Time structure of the activity in neural network models

Wulfram Gerstner
Institut für Theoretische Physik,
Physik-Department der Technischen Universität München, D-85748 Garching bei München, Germany
(Received 10 May 1994)

Several neural network models in continuous time are reconsidered in the framework of a general mean-field theory which is exact in the limit of a large and fully connected network. The theory assumes pointlike spikes which are generated by a renewal process. The effect of spikes on a receiving neuron is described by a linear response kernel which is the dominant term in a weak-coupling expansion. It is shown that the resulting "spike response model" is the most general renewal model with linear inputs. The standard integrate-and-fire model forms a special case. In a network structure with several pools of identical spiking neurons, the global states and the dynamic evolution are determined by a nonlinear integral equation which describes the effective interaction within and between different pools. We derive explicit stability criteria for stationary (incoherent) and oscillatory (coherent) solutions. It is shown that the stationary state of noiseless systems is 'almost always' unstable. Noise suppresses fast oscillations and stabilizes the system. Furthermore, collective oscillations are stable only if the firing occurs while the synaptic potential is increasing. In particular, collective oscillations in a network with delayless excitatory interaction are at most semistable. Inhibitory interactions with short delays or excitatory interactions with long delays lead to stable oscillations. Our general results allow a straightforward application to different network models with spiking neurons. Furthermore, the theory allows an estimation of the errors introduced in firing rate or "graded-response" models.

PACS number(s): 87.10.+e, 05.90.+m, 82.40.Bj

I. INTRODUCTION

A. Temporal structure of neural activity

The theoretical description of neuronal activity has a long tradition in both mathematics and biology [1–3] The interest of physicists, however, seems to be more recent and largely due to the work of Hopfield [4,5] which is a continuation of earlier studies on associative memory networks; cf. the reprint volumes [6,7]. The formulation of Hopfield opened the path for an application of statistical mechanics to the theory of neural networks [8,9]. In the following years a large number of researchers studied the associative properties and the storage capacity of the Hopfield network and extensions thereof. For reviews consult [10–14].

In this paper we discuss a more recent development in the field, that is, the role of time in neuronal activity. The potential relevance of temporal structure arises from two different sources. First, the signal of *single* neurons has a pulselike structure. More precisely, it consists of a sequence of action potentials, short spikes of the membrane voltage. It is intuitively clear that the exact spike times or the precise length of interspike intervals may contain information which is not contained in the time averaged mean firing rate [15–20]. In particular, in all situations where fast reaction of a system is required there is no time for temporal averaging and single spikes should be important. A careful evaluation of experimental results has given some support to this idea [21–24]. Second, correlated changes in the activity of *several* neurons can convey additional information beyond the single neuron activity [25–30]; for reviews see [31,32]. For example, synchronous activity of a population of neurons could have the additional meaning that these neurons somehow relate to the same object in the outside world [33–35]. Coherent oscillations as found in the visual cortex of cat and monkey can be interpreted in the light of this hypothesis [36–45].

Both aspects of time have been disregarded in the standard approach to neural networks. In all standard models, the output of the model neuron, be it discrete or analog, is usually interpreted as the mean firing rate, a time averaged quantity; e.g., [4,13]. This is also true for graded-reponse neurons where a membrane time constant of the neurons is taken into account [5,46]. In formal associative memory networks, the role of time is mainly restricted to the retrieval dynamics, except for those networks where temporal sequences of patterns have been learned [47–54]. Some years ago, a formal model with two time scales only was analyzed by Choi [55]. Recently, however, a couple of more detailed models have been investigated with respect to the temporal structure of neuronal activity. In particular, several model networks of spiking neurons in continuous time have been studied analytically [56–62].

Some of those investigations concern fully connected networks of *integrate-and-fire* neurons. In the one-step version of this model, pointlike spikes are directly fed into the central membrane equation, whereas in the two-step or three-step model, spikes are low-pass filtered first. The theory of Mirollo and Strogatz [56] which is applicable to the one-step integrate-and-fire model in continu-

ous time shows that almost all initial conditions eventually lead to perfect synchrony. This is different in the discrete-time version of the one-step integrate-and-fire model without external input current [63]. The global behavior of integrate-and-fire networks in continuous time can be described with dynamic mean-field equations as discussed by Kuramoto [57] for the one-step, by Tsodyks *et al.* [60] and Treves [61] for the two-step, and by Abbott and van Vreeswijk [62] for the three-step model. A similar approach for discrete-time dynamics has been taken by Usher *et al.* [64]. Tsodyks *et al.* [60] discuss the effect of small inhomogeneities induced by variations in the parameters in a standard two-step model. Treves [61] has studied transients and adaptation in the generalized two-step model with two different ion currents. His approach applies not only to homogeneous networks, but also to systems consisting of several pools; cf. Sec. II C. Abbott and van Vreeswijk [62] have analyzed the stability of stationary states in the generalized form of the three-step model. Finally, simulation studies show that a network of integrate-and-fire neurons can be used as an associative memory [65,66].

In contrast to integrate-and-fire models which are usually stated in the form of differential equations, the spike response model [58,59,67–71] is based on the idea of response kernels which describe the integrated effect of spike reception or spike emission on the membrane potential. The properties of single spike response neurons and networks thereof have been investigated in [58,67]. A large and fully connected network of spike response neurons has been used as an associative memory for a finite number of patterns [58]. Collective states of such a network have been studied both in simulations and analytically [58,59]. In particular, the dynamic evolution of homogeneous networks has been described by an integral equation for the instantaneous activity averaged over the network. The stability of stationary and oscillatory solutions has been analyzed in the low-noise limit [59,70,71]. In addition to stationary and oscillatory states, the potential relevance of spatiotemporal spike patterns has been demonstrated [69]. An extended version of the model that includes local inhibition and some inhomogeneities by variations in the parameters can be found in [68,72]. A homogeneous model in discrete time and nonlinear synaptic transmission has been studied by Bauer and Pawelzik [73].

A couple of other models have been used in studies of large neural systems. The standard model of single neuron activity is the Hodgkin-Huxley model [3], a set of four coupled differential equations to describe action potentials in the giant axon of the squid. Compartmental versions of the model with additional ion currents have been used by several groups [74–77]. Due to the complicated system of nonlinear equations, network studies have been limited to simulations, usually with a small number of neurons only [78–80]. There are several other simulation studies based on different versions of nonlinear differential equations [39,81–85]. Analytical results with a piecewise linear neuron model have been achieved by Abbott [86]. Oscillations can also be found in networks of formal two-state neurons with local inhibition [87]. Finally, several researchers have used the Kuramoto model [88,89] to describe collective oscillations in neuronal networks [90–94]. A slightly different version of this model has been studied by Kurrer and Schulten [95]. Depending on the level of intrinsic noise they find a phase transition from the quiescent state to a collective oscillation.

B. Overview

In this paper we address four major topics.

(1) The first and central question concerns the *relation* between different model neurons frequently used in neural network studies. In Sec. I C we give a short review of some of the more popular model neurons and introduce our notation. In Secs. II A and II B we show that, with a simple renewal assumption, all of the *spiking* model neurons can be analyzed in terms of two linear response kernels which describe the effect of a spike on the emitting and the receiving neuron. The kernels allow us to establish a relation between various model neurons, in particular between different versions of the integrate-and-fire neuron and the spike response model. Another popular model, the graded-response neuron, is based on a *firing rate* description. How rate models can be related to spike models is the question which we address in Sec. III B. In particular, we derive a systematic estimation of the errors introduced, if neuronal activity is described by the graded-response model.

(2) If a relation between different neural network models can be established, it would be desirable to have a *general theory* of the global network dynamics which contains the various models as special cases. This is the second topic which is addressed in this paper. A possible approach to this question is presented in Sec. II C where we introduce the concept of *pools* of equivalent neurons. The pool concept in combination with the renewal assumption allows us to derive a general equation for the collective pool dynamics. Specification of the response kernels leads back to the different model networks.

(3) Our approach by pool equations allows us to address a third topic. In some models, the basic units of a network are not considered to represent single neurons, but rather localized *populations* of neurons. Wilson and Cowan [96] have derived macroscopic equations for the dynamics and interactions of pools of neurons, but their results are limited to a rather special case. In many other models, *ad hoc* nonlinear differential equations have been used to describe the effective pool dynamics. Here we address the question of whether a systematic derivation of the effective dynamics and interaction of pools of neurons is possible, if the single neuron dynamics is known. The central ideas and results regarding this question are presented in Sec. II C.

(4) The final topic of this paper concerns the stability of collective network states, in particular, of coherent oscillations. Several researchers have studied collective oscillations using rather specific network models. Here we address the question of whether it is possible to find *universal*, that is, model independent, *stability criteria*. A fairly general answer is derived in Sec. III D. The results can be applied to various versions of the integrate-

and-fire model and the spike response model, both in the case of excitatory and inhibitory interaction. This also solves the question of whether excitatory or inhibitory interactions are more appropriate to sustain collective oscillations.

In Sec. IV, our results regarding those issues are summarized in a couple of conclusions.

