakim (2003), From subthreshold to firing-rate resonance, *J. Neurophysiol.,* **89**: in press.

[99] H. Risken (1984), *The Fokker-Planck Equation: Methods of Solution and Applications,* Springer-Verlag, Berlin.

[100] K. Sakai and Y. Miyashita (1991), Neural organization for the long-term memory of paired associates, *Nature*, **354**: 152-155.

[101] E. Salinas and T. J. Sejnowski (2000), Impact of correlated synaptic input on output firing rate and variability in simple neuronal models, *Journal of Neuroscience*, **20**: 6193-6209.

[102] E. Salinas and T. J. Sejnowski (2002), Integrate-and-fire neurons driven by correlated stochastic input, *Neural Computation,* **14**: 2111-2155.

[103] H. S. Seung (1996), How the brain keeps the eyes still, *Proc. Natl. Acad. Sci. USA,* **93**: 13339-13344.

[104] H. S. Seung and D. D. Lee and B. Y. Reis and D. W. Tank (2000), Stability of the memory of eye position in a recurrent network of conductance-based model neurons, *Neuron,* **26**: 259-271.

[105] M. N. Shadlen and W. T. Newsome (1994), Noise, neural codes and cortical organization, *Current opinion in Neurobiol.,* **4**: 569-579.

[106] M. N. Shadlen and W. T. Newsome (1998), The variable discharge of cortical neurons: implications for connectivity, computation, and information coding, *J. Neurosci.,* **18**: 3870-3896.

[107] M. J. Shelley and D. McLaughlin and R. Shapley and J. Wielaard (2002), States of high conductance in a large-scale model of the visual cortex, *J. Comput. Neurosci.,* **13**: 93-109.

[108] S. Shinomoto and Y. Sakai and S. Funahashi (1999), The Ornstein-Uhlenbeck process does not reproduce spiking statistics of neurons in prefrontal cortex, *Neural Comput.*, **11**: 935-51.

[109] O. Shriki and D. Hansel and H. Sompolinsky (2003), Rate models for conductance based cortical neuronal networks, *Neural Comput.*, **15**: in press.

[110] W. R. Softky and C. Koch (1993), The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs, *J. Neurosci.*, **13**: 334-350.

[111] H. Sompolinsky and R. Shapley (1997), New perspectives on the mechanisms for orientation selectivity, *Curr. Opin. Neurobiol.,* **7**: 514-522.

[112] J. Tegnér and A. Compte and X.-J. Wang (2002), Dynamical stability of reverberatory neural circuits, *Biol. Cybern.* **87**: 471-481.

[113] A. M. Thomson (2000), Facilitation, augmentation and potentiation at central synapses, *Trends in Neurosciences,* **23**: 305-312.

[114] A. Treves (1993), Mean-field analysis of neuronal spike dynamics, *Network,* **4**: 259-284.

[115] M. V. Tsodyks and K. Pawelzik and H. Markram (1998), Neural networks with dynamic synapses, *Neural Computation,* **10**: 821-835.

[116] M. V. Tsodyks and and W. E. Skaggs and T. J. Sejnowski and B. L. McNaughton (1997), Paradoxical effects of external modulation of inhibitory interneurons, *J. Neurosci.*, **17**: 4382-4388.

[117] H. C. Tuckwell (1988), *Introduction to Theoretical Neurobiology,* Cambridge: Cambridge University Press.

[118] C. van Vreeswijk and H. Sompolinsky (1996), Chaos in neuronal networks with balanced excitatory and inhibitory activity, *Science*, **274**: 1724-1726.

[119] C. van Vreeswijk and H. Sompolinsky (1998), Chaotic balanced state in a model of cortical circuits, *Neural Computation,* **10**: 1321-1371.

[120] X.-J. Wang (1999), Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory, *J. Neurosci.,* **19**: 9587-9603.

[121] X.-J. Wang (2001), Synaptic reverberation underlying mnemonic persistent activity, *Trends Neurosci.,* **24**: 455-463.

[122] X.-J. Wang and G. Buzsáki (1996), Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model, *J. Neurosci.,* **16**: 6402-6413.

[123] N. Wax (1954), *Selected Papers on Noise and Stochastic Processes,* Dover Publications.

[124] G. V. Williams and P. S. Goldman-Rakic (1995), Modulation of memory fields by dopamine D1 receptors in prefrontal cortex, *Nature*, **376**: 572-575.

[125] H. R. Wilson and J. D. Cowan (1972), Excitatory and inhibitory interactions in localized populations of model neurons, *Biophys. J.,* **12**: 1-24.

[126] H. R. Wilson and J. D. Cowan (1973), A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue, *Kybernetik,* **13**: 55-80.

[127] Z. Xiang and J. R. Huguenard and D. A. Prince (1998), $GABA_A$ receptor mediated currents in interneurons and pyramidal cells of rat visual cortex, *J. Physiol.,* **506**: 715-730.

[128] K. Zhang (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensembles: A theory, *J. Neurosci.,* **16**: 2112-2126.

[129] F.-M Zhou and J. J. Hablitz (1998), AMPA receptor-mediated EPSCs in rat neocortical layer II/III interneurons have rapid kinetics, *Brain Research,* **780**: 166-169.

[130] E. Zohary and M. N. Shadlen and W. T. Newsome (1994), Correlated neuronal discharge rate and its implications for psychophysical performance, *Nature*, **370**: 140-3.

[131] R. S. Zucker and W. G. Regehr (2002), Short-term synaptic plasticity, *Annu. Rev. Physiol.*, **64**: 355-405.