C. Review of current models

1. *Firing rate models*

Neurons which are driven by some stimulus usually emit action potentials, that is, short solitonlike pulses of electrical activity which travel along the axon to the synapses on the dendritic trees of other neurons and thus carry the signal to them. If we count the number of action potentials emitted by a neuron in some time interval and divide by the length of the interval, we find the mean firing rate of the neuron. By definition it is a time averaged quantity, typical time windows ranging from 100 to over 1000 ms. The dependence of the mean firing rate ν_i of neuron i upon its input h_i,

$$\nu_i = g(h_i), \tag{1}$$

is called the gain function of the neuron. In all standard approaches to neural networks, be it a feedforward or a recurrent net, the input h_i to a postsynaptic neuron i is the sum of the incoming activity

$$h_i = \sum_{j\in\Gamma_i} J_{ij}\nu_j, \tag{2}$$

where the sum runs over all neurons which are presynaptic to neuron i, i.e., $\Gamma_i = \{j \mid j \text{ sends signals to } i\}$. The prefactor J_{ij} denotes the synaptic efficacy of a synapse from j to i and is adjusted by some learning rule. The gain function $g(h)$ is usually taken to be a sigmoidal, e.g., $g(h) = (1+\tanh\beta h)/2$. Equations (1) and (2) define the *rate model of neural activity* as used in most standard approaches; cf. [13]. The solutions of the fixed-point equation

$$\nu_i = g\left(\sum_{j\in\Gamma_i} J_{ij}\nu_j\right) \tag{3}$$

define the collective states in firing rate models. Note that all quantities are stationary, that is, independent of time.

There are several ways to introduce a dynamics with (3) as the fixed points. One possibility is by discretizing time and reading Eq. (3) as a mapping from one time step to the next, that is, $\nu_i(t+1) = g[\sum_{j\in\Gamma_i} J_{ij}\nu_j(t)]$. In the following we concentrate on models in continuous time. In this case, a rate dynamics can be defined by introduction of a time constant τ turning (3) into a differential equation

$$\tau\frac{d\nu_i}{dt} = -\nu_i + g\left(\sum_{j\in\Gamma_i} J_{ij}\nu_j\right). \tag{4}$$

The time constant τ can be interpreted as a membrane parameter $\tau = RC$ where R and C are resistance and, respectively, capacitance of the neuron. For later convenience, we transform (4) into a differential equation for the membrane potential. With $h_i = \sum_{j\in\Gamma_i} J_{ij}\nu_j$ we have

$$\tau\frac{dh_i}{dt} = -h_i + \sum_{j\in\Gamma_i} J_{ij}g(h_j). \tag{5}$$

This is the neuron equation in the well known "graded-response" model as discussed by Hopfield [5] and similarly by Cohen and Grossberg [46].

In some models, Eq. (4) or (5) is used in a different interpretation. The time averaged rate ν_i in (4) is replaced by the *spatially averaged activity* $A_i(t)$ of a localized population of neurons. Thus instead of (4) we have

$$\tau\frac{dA_i}{dt} = -A_i + g\left(\sum_{j\in\Gamma_i} J_{ij}A_j\right). \tag{6}$$

Feldman and Cowan [97] have shown that (6) can be derived from microscopic equations of a neuronal threshold dynamics under the assumption that all quantities vary only slowly in time. In other words, oscillations and fast transients are neglected.

2. *The model of Wilson and Cowan*

A decade before the recent interest in neural networks started, Wilson and Cowan [96] studied the dynamics of interacting populations of neurons from a slightly different point of view. They have derived an integral equation governing the dynamics of the local activity $A(\mathbf{x},t)$ which, in our notation, reads

$$A(\mathbf{x},t+\Delta t) = \left\{1 - \int_0^{\gamma^{\text{refr}}} ds\, A(\mathbf{x},t-s)\right\} S_{\mathbf{x}}[h(\mathbf{x},t)], \tag{7}$$

with $h(\mathbf{x},t) = \int_{-\infty}^{t} dt' \epsilon(t-t')\,[\sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y})A(\mathbf{y},t') + P(t')]$. Here, $A(\mathbf{x},t)$ is the activity of an excitatory or inhibitory population labeled by the index $\mathbf{x}$, $P(t)$ is some external stimulus, $\epsilon(s')$ is the synaptic response function, γ^{refr} is the absolute refractory period, and $S_{\mathbf{x}}(h)$ is a monotonically increasing function, $0 \le S_{\mathbf{x}} \le 1$, which denotes the portion of neurons that fire at a stimulation level h. The prefactor in curly brackets arises since neurons which are refractory are insensitive to stimulation. Based on a "time coarse-graining" procedure, Wilson and Cowan transformed the integral equation (7) into a differential equation.

$$\tau\frac{d}{dt}A(\mathbf{x},t) = -A(\mathbf{x},t) + \left\{k - \gamma^{\text{refr}}A(\mathbf{x},t)\right\} S_{\mathbf{x}}[h(\mathbf{x},t)], \tag{8}$$

with some constant k and the membrane potential $h(\mathbf{x},t) = \sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y})A(\mathbf{y},t) + P(t)$. As before, the term in curly brackets arises due to the absolute refractory period of the neurons. In the limit of vanishing refractoriness ($\gamma^{\text{refr}} \to 0$), (8) is equivalent to the graded-response model in the interpretation (6). Wilson and Cowan [96] studied the macroscopic dynamics of two coupled populations, one excitatory and one inhibitory, by a phase plane analysis of (8). The collective states exhibit a rich behavior including hysteresis and limit cycles.

3. *Integrate-and-fire models*

Firing rate models like the graded-response model presuppose that all relevant information is contained in the mean firing rate, a time averaged quantity. In population models, we study the local activity, a spatially averaged quantity. In networks with spiking model neurons, we take single spikes of single neurons as the essential information allowing for a much better spatial and temporal resolution. In the following spikes are always assumed to be short pulses described by a δ function. Thus the spike train of neuron i is

$$S_i^{(F)}(t) = \sum_{f=1}^{F} \delta(t - t_i^f), \tag{9}$$

where t_i^f with $1 \le f \le F$ are the firing times of neuron i. Spikes are labeled such that $f = 1$ denotes the most recent spike, $f = 2$ the second last, and so forth.

In the integrate-and-fire model [1], the neuron is described as a leaky integrator that fires, if the membrane potential $h(t)$ reaches a threshold $\theta > 0$ (for a review see [98,99]). After firing the membrane potential is reset to its initial value $h(t_0) = 0$. Between two spikes the change of the membrane potential is described by the differential equation of a RC circuit,

$$\tau_1 \frac{dh_i}{dt} = -h_i + R\,(I_i^{\text{syn}} + I_i^{\text{ext}}), \tag{10}$$

where $\tau_1 = RC$ is the membrane time constant of a neuron with resistance R and capacitance C. The currents $I_i^{\text{syn}} + I_i^{\text{ext}}$ describe synaptic and external input, respectively.

As mentioned before, spikes are assumed to be δ-function pulses. In the simple version of the model, spikes of presynaptic neurons directly feed into neuron i. Thus

$$I_i^{\text{syn}}(t) = \sum_{j\in\Gamma_i} K_{ij} \sum_{f=1}^{F} \delta(t - t_j^f), \tag{11}$$

where K_{ij} is a coupling constant and t_j^f denotes the firing time of neuron j. We call this the one-step version of the integrate-and-fire model.

Whereas a presynaptic action potential is a comparatively short pulse, the effective input to the postsynaptic neuron is much broader due to the synaptic transmission process. In the two-step version, we allow for an additional synaptic integration time τ_2. To be more precise, we have

$$I_i^{\text{syn}}(t) = \int_0^\infty ds' \frac{1}{\tau_2} \exp\left(-\frac{s'}{\tau_2}\right) \sum_{j\in\Gamma_i} K_{ij} \times \sum_{f=1}^{F} \delta(t - t_j^f - s'). \tag{12}$$

In other words, the input current to neuron i consists of pulses that decay with a time constant τ_2. Similarly we can define a three-step model, if we introduce yet another time constant and take

$$I_i^{\text{syn}}(t) = \frac{1}{\tau_2 - \tau_3} \int_0^\infty ds' \left[\exp\left(-\frac{s'}{\tau_2}\right) - \exp\left(-\frac{s'}{\tau_3}\right)\right] \times \sum_{j\in\Gamma_i} K_{ij} \sum_{f=1}^{F} \delta(t - t_j^f - s'). \tag{13}$$

For $\tau_2 = \tau_3$, each presynaptic spike induces an input current proportional to $s' \exp(-s'/\tau_2)$. This is considered a fairly realistic description of the synaptic input [98].

The standard ansatz in integrate-and-fire models is the linear differential equation (10). It describes the dynamics in a RC ciruit with constant R and C. In real neurons both R and C have an intricate voltage dependence, since ion channels open and close as a function of the membrane voltage and/or the ion concentration. Thus we have to study a generalized version of (10),

$$\frac{dh_i}{dt} = -\frac{1}{\tau(h_i)} h_i + \frac{R(h_i)}{\tau(h_i)} \left[I_i^{\text{syn}}(t) + I_i^{\text{ext}}(t)\right]. \tag{14}$$

As before, I_i^{ext} is the external input current and I_i^{syn} is the input from presynaptic neurons. As an example we can take $R(h) \propto (h - h_0)$ where $h_0 > \theta$ is the reversal potential for a typical ion, say Na^+. Inclusion of several ion currents with specific reversal potentials leads to an even more realistic neuron model (e.g., [74–77]). The generalized integrate-and-fire model (14) is equivalent to a *pulse-coupled phase model.* To show this we set $\alpha(h) \equiv R(h)/\tau(h)$, $f(h) = -h/\tau(h)$, and require that $\alpha(h) \ge 0$ and $f(h) \ne 0$. After a nonlinear transformation, (14) can be written in the form

$$\frac{d\psi_i}{dt} = 1 + \gamma(\psi_i) I_i(t), \tag{15}$$

with $I_i = I_i^{\text{syn}} + I_i^{\text{ext}}$, $\psi = F(\phi) \equiv \int_{\phi_0}^{\phi} d\phi'/f(\phi')$, and $\gamma(\psi) = \alpha[F^{-1}(\psi)]/f[F^{-1}(\psi)]$ where F^{-1} denotes the inverse of the function $F(\phi)$. We call (15) the standard form I. Alternatively, we can study the model in the standard formulation II, i.e.,

$$\frac{d}{dt} V_i = G(V_i) + I_i(t), \tag{16}$$

with $V = u(\psi) \equiv \int_0^{\psi} d\psi'/\gamma(\psi')$ and $G(V) = 1/\gamma[u^{-1}(V)]$.

Equation (16) has an immediate connection to the model of Mirollo and Strogatz [56] which we sketch here

briefly. The state variable V_i of a unit i is bound to the trajectory

$$V_i = u(\psi_i), \tag{17}$$

where ψ is some phase variable and u is a monotonically increasing function with $u''(\psi) < 0$. If V_i (which is related to the membrane potential) reaches the threshold $\theta = 1$, then the unit fires and V_i is reset to zero. A free unit evolves according to $\psi_i = t$ and $V_i = u(t)$. If another unit j fires at a time t_j^f, we have an input $I_i(t) = K_{ij}\,\delta(t - t_j^f)$ and the potential at time t_j^f is raised from V_i to

$$V_i' = \min(V_i + K_{ij}, 1). \tag{18}$$

Since the neuron is bound to its trajectory, this causes a phase shift

$$\Delta = \psi' - \psi = u^{-1}(V_i') - u^{-1}(V_i), \tag{19}$$

where $u^{-1}(V)$ is the inverse of $u(\psi)$. Mirollo and Strogatz [56] could show that in a homogeneous and fully coupled population of units defined by (17) and (18) almost all initial conditions eventually lead to synchrony. This is an important result since it is based on a rigorous argument and concerns the full state space of the system, in contrast to local arguments based on a linear stability analysis.

4. Spike response model

As before, spikes are generated by a threshold process, or, more precisely, the firing time t_i^f is given by the condition that $h_i(t_i^f) = \theta$ where $h(t)$ is the membrane potential and θ the firing threshold. In the spike response model, [58,59,67–71], the net effect that firing has on the emitting and the receiving neuron is described by two different response functions, $\epsilon(s')$ and $\eta^{\rm refr}(s)$. The *refractory function* $\eta^{\rm refr}(s)$ describes the response of the firing neuron to its own spike. Due to causality, $\eta^{\rm refr}(s)$ vanishes for $s \le 0$. In principal, $\eta^{\rm refr}(s)$ could be adjusted to experimental data. The simple ansatz

$$\eta^{\rm refr}(s) = \begin{cases} -\infty & \text{for } 0 < s \le \gamma^{\rm refr} \\ \eta_0/(s - \gamma^{\rm refr}) & \text{for } s > \gamma^{\rm refr} \end{cases} \tag{20}$$

has proven to be useful for simulations. Here, $\gamma^{\rm refr}$ is the absolute refractory period during which the neuron cannot fire again. For $s > \gamma^{\rm refr}$ firing is possible, but more difficult (relative refractory period).

If a neuron fires several spikes in sequence, the effects of refractoriness can add up. Assuming linear superposition we have

$$h_i^{\rm refr}(t) = \int_0^\infty ds\, \eta^{\rm refr}(s) S_i^{(F)} = \sum_{f=1}^{F} \eta^{\rm refr}(t - t_i^f). \tag{21}$$

Summation over several recent spikes ($F > 1$) is a simple procedure to incorporate adaptation effects [58,100,101]. In fact, most neurons do exhibit a pronounced adaptation; see, e.g., [102]. For the sake of analytical transparency, however, it is more convenient to neglect adaptation and take $F = 1$. This is a rather strong simplification and implies that refractoriness depends on the most recent spike only,

$$h_i^{\rm refr}(t) = \eta^{\rm refr}(t - t_i^1). \tag{22}$$

This is a renewal property and will be used throughout Secs. II and III.

The second response function is the *synaptic kernel* $\epsilon(s')$. It describes the effect of an incoming spike on the membrane potential at the soma of the postsynaptic neuron. If we add the contributions of all input signals we have

$$h_i^{\rm syn}(t) = \sum_{j \in \Gamma_i} J_{ij} \int_0^\infty ds' \epsilon(s') S_j^{(F)}(t - s'). \tag{23}$$

The sum runs over all neurons j which are presynaptic to neuron i and J_{ij} is the synaptic efficacy of a connection from j to i. In contrast to (22), here a superposition of an arbitrary number ($F \ge 1$) of spikes of a presynaptic neuron is possible. The shape of the synaptic response function can be chosen according to experimental results. A useful approximation is a delayed α function [98,103],

$$\epsilon(s') = \theta(s' - \Delta^{\rm ax}) \left(\frac{s' - \Delta^{\rm ax}}{\tau_s^2}\right) \exp\left(-\frac{s' - \Delta^{\rm ax}}{\tau_s}\right), \tag{24}$$

where $\Delta^{\rm ax}$ is the transmission delay and τ_s is a membrane time constant. The step function $\theta(x)$ is equal to one for $x > 0$ and vanishes for $x \le 0$.

In principle, the synaptic response can also depend on the state of the receiving neuron, in particular the time s that has passed since the last *postsynaptic* spike. This effect is included if we take a general response kernel $\kappa(s, s')$ instead of $\epsilon(s')$. As will be shown in Sec. II B, this is the most general linear ansatz compatible with renewal theory. As an example, we consider multiplicative dependence, i.e., $\kappa(s, s') = \rho(s)\,\epsilon(s')$ with two arbitrary functions $\rho(s)$ and $\epsilon(s')$ for the postsynaptic and presynaptic contribution, respectively. The postsynaptic part $\rho(s)$ describes the receptivity of the receiving neuron and $\epsilon(s')$ is the synaptic response kernel introduced above. Note that the receptivity $\rho(s)$ is multiplicative whereas the refractory kernel $\eta^{\rm refr}(s)$ is an additive contribution.

The total membrane potential of a neuron i which has fired its last spike at time $t_i^1 = t - s$ is the sum of the refractory potential (22) and the synaptic potential (23),

$$h_i(s, t) = h_i^{\rm ext}(t) + \eta^{\rm refr}(s) + \sum_{j \in \Gamma_i} J_{ij} \int_0^\infty \kappa(s, s')\, S_j^{(F)}(t - s') ds'. \tag{25}$$

The additional contribution $h_i^{\rm ext}(t)$ allows for external input.

Equation (25) together with the threshold process for firing defines the noise-free spike response model. In the case of noise, the neuron can fire even though the membrane potential (25) has not reached the threshold yet or

it may pass the threshold for a short time without firing. A simple way of dealing with noise is by introduction of a firing probability

$$P_F(h;\delta t) = \tau^{-1}(h)\,\delta t, \tag{26}$$

where δt is an infinitesimal time interval. The instantaneous "rate" $\tau^{-1}(h)$ is an arbitrary function with $\tau^{-1} \to \infty$ for $h \gg \theta$ and $\tau^{-1} \to 0$ for $h \ll \theta$. A simple choice is [58]

$$\tau^{-1}(h) = \tau_0^{-1} \exp[\beta(h-\theta)] \tag{27}$$

in analogy to the Arrhenius formula of chemical reaction kinetics. The parameter β is a measure of the noise. For $\beta \to \infty$ we are back to the noiseless threshold behavior.

Note that in simulations, a discrete version of the model with finite time steps Δt must be used. In order to find the firing probability in a finite interval Δt we integrate (26) over time. This yields [58]

$$P_F(h;\Delta t) = 1 - \exp[-\tau^{-1}(h)\,\Delta t]. \tag{28}$$

The exponential factor arises since neurons can fire at most once in a sufficiently short interval of, let us say, $\Delta t = 1$ ms. The function $P_F(h;\Delta t)$ exhibits the common sigmoidal dependence upon h with $0 \le P_F(h;\Delta t) \le 1$.

In (26) we have assumed that the firing probability at time t depends on the momentary value of $h(t)$ only [58,67,71]. In principle, it could also depend on other quantities like the derivatives of h (cf. below, Sec. III E) or the time s that has passed since the last (postsynaptic) spike [104]. Since the latter causes no difficulties in the mathematical analysis, we will include it in Secs. II and III, unless stated otherwise. Thus, instead of (26), we have

$$P_F(s,h;\delta t) = \tau^{-1}(s,h)\delta t. \tag{29}$$

As before, the membrane potential $h(s,t) = h^{\text{ext}}(t) + \eta^{\text{refr}}(s) + h_i^{\text{syn}}(s,t)$ introduces another, *implicit*, s dependence into τ^{-1}.

II. THEORETICAL FRAMEWORK

A. The renewal hypothesis

Renewal theory and the mathematics of point processes have repeatedly been applied to the phenomena of neural activity [17,18,105–107]; for an extensive discussion of single neuron renewal models see [99,108,109]. Here we review the basic ideas and introduce our notation. We suppose that the state of a neuron can be described by a set of variables $v^k(t)$ with $1 \le k \le m$. The number m of state variables depends on the specific model. In the Hodgkin-Huxley model there are four dynamic variables [3], in compartmental models there may be 20 or more (e.g., [76]). In the following we gather the variables into a vector $\mathbf{v}(t) \in \mathbf{R}^m$.

In general, the state of a neuron i depends on the present and past input from all other neurons and external sources as well as on its own history. The synaptic input consists of spike trains, a set of firing times t_j^f; cf. Eq. (9). Similarly, the past of neuron i is reflected in a set of spike times t_i^f. Regarding the external input we assume a current with time course $I_i^{\text{ext}}(t)$. Thus the state of neuron i is given by

$$\mathbf{v}_i(t) = \mathbf{f}\left[t_i^1, t_i^2, \ldots; \{t_j^1, t_j^2, \ldots | j \in \Gamma_i\}, \{I_i^{\text{ext}}(t') | t' \le t\}\right], \tag{30}$$

where j runs over all neurons presynaptic to neuron i and $\mathbf{f} \in \mathbf{R}^m$ is some arbitrary function.

It is the central hypothesis of a renewal theory of neural activity that the dependence of $\mathbf{v}_i$ upon neuron i's *own* history goes back to its *most recent spike only*. This allows us to drop the dependence upon $t_i^2, t_i^3, \ldots$ from (30). With $\{S_j(t')\} \equiv \{\sum_{f=1}^{F} \delta(t' - t_j^f) \,|\, j \in \Gamma_i\}$ we can replace (30) by

$$\mathbf{v}_i(t) = \mathbf{F}[t - t_i^1; \{S_j(t')\}; \{I_i^{\text{ext}}(t') \,|\, t' \le t\}]. \tag{31}$$

In this sense we can say that the internal variables are "reset" after each spike. This is a rather crude approximation to biology since neurons are known to have strong adaptation properties (consult, e.g., [102]). Nevertheless, for the sake of mathematical transparency we assume throughout the following that neurons can indeed be described by a renewal dynamics. As will be shown in the following subsections, this assumption allows an enormous simplification of the model equations.

In the context of renewal theory, it is convenient to consider the survivor function which we introduce now. Let us consider an ensemble of N_1 independent and identical neurons which all receive the same input from a second ensemble $S_j(t), 1 \le j \le N_2$. We assume that all neurons in the first ensemble $1 \le i \le N_1$ have fired their *last* spike at $t_i^1 = t^f$. Due to the renewal assumption, earlier spikes are of no influence. Thus all neurons $1 \le i \le N_1$ have the *same* evolution of the internal state $\mathbf{v}(t')$ for $t' \ge t^f$ as long as they do not fire again. If we allow for a certain amount of noise, some of these neurons fire their next spike a little earlier, others a little later. The probability that a neuron "survives" the time from t^f to $t^f + s$ without firing again is called the "survivor function"

$$S_{\mathbf{v}}(t^f + s | t^f). \tag{32}$$

The index $\mathbf{v}$ is a reminder that the survivor function depends on the state $\mathbf{v}(t')$ during the interval $t^f \le t' \le t^f + s$. Since the portion of neurons which stay quiescent for the time s decreases due to firing, the survivor function is a monotonically decreasing function with $\lim_{s\to 0} S_{\mathbf{v}}(t^f + s|t^f) = 1$ and $\lim_{s\to\infty} S_{\mathbf{v}}(t^f + s|t^f) = 0$. Given the function $S_{\mathbf{v}}(t^f + s|t^f)$, other quantities of interest can be calculated. For example, the distribution $D_{\mathbf{v}}(t^f + s|t^f)$ of interspike intervals is given by

$$D_{\mathbf{v}}(t^f + s|t^f) = -\frac{d}{ds} S_{\mathbf{v}}(t^f + s|t^f). \tag{33}$$

The mean duration of the interspike interval is

$$\langle s \rangle = \int_0^\infty ds\, s\, D_{\mathbf{v}}(t^f + s|t^f) = \int_0^\infty ds\, S_{\mathbf{v}}(t^f + s|t^f). \quad (34)$$

The mean firing rate ν is defined as the inverse of the mean interspike interval

$$\nu = \langle s \rangle^{-1}. \quad (35)$$

Our considerations here have concentrated on an *ensemble* of equivalent neurons which all have fired their last spike at the same time t^f. Note, however, that due to the renewal assumption a long spike train of a *single* neuron driven by a constant input exhibits the same interval statistics as an ensemble of neurons driven by the same input. Thus, for constant input, the mean firing rate defined in (35) can also be determined from the spike train of a single neuron by the standard procedure of counting the number of spikes and dividing by the measurement interval.

This closes our review of renewal theory. It will be shown below that the survivor function determines the dynamics in large pools of neurons. In this sense the *sole* purpose of specific neuron models is to calculate the quantity $S_{\mathbf{v}}(t^f + s|t^f)$.

B. Classification of pulse-coupled neuron models

The connection between the internal state $\mathbf{v}_i(t)$ and the synaptic input $\{S_j(t)\}$ is given by (31). We use this equation to classify the model networks with spiking neurons which have been introduced in Sec. I C.

First, we take the external input as a constant, $I_i^{\text{ext}}(t') \equiv I_0$. Furthermore, we assume that the functional derivative of $\mathbf{F}$ with respect to the synaptic input functions $S_j(t')$ exists. A Taylor-Volterra expansion of (31) in $S_j(t')$ yields

$$\mathbf{v}_i(t) = \eta_{I_0}(t - t_i^1) + \sum_{j \neq i} \int_0^\infty ds' J_{ij}\, \kappa^{(1)}_{I_0;ij}(t - t_i^1; s')\, S_j(t - s') + \sum_{j,k \neq i} \int_0^\infty ds' \int_0^\infty ds'' J_{ij} J_{ik} \kappa^{(2)}_{I_0;ijk}(t - t_i^1; s', s'')\, S_j(t - s') S_k(t - s'') + O(S_j S_k S_l). \quad (36)$$

For the sake of notational convenience we have omitted the specification $j \in \Gamma_i$ in the subscript of the sum. Also, we have split the "Taylor coefficients" into two factors, i.e., a response kernel κ which depends on the times s and s' and a prefactor which we interpret as the synaptic efficacy J_{ij}. In the limit of weak coupling ($J_{ij} \ll 1$), the linear term dominates the expansion in S_j. In the following we keep only terms to first order in J_{ij}.

1. Spike response model

As mentioned above, we work in the limit of weak coupling ($J_{ij} \ll 1$) and neglect all nonlinear terms. As a first specification we assume that the time course $\kappa^{(1)}$ is the same for all pairs of neurons. This allows us to omit the subscript ij. Since the external input current is kept constant, we also neglect the index I_0. Furthermore we restrict the model to a single variable $\mathbf{v}(t) = h(t)$ only. What remains is the most general one-dimensional linear model compatible with renewal theory. It is characterized by the synaptic response kernel $\kappa^{(1)}(s, s')$ and the free trajectory $\eta(s)$ where s and s' are the time after the last postsynaptic and presynaptic spike, respectively. If we identify $\eta(s)$ with the refractory function $\eta^{\text{refr}}(s)$ we have the spike response model as discussed in Sec. I C 4. In particular, if the synaptic response kernel is independent of the state of the postsynaptic neuron, $\kappa^{(1)}(s, s') = \epsilon(s')$, we can give a simple interpretation of the response kernel, i.e., it describes the excitatory or inhibitory postsynpatic potential evoked by an incoming spike; cf. Fig. 1. This is the standard form of the spike response model which has been studied earlier [58,59,67–71]. Most of our discussion in Sec. III concerns the spike response model with arbi-

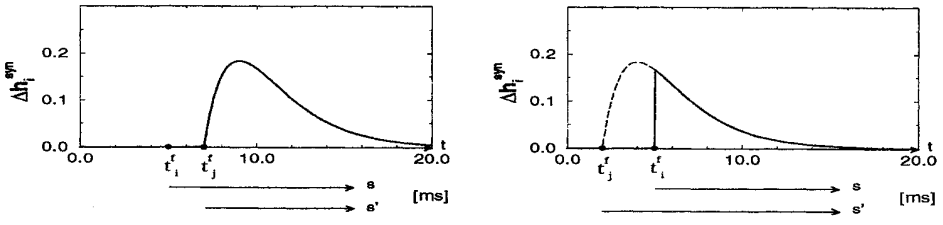

FIG. 1. Response kernel of the spike response model. The figure shows the change $\Delta h_i^{\text{syn}}(t)$ of the synaptic potential of neuron i in response to a single spike of a presynaptic neuron j. Neuron j fires at time t_j^f and the firing of i occurs at t_i^f. We define $s' = t - t_j^f$ and $s = t - t_i^f$. Here we assume that the synaptic response kernel has the standard form $\kappa(s, s') = \epsilon(s')$ with $\epsilon(s') = (s'/\tau_s^2)\exp(-s'/\tau_s)\theta(s')$ and $\tau_s = 2$ ms. In this case $\Delta h_i^{\text{syn}}(t) = \epsilon(t - t_j^f)$, independent of s; left: $s < s'$; right: $s \langle s'$ (dashed line). The response kernel is defined for $s > 0$ only (solid line).

trary synaptic response kernel $\kappa^{(1)}(s,s')$ and arbitrary refractory function $\eta^{\rm refr}(s) \equiv \eta(s)$.

2. *Integrate-and-fire models*

The standard integrate-and-fire model (10) is a special case of the spike response model. In order to compare the two models, we introduce $J_{ij} = R K_{ij}$. For the refractory function, we take

$$\eta(s) = R I_0 \, [1 - \exp(-s/\tau_1)]. \tag{37}$$

The synaptic response function depends on the version of the integrate-and-fire model under consideration. It is

$$\kappa^{(1)}(s,s') = \frac{\theta(s')}{\tau_1} \exp\left(-\frac{s'}{\tau_1}\right) \theta(s-s') \tag{38}$$

for the one-step model and

$$\kappa^{(1)}(s,s') = \frac{\theta(s')}{\tau_1 - \tau_2} \left\{ \left[\exp\left(-\frac{s'}{\tau_1}\right) - \exp\left(-\frac{s'}{\tau_2}\right) \right] \times \theta(s-s') + \exp\left(-\frac{s'}{\tau_2}\right) \times \left[\exp\left(-\frac{s}{\tau_1} + \frac{s}{\tau_2}\right) - 1 \right] \theta(s'-s) \right\}. \tag{39}$$

for the two-step model. The three-step model with $\tau_1 = \tau_2 = \tau_3$ yields

$$\kappa^{(1)}(s,s') = \frac{\theta(s')}{2\tau_1^3} \left\{ s'^2 \exp\left(-\frac{s'}{\tau_1}\right) \theta(s-s') + (2s\,s' - s^2) \exp\left(-\frac{s'}{\tau_1}\right) \theta(s'-s) \right\}. \tag{40}$$

The θ functons in (37)–(40) take care of causality and the fact that the membrane potential is reset at $s = 0$. This is most easily seen in the one-step version. Here only those spikes contribute to the synaptic potential for $s > 0$ which have been fired *after* the last spike of the postsynaptic neuron at $s = 0$. Thus we have a factor $\theta(s - s')$. Similarly, in the two-step and three-step version, only the portion on the synpatic input current which arrives after the last postsynaptic spike contributes; cf. Fig. 2. In Sec. III, solutions for arbitrary $\kappa^{(1)}(s,s')$ are given. With the special form (38)–(40), we get the solutions of the integrate-and-fire model in the standard form (10). For the generalized form (14) of the integrate-and-fire model, however, we have to use different kernels which we discuss now.

In order to relate the "phase model" formulation, (15) and (16), to the previous approach, we have to calculate the response of the neuron to a single incoming spike $I_i(t) = K_{ij}\,\delta(t - t_j^f)$. According to (15), a unit which receives no spikes has a phase evolution $\psi_i^0(t) = t$. On the other hand, a unit which receives a spike at time t_j^f suffers a phase shift $\Delta\psi \equiv \psi_i' - \psi_i = u^{-1}[V'] - u^{-1}[V_0]$ where $V_0 = V_i(t_j^f) = u(\psi_i)$ is the value of the variable V immediately before firing and V' the new value after firing; cf. (17)–(19). Consequently, its phase evolution is $\psi_i^0(t) + \Delta\psi\,\theta(t - t_j^f)$. Note that the phase shift $\Delta\psi$ in response to a single spike depends on the value of the phase variable immediately before the spike arrives. If several spikes have arrived in the time interval $t_i^1 \le t' \le t_j^f$, then the phase ψ_i has been shifted repeatedly. It follows that the shift $\Delta\psi$ due to the spike at t_j^f is a function of the input $I(t')$ for $t_i^1 \le t' \le t_j^f$. Thus, even though (15) looks linear in I, the phase has an implicit nonlinear dependence upon the input and a linear response kernel is not sufficient to describe the dynamic evolution. We conclude that, in general, phase models are in a model class different from the spike response model.

In the limit of weak coupling ($K_{ij} \ll 1$), however, we can expand ψ in (15) in powers of I. With $\psi(t) = t + \Delta_1(t) + \Delta_2(t) + \cdots$ we find

$$\Delta_1(t) = \int_0^t dt'\,\gamma(t')I(t'), \tag{41}$$

$$\Delta_2(t) = \int_0^t dt' \left(\frac{d}{ds'}\gamma(s)|_{t'}\right) I(t') \int_0^{t'} dt''\,\gamma(t'')I(t''). \tag{42}$$

Equation (41) allows us to calculate the synaptic response kernel $\kappa^{(1)}(s,s')$ as will be shown now.

3. *Mirollo-Strogatz model*

Specifically, let us consider the linear response to an input $I_i(t) = K_{ij}\delta(t - t_j^1 - \Delta^{\rm ax})$ where t_j^1 is the last spike of a presynaptic neuron j and $\Delta^{\rm ax}$ is the signal transmission time from j to i. We make the identification $K_{ij} = J_{ij}/\tau$ and introduce $s = t - t_i^1$ and $s' = t - t_j^1$. We also take care of the fact that the phase has been reset to 0 at t_i^1, the last firing of neuron i. From (41) we find the linear response kernel

$$\kappa^{(1)}(s,s') = \theta(s' - \Delta^{\rm ax})\,\frac{\gamma(s - s' + \Delta^{\rm ax})}{\tau}\,\theta(s - s' + \Delta^{\rm ax}). \tag{43}$$

Thus, for low input level, the phase model in standard form I, cf. (15), corresponds to a spike response model with $\eta(s) = s$ and $\kappa^{(1)}(s,s')$ given by (43); cf. Fig. 3. With the relation $h(t) = u[\psi(t)]$ and $\psi(t) = t + \Delta_1(t)$ we can transform our result into the standard form II; cf. (16). This yields a linear response kernel

$$\kappa^{(1)}(s,s') = \theta(s' - \Delta^{\rm ax})\,\frac{\gamma(s - s' + \Delta^{\rm ax})}{\tau\,\gamma(s)}\,\theta(s - s' + \Delta^{\rm ax}) \tag{44}$$

and the free trajectory $\eta(s) = u(s)$.

As a special case let us consider $u(s) = R I_0\,[1 - \exp(-s/\tau_1)]$. This is again the trajectory of the standard integrate-and-fire model (10). Taking the derivative of u, we find $\gamma(s) = 1/u'(s) = (\tau_1/R I_0)\exp(s/\tau_1)$. Equation

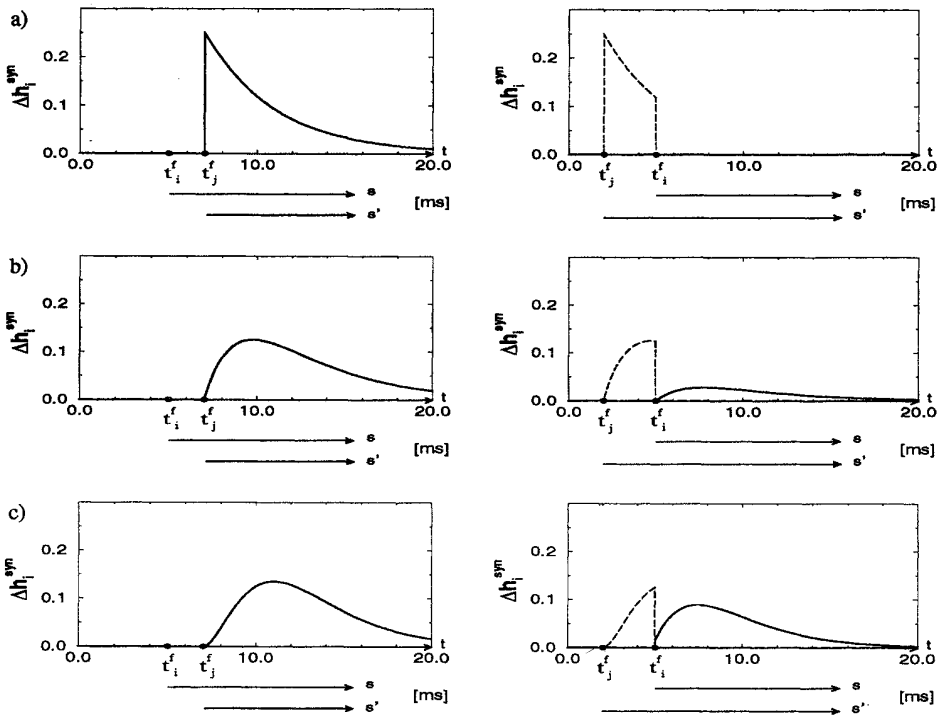

FIG. 2. Response kernel of the integrate-and-fire model. As before, we consider a single firing event of a presynaptic neuron j at time t_j^f and plot the change $\Delta h_i^{\rm syn}$ in response to the firing. For $s > 0$ (or $t > t_i^f$), the change $\Delta h_i^{\rm syn}$ is given by the response kernel, namely, $\Delta h_i^{\rm syn}(t) = \kappa(t-t_i^f, t-t_j^f)$ (thick lines). The response for $t_j^f < t < t_i^f$ is indicated by dashed lines. Left column: Neuron j fires later than i ($s < s'$). The postsynaptic neuron experiences the full effect of the presynaptic spike. Right column: Neuron i fires later than j ($s < s'$). Due to the reset of the potential at time t_i^f, the postsynaptic neuron experiences only a partial input for $s > 0$ (or none at all). (a) One-step integrate-and-fire model: The response is an exponentially decaying pulse for $s > s'$ and vanishes for $s < s'$ ($\tau_1 = 4$ ms). (b) Two-step integrate-and-fire model: The response is a smooth function of time, but it is much smaller for $s < s'$ ($\tau_1 = 2$ ms, $\tau_2 = 4$ ms). (c) Three-step integrate-and-fire model: The response starts even smoother, otherwise it is similar to case (b) ($\tau_1 = \tau_2 = \tau_3 = 2$ ms).

(44) leads to the response kernel

$$\kappa^{(1)}(s,s') = \frac{\theta(s'-\Delta^{\rm ax})}{\tau} \exp\left(-\frac{s'-\Delta^{\rm ax}}{\tau_1}\right) \times\theta(s-s'+\Delta^{\rm ax}). \quad (45)$$

In the case of delayless interaction $\Delta^{\rm ax} = 0$, Eq. (45) is identical to the response kernel which we have found earlier (38). It is the exact kernel and not a weak-coupling approximation, since (10) describes a linear model.

C. Dynamic mean-field theory

In Sec. II A we have considered an ensemble of independent neurons which receive sequences of spikes from a different set of neurons. Now we turn to networks. In this case, the input to a given neuron consists of the spikes of other neurons in the same net. In order to derive dynamic equations for the macroscopic behavior of the system, we have to restrict the network structure.

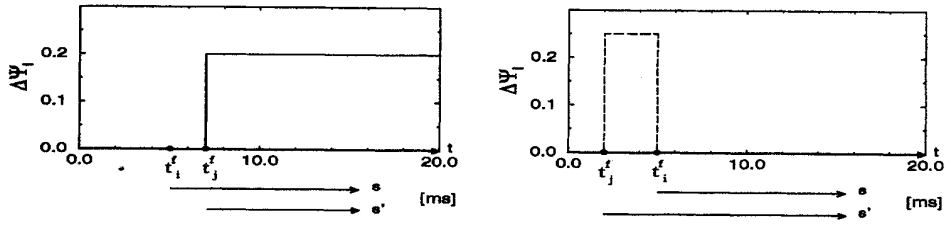

FIG. 3. Response kernel of the Mirollo-Strogatz model. In the "phase" description (standard form I), the response kernel is a θ function if $s > s'$ (left) and vanishes for $s < s'$ (right). In the "voltage" description, the response kernel is similar to the one-step integrate-and-fire model; see Fig. 2.

1. The pool concept

Real neural systems have a high degree of connectivity. In the mammalian cortex, for example, each neuron is connected to approximately 10^4 other neurons and connections can span a considerable distance (see, e.g., [27,110]). Thus, in contrast to most standard models of statistical or solid state physics, an ansatz with nearest neighbor interaction is a bad approximation to real cortical connectivity. Rather we should consider a fully connected system as a first approximation. This naturally leads to an approach by mean-field theory which is exact in the limit of a large and fully connected system.

In the following we study a network of N neurons in the limit of $N \to \infty$. We assume extensively many connections, i.e., every neuron is connected to a finite fraction of all other neurons. Furthermore, we assume that the network consists of a finite number of pools $L(\mathbf{x})$ where $\mathbf{x} \in \mathbf{K}$ with $\mathbf{K}$ to be specified.

The concept of pools and the related concept of assemblies have been applied repeatedly to neural systems, both in biological considerations (for a review see [111]) and in mathematical theories (see, e.g., [96,112]). In the biological literature, pools or assemblies are sometimes defined by common input (e.g., similar receptive fields), by common output (e.g., targeting the same region of the cortex), or by a common task (e.g., dealing with the same object). In the following we avoid the term "assembly" which is multifaceted and prone to misunderstandings. We stick to the simple term "pool," take the most restrictive approach, and assume that pools have both common input *and* output connections. More precisely, all neurons of a given pool are equivalent in the following sense. For any neuron $i \in L(\mathbf{x})$ and an arbitrary neuron $j \in L(\mathbf{y})$ the interaction strength J_{ij} and time course $\boldsymbol{\kappa}^{(1)}$ depend on $\mathbf{x}$ and $\mathbf{y}$ only, i.e.,

$$J_{ij} = |L(\mathbf{y})|^{-1} J(\mathbf{x},\mathbf{y}), \quad \text{for } i \in L(\mathbf{x}), j \in L(\mathbf{y}),$$
$$\boldsymbol{\kappa}^{(1)}_{ij}(s,s') = \boldsymbol{\kappa}^{(1)}(\mathbf{x},\mathbf{y};s,s') \quad \text{for } i \in L(\mathbf{x}), j \in L(\mathbf{y}). \tag{46}$$

The normalization $|L(\mathbf{y})|$ denotes the total number of neurons in pool $L(\mathbf{y})$. If we allow nonlinear terms in (36) we require similar relations for $\boldsymbol{\kappa}^{(2)}_{ijk}$. Equation (46) is the central assumption of the pool concept. Equivalent or similar definitions have been used by a number of authors [51,52,58,59,61,71,96,112–114]. For the sake of simplicity, we assume throughout the following $\boldsymbol{\kappa}^{(1)}(\mathbf{x},\mathbf{y};s,s') = \boldsymbol{\kappa}^{(1)}(s,s')$, independent of the pool indices $\mathbf{x}$ and $\mathbf{y}$.

To be specific, we discuss two examples. First, let us consider a two-dimensional sheet of cortex subdivided in columns. According to the concept of Hubel and Wiesel [115,116], all neurons in a given column have similar receptive fields characterized by the location $\mathbf{r} \in \mathbf{R}^2$ in the visual field (or on the retina) and, possibly, some additional parameter such as the direction $\theta \in [0, 2\pi)$ of a moving bar. Both the receptive field concept and the specific spatial organization suggested by Hubel and Wiesel have repeatedly been questioned; see, e.g., [31,117]. Here, we take the columnar structure as a theoretical assumption in an even more restrictive formulation. In order to apply the pool concept we have to require that all neurons in a column have *identical* receptive fields *and* connections. Then each column can be labeled by $\mathbf{x} = (\mathbf{r},\theta) \in \mathbf{K} \equiv \mathbf{R}^2 \times [0, 2\pi)$ and an obvious ansatz for the interaction between $i \in L(\mathbf{x})$ and $j \in L(\mathbf{y})$ is

$$J_{ij} = J_1(\mathbf{r},\mathbf{r}') J_2(\theta - \theta'). \tag{47}$$

In the pool concept, all details of the real cortical connectivity are neglected. The result is a network structure which describes some basic features of the "mean cortical anatomy." In particular, we can take care of the mean connectivity between different areas and layers of the cortex with a similar ansatz [118]. Instead of adding more details to (47), we can also simplify even further and neglect the angle dependence. In this case, we can identify the column $L(\mathbf{x})$ with neurons in the neighborhood of a location $\mathbf{x} \in \mathbf{R}^2$ on the two-dimensional sheet of cortex. In this interpretation, an ansatz $J_{ij} = J(\mathbf{x} - \mathbf{y})$ yields the spatial connectivity from neurons near point $\mathbf{y}$ to neurons near $\mathbf{x}$. This can be used to derive spatially continuous modes as studied by Wilson and Cowan [119], Feldman and Cowan [97], and others. For a derivation of the spatially continuous formulation, see Sec. III A.

We now turn to a second example of the pool concept. Let us consider an associative network which has to learn a finite number of patterns $\{\xi^\mu\} = \{\xi^\mu_i = \pm 1; 1 \le i \le N\}$ with $1 \le \mu \le q$. A fairly general Hebb rule yields synaptic efficacies of the form [120]

$$J_{ij} = \sum_{\mu=1}^{q} f_1(\xi^\mu_i) f_2(\xi^\mu_j). \tag{48}$$

Examples include the asymmetric coupling $J_{ij} \propto \sum_{\mu=1}^{q} \xi^\mu_i (\xi^\mu_j - a)$ or the symmetric coupling $\propto (\xi^\mu_i - a)(\xi^\mu_j - a)$ which have both been used for low activity random patterns with $a = \lim_{N\to\infty} (N q)^{-1} \sum_{j=1}^{N} \sum_{\mu}^{q} \xi^\mu_j < 0$. For $a = 0$ we have the coupling of the Hopfield model [4]. We focus on the "information vector" $\boldsymbol{\xi}_i = (\xi^1_i, \ldots, \xi^q_i)$ which neuron i has to store. For q patterns, 2^q different information vectors. If we keep q fixed while taking $N \to \infty$, there are many neurons which deal with the same information vector. This allows us to introduce *sublattices* [113,114,121,122] or pools of equivalent neurons by the definition $L(\mathbf{x}) = \{i | \boldsymbol{\xi}_i = \mathbf{x}\}$. The interaction (48) for $i \in L(\mathbf{x})$ and $j \in L(\mathbf{y})$ reduces to $J_{ij} = J(\mathbf{x},\mathbf{y}) = \sum_{\mu=1}^{q} f_1(x^\mu) f_2(y^\mu)$ and we are back to the pool interaction (46).

2. Pool dynamics

In this section we derive dynamic equations for the pool activity

$$A(\mathbf{x},t) = \lim_{\Delta t \to 0} \lim_{N\to\infty} |L(\mathbf{x})|^{-1} \times \sum_{j \in L(\mathbf{x})} \frac{1}{\Delta t} \int_t^{t+\Delta t} dt'\, S^{(F)}_j(t'). \tag{49}$$

We can use this definition in Eq. (36) and find for a neuron $i \in L(\mathbf{x})$

$$\mathbf{v}(\mathbf{x},s,t) = \eta(s) + \sum_{\mathbf{y}\in\mathbf{K}} \int_0^\infty ds'\, J(\mathbf{x},\mathbf{y})\kappa^{(1)}(\mathbf{x},\mathbf{y};s,s')A(\mathbf{y},t-s') + \sum_{\mathbf{y},\mathbf{z}\in\mathbf{K}} \int_0^\infty ds' \int_0^\infty ds''\, J(\mathbf{x},\mathbf{y})\, J(\mathbf{x},\mathbf{z})\kappa^{(2)}(\mathbf{x},\mathbf{y},\mathbf{z};s,s',s'')\, A(\mathbf{y},t-s')\, A(\mathbf{z},t-s'') + O(A^3). \quad (50)$$

The sum runs over all pools including $\mathbf{x}$ since two different neurons $i,j \in L(\mathbf{x})$ indeed interact with $J_{ij} = J(\mathbf{x},\mathbf{x})$. In (50) we have also set $J_{ii} = J(\mathbf{x},\mathbf{x})$, introducing an error of order $|L(\mathbf{x})|^{-1}$ which vanishes for $N \to \infty$.

We note that all neurons $j \in L(\mathbf{x})$ which have fired their last spike at the same time, i.e., $t_j^1 = t^f$, have the same internal state; cf. (50). According to (49), the portion of those neurons is $A(\mathbf{x},t^f)$. The probability that one of these neurons fires its next spike at $t^f + s$ is given by $D_{\mathbf{v}}(t^f + s|t^f)$; cf. (33). Thus neurons which have fired their last spike between t^f and $t^f + ds$ make a contribution $D_{\mathbf{v}}(t^f + s|t^f)A(\mathbf{x},t^f)\,ds$ to the activity of pool $L(\mathbf{x})$ at time $t^f + s$. If we keep $t = t^f + s$ fixed and integrate over s we find

$$A(\mathbf{x},t) = \int_0^\infty ds\, D_{\mathbf{v}}(t|t-s)\, A(\mathbf{x},t-s). \quad (51)$$

This is an integral equation which describes the evolution of the pool activity $A(\mathbf{x},t)$. Note that (51) is invariant under a rescaling of the activity $A(\mathbf{x},t) \longrightarrow CA(\mathbf{x},t)$. An additional equation that yields a normalization condition for $A(\mathbf{x},t)$ can be found as follows.

Let us consider the portion of neurons in $L(\mathbf{x})$ which have fired their last spike at $t-s$ and have stayed quiescent from $t-s$ to t. With (32) this portion is $S_{\mathbf{v}}(t|t-s)A(\mathbf{x},t-s)ds$. If we look at the system at time t, every neuron must have fired its last spike at *some* point back in time. Integration over time yields the normalization condition

$$1 = \int_0^\infty ds\, S_{\mathbf{v}}(t|t-s)\, A(\mathbf{x},t-s). \quad (52)$$

Division of (51) by (52) yields

$$A(\mathbf{x},t) = \frac{\int_0^\infty ds\, D_{\mathbf{v}}(t|t-s)\, A(\mathbf{x},t-s)}{\int_0^\infty ds\, S_{\mathbf{v}}(t|t-s)\, A(\mathbf{x},t-s)}. \quad (53)$$

This is the central equation which will be used throughout the rest of this paper.

For a discussion of the equation we have to keep in mind that both the interval distribution and the survivor function depend on the internal state $\mathbf{v}(\mathbf{x};s,t')$ during the interval $t-s \le t' \le t$. The internal state $\mathbf{v}(\mathbf{x};s,t')$ in turn is dependent upon the activity of all pools $A(\mathbf{y},t'')$ for $t'' \le t'$, including the pool $\mathbf{y} = \mathbf{x}$; cf. (50). If the connection between the survivor function and the internal state is known, we have a closed set of equations. The missing link is supplied by specific network models as discussed in Sec. III.

The combination of (53) with (50) yields a solution of the third problem outlined in Sec. I B, namely, it gives a systematic link between the single neuron dynamics and the effective pool dynamics. Our derivation shows that this connection can be found in the rather abstract setting of renewal theory. Specific examples will be discussed in Sec. III. In particular, the relation of this result with the model equations of Wilson and Cowan will be explained in Sec. III F.

We close this section with a final remark regarding the relevance of the renewal assumption (Sec. II A) for the derivation of Eqs. (51) and (52). Based on the renewal hypothesis, we have been able to classify neurons by their *last* firing time t^f and by their pool index $\mathbf{x}$ and we have concentrated on the activity $A(\mathbf{x},t^f)$. In order to make the relation between Eq. (51) and renewal theory a little bit clearer, let us assume for the moment that the internal state $\mathbf{v}(t)$ depends not only on the last firing but also on the second last. In this case, the expression $D_{\mathbf{v}}(t|t-s)A(\mathbf{x},t-s)ds$ in (51) is not well defined, since the interval distribution $D_{\mathbf{v}}$ depends on $\mathbf{v}$ and the internal state $\mathbf{v}$ may and, in general, does vary within the group of neurons contributing to $A(\mathbf{x},t-s)ds$. We can take care of this problem, if we introduce the time s_2 which has passed since the second last spike as an additional parameter and work with the generalized quantities $D_{\mathbf{v}}(t|t-s,t-s_2)$ and $A(\mathbf{x},t-s,t-s_2)$. Similarly we can generalize our approach to the situation where the last k spikes are relevant [70]. Here, we do not want to go into detail but we return to the renewal case and apply Eq. (53) to specific network models.

III. APPLICATION TO SPECIFIC NETWORK MODELS

Throughout the following we assume that the internal state (50) has a linear dependence upon the input only. In other words we neglect terms of second and higher order in A. We also restrict our discussion to models with a single internal variable $\mathbf{v}(t) = h(t)$ only. For the sake of simplicity, we make the additional assumption $\kappa^{(1)}(\mathbf{x},\mathbf{y};s,s') \equiv \kappa(s,s')$, independent of the pool indices $\mathbf{x}$ and $\mathbf{y}$. Equation (50) is then reduced to

$$h(\mathbf{x};s,t) = \eta(s) + h^{\mathrm{syn}}(\mathbf{x},s,t), \quad (54)$$

with

$$h^{\mathrm{syn}}(\mathbf{x},s,t) = \sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y}) \int_0^\infty ds'\kappa(s,s')A(\mathbf{y},t-s'). \quad (55)$$

The first term in (54) describes the internal dynamics, the second term is the total input h^{syn}. We now turn to the specific network models of Sec. I C.

A. Spatially continuous networks

We remind the reader that $\mathbf{x}$ in (54) and (55) is a formal vector index denoting neurons in pool $L(\mathbf{x})$. In the simplified columnar interpretation (neglecting any angle dependence) as sketched in Sec. II C 1, however, we can identify $\mathbf{x} \in \mathbf{R}^2$ with a location on the two-dimensional sheet of cortex and $L(\mathbf{x})$ with the ensemble of neurons in the neighborhood of $\mathbf{x}$. Instead of the index i, neurons can now be labeled by their coordinate $\mathbf{x}_i = (x_{i,1}, x_{i,2})$. A pool $L(\mathbf{x})$ with $\mathbf{x} = (x_1, x_2)$ is defined by

$$L(\mathbf{x}) = \{i \,|\, x_1 \le x_{i,1} \le x_1 + \epsilon; x_2 \le x_{i,2} \le x_2 + \epsilon\}. \quad (56)$$

In a homogeneous, two-dimensional network, the number of neurons in $L(\mathbf{x})$ is proportional to ϵ^2, thus

$$|L(\mathbf{x})| = \rho\,\epsilon^2, \quad (57)$$

where ρ is the spatial densitiy of the neurons. It is convenient to renormalize the weights (46) and introduce scaled connections

$$J(\mathbf{x},\mathbf{y}) = \epsilon^2 \hat{J}(\mathbf{x},\mathbf{y}). \quad (58)$$

With these definitions, we find from (55)

$$h^{\mathrm{syn}}(\mathbf{x},s,t) = \sum_{\mathbf{y}} \epsilon^2 \hat{J}(\mathbf{x},\mathbf{y}) \times \int_0^\infty ds' \kappa(s,s') A(\mathbf{y}, t-s'). \quad (59)$$

Taking the continuum limes $\epsilon \to 0$, we find

$$h^{\mathrm{syn}}(\mathbf{x},s,t) = \int d\mathbf{y} \hat{J}(\mathbf{x},\mathbf{y}) \times \int_0^\infty ds' \kappa(s,s') A(\mathbf{y}, t-s'). \quad (60)$$

A spatially continuous formulation has been used by a large number of researchers. von Seelen [123] has studied linear field theories of homogeneous continuous networks. A linear network including specific details of the mean cortical anatomy has been studied by Krone *et al.* [118]. Wilson and Cowan [119] have applied their population equations to the continuous case. Feldman and Cowan [97], Ermentrout and Cowan [124,125], and others have combined the spatially continuous formulation with the graded-response model (5) and studied spatiotemporal activation patterns. Recently, Idiart and Abbott [126] have analyzed the propagation velocity of excitation over a continuous net of graded-response neurons. In principle, the spatially continuous formulation can be combined with any of the model neurons introduced in Sec. I C. In the following subsections, these model neurons will be reconsidered from the viewpoint of the general theory developed in Sec. II.

B. Firing rate models

In order to connect the firing rate model, Sec. I C 1, with the theoretical framework of Sec. II, we assume constant pool activities, i.e., $A(\mathbf{x},t) \equiv A(\mathbf{x})$. Using $\int_0^\infty ds D_h(t|t-s) = 1$ and (34) we find from (53)

$$A(\mathbf{x}) = \left[\int_0^\infty ds\, S_h(t|t-s)\right]^{-1} = \langle s \rangle^{-1}. \quad (61)$$

The survivor function $S_h(t|t-s)$ in (61) has to be evaluated with the internal variables

$$h(\mathbf{x};s,t) = \eta(s) + \overline{\kappa}(s) \sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y}) A(\mathbf{y}), \quad (62)$$

where we have introduced $\overline{\kappa}(s) = \int_0^\infty ds' \kappa(s,s')$. Equations (61) and (62) are the first result of this section and will now be discussed from a different point of view.

Let us consider a *single* neuron driven by a constant input $h^{\mathrm{in}} = \sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y}) A(\mathbf{y})$. After a spike at a time $t^f = t - s$, its internal state at time t will be given by (62). Therefore it behaves identically to the neurons in pool $L(\mathbf{x})$. By definition, the firing rate of a neuron driven by a constant input h^{in} can be found from the gain function, that is, $\nu = g(h^{\mathrm{in}})$; cf., (1). On the other hand, the mean firing rate is the inverse of the mean interspike interval [17], i.e., $\nu = \langle s \rangle^{-1}$; cf. Eq. (35). Thus Eqs. (61) and (62) can be reduced to a single fixed-point equation

$$A(\mathbf{x}) = g\left(\sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y}) A(\mathbf{y})\right) \quad (63)$$

and we are back to the firing rate model. Our derivation shows that a rate description is valid for the *stationary* activity in large pools of neurons.

So far, the synaptic reponse kernel $\kappa(s,s')$ need not be specified. We now turn to a rather specific response function

$$\kappa(s,s') = \frac{1}{\tau} \exp\left(-\frac{s'}{\tau}\right) \theta(s'). \quad (64)$$

In particular, we have $\kappa(s,s') = \epsilon(s')$ independent of the time s. The derivative of the synaptic potential (55) with respect to t is

$$\tau \frac{d}{dt} h^{\mathrm{syn}} = -h^{\mathrm{syn}} + \sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y}) A(\mathbf{y},t), \quad (65)$$

where we have used (64). If we now replace $A(\mathbf{y},t)$ by $g[h^{\mathrm{syn}}(\mathbf{y},t)]$, as suggested by (63), then (65) becomes identical to the standard equation of the graded-response model; cf. Eq. (5). This change is possible under *stationary* conditions; cf. Eq. (63). We remark that this procedure suggests the pool interpretation (6) with the ensemble averaged rate A rather than the single neuron interpretation (4) with a time averaged mean firing rate ν.

Let us find the error introduced when replacing $A(\mathbf{y},t)$ by $g[h^{\mathrm{syn}}(\mathbf{y},t)]$. To this end we use our dynamic equation (53) and approximate

$$A(t-s) = A(t) - A'|_t s + \frac{1}{2} A''|_t s^2 - \cdots \quad (66)$$

neglecting all higher derivatives. This yields

$$A(t) = \langle s\rangle^{-1}\left[1 - \frac{A'|_t\langle s\rangle}{A}\left(1 - \frac{\langle s^2\rangle}{2\langle s\rangle^2}\right) + \cdots\right], \quad (67)$$

where

$$\langle s\rangle = \int_0^\infty ds\, s\, D_h(t|t-s) \quad (68)$$

and

$$\langle s^2\rangle = \int_0^\infty ds\, s^2\, D_h(t|t-s) \quad (69)$$

are the first and second moments of the interval distribution. Note that the interval distribution depends on the time course of $h^{\rm syn}(t')$ for $t - s \le t' \le t$. If we evaluate $D_h(t|t-s)$ for $h^{\rm syn}(t') \approx h^{\rm syn}(t)$, then the first moment is the mean interspike interval at input $h^{\rm syn}(t)$. As before, we have $\langle s\rangle^{-1} = g[h^{\rm syn}(t)]$. The expression in the square brackets in Eq. (67) yields a correction to the graded-response model which is small only if $A'|_t\langle s\rangle \ll A$. It will be shown in the following sections that $A(t)$ may change rapidly during one period $\langle s\rangle$. In particular, during collective oscillations, $A'|_t\langle s\rangle/A$ is of order one or larger. Thus the approximation $A(\mathbf{x},t) = g[h^{\rm syn}(\mathbf{x},t)]$ is no longer justified. Feldman and Cowan [97], who derived graded-response equations 20 years ago by a procedure of "time coarse graining," have been aware of the fact that fast transients and oscillations in A are neglected. Others have used the graded-response ansatz with a qualitative argument claiming the "quasistationarity" of the membrane potential h. Our derivation of Eq. (67) shows that an explicit estimation of the error due to "time coarse graining" or "quasistationarity" is possible. In fact, the assumption of quasistationarity is a poor approximation during transients or collective oscillations.

Our present results can be summarized as follows.

(i) If we have a large neural system and if we know that it is in a *stationary* state with *constant pool activities*, then the time average and the pool average are equivalent and the system is fully determined by the gain function of the neurons. In this case, higher moments of the interval distribution are not important. This result is independent of specific model assumptions [127].

(ii) The graded-response model can be derived from the macroscopic pool dynamics (53), if the synaptic response function is a simple exponential and if we can evaluate the dynamic equation for quasistationary h. More precisely, the graded-response equations are valid, if $A(t)$ changes only slowly during a typical interspike interval $\langle s\rangle$. In general, this is not the case even though $h^{\rm syn}$ may change only slowly. Thus the graded-response model cannot provide a proper dynamic description of neural activity. In particular, fast transients and oscillations are neglected. It may be used, however, as a simple dynamics to find the stationary states in firing rate models.

C. Spike response model

If we identify the internal contribution in (54) with the refractory potential $\eta^{\rm refr}(s) = \eta(s)$ and allow an arbitrary synaptic response kernel $\kappa(s,s')$ we have the generalized version of the spike response model as defined in Sec. I C 4. The standard formulation of the spike response model is found for $\kappa(s,s') = \epsilon(s')$ independent of the state s of the postsynaptic neuron. In order to keep the arguments general, we allow throughout the following an arbitrary response $\kappa(s,s')$.

Using the definition of the noise, (26), we can calculate the survivor function $S_h(t^f + s|t^f)$ that enters the evolution equation (53). Let us consider all neurons in $L(\mathbf{x})$ that have fired at $t = t^f$ and have been quiescent thereafter. At a later time $t = t^f + s$, all of these neurons receive the same input $h^{\rm syn}(\mathbf{x},s,t) = \sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y}) \int_0^\infty \kappa(s,s')A(\mathbf{y},t-s')ds'$ and they all have the same refractory potential $\eta(s)$. The portion of neurons that remains quiescent decays due to firing with probability $P_F[s, h(t_f+s); \delta t]$ where $h(t^f+s) = h^{\rm syn}(\mathbf{x},s,t^f + s) + \eta(s)$; cf. Eq. (26). Thus the surviving portion is

$$S_h(t^f + s|t^f) = \exp\left\{-\int_0^s ds'\tau^{-1}[s', h(t^f + s')]\right\}. \quad (70)$$

If the survivor function (70) is used in (53) we find

$$A(\mathbf{x},t) = \frac{\int_0^\infty ds\,\tau^{-1}[s,h(\mathbf{x},s,t)]S_h(t|t-s)A(\mathbf{x},t-s)}{\int_0^\infty ds\, S_h(t|t-s)A(\mathbf{x},t-s)}. \quad (71)$$

If we drop the explicit s dependence of τ^{-1}, (71) is identical to [59] since the denominator is normalized to unity; cf. Eq. (52). Equations (70) and (71) combined with the equation for the membrane potential (54) and (55) form a closed set. They describe the effective interaction within and between different pool activities $A(\mathbf{x},t)$ and $A(\mathbf{y},t)$. We note that the interaction is highly nonlinear due to (70) and involves three iterated integrations over time. Thus the transition from single neurons to pools of neurons results in a much more complicated dynamics.

In the case of stationary activity $A(\mathbf{x},t) \equiv A(\mathbf{x})$, Eq. (71) is reduced to a firing rate model

$$A(\mathbf{x}) = g[h^{\rm in}(\mathbf{x})], \quad (72)$$

with

$$h^{\rm in}(\mathbf{x}) = \sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y})\, A(\mathbf{y}). \quad (73)$$

In contrast to Eq. (63) of Sec. III B, we now have an explicit expression for the gain function, namely,

$$g(h) = \left\{\int_0^\infty ds \exp\left(-\int_0^s ds'\tau^{-1}[s', \overline{\kappa}(s')h + \eta(s')]\right)\right\}^{-1}, \quad (74)$$

where $\overline{\kappa}(s) = \int_0^\infty \kappa(s,s')ds'$.

Summarizing this subsection we have the following result. With the specific form of the noise (26), we can calculate the survivor function, the interval distribution, and the gain function of the model neurons. Insertion of the survivor function into (53) combined with (54) and (55) allows us to predict the network behavior. This is possible for arbitrary response kernels.

D. Collective states in noise-free networks with linear input

For the sake of convenience, we restrict our arguments throughout the following to the case of a homogeneous network. In other words, we have a single pool with internal interaction $J(\mathbf{x},\mathbf{x}) = J_0$. A generalization to several pools is straightforward. To be specific, we always assume $J_0 > 0$. In the case of a network with inhibitory coupling, the minus sign is included in the kernel $\kappa(s,s')$. We also concentrate on the noise-free limit $\beta \to \infty$. Some remarks on noisy systems are made in passing. All arguments apply for arbitrary $\kappa(s,s')$ unless stated otherwise. Specification of the response kernels $\kappa(s,s')$ and $\eta(s)$ yields different pulse-coupled models as discussed in Sec. II B.

In the noiseless case we take advantage of the fact that $\tau^{-1}(s,h) \to 0$ if $h < \theta$ and $\tau^{-1}(s,h) \to \infty$ if $h > \theta$, independently of s. Thus (71) is reduced to

$$A(\mathbf{x},t) = \frac{\int_0^\infty ds\, \delta[s - s_F(t-s)]A(\mathbf{x},t-s)}{\int_0^{s_F} ds'\, A(\mathbf{x},t-s')}, \tag{75}$$

where $s_F(t)$ is given by the threshold condition

$$s_F(t_f) = \inf\{s \mid h(\mathbf{x},s,t^f+s) = \theta\}. \tag{76}$$

As a first step we go back to the stationary state $A(t) \equiv A_0$. Equations (76) and (77) yield

$$A_0 = T^{-1}, \tag{77}$$

where the interspike interval T is given by the threshold condition

$$T = \inf\{s \mid \eta(s) + \overline{\kappa}(s)J_0A_0 = \theta\}. \tag{78}$$

1. *Stability of the stationary state*

We consider a small perturbation of the stationary state

$$A(t) = A_0 + A_1 \exp\{i\tilde{\omega}t\}, \tag{79}$$

with $\tilde{\omega} = \omega + i\lambda$ where $\lambda T \ll 1$. This induces an oscillation of the synaptic potential

$$h^{\text{syn}}(s,t) = h_0(s) + h_1(s)\exp\{i\tilde{\omega}t - i\alpha\}, \tag{80}$$

with $h_0(s) = \overline{\kappa}(s)J_0A_0$ and $h_1(s) = J_0A_1|\tilde{\kappa}(s,\tilde{\omega})|$, where

$$\tilde{\kappa}(s,\tilde{\omega}) \equiv |\tilde{\kappa}(s,\tilde{\omega})|e^{-i\alpha} = \int_0^\infty ds'\kappa(s,s')e^{-i\tilde{\omega}s'} \tag{81}$$

is the generalized Fourier transform of the synpatic response function evaluated at time s after the last postsynpatic spike. The phase α is understood to depend on $\tilde{\omega}$. The change in h^{syn} influences the interspike interval given by the threshold condition (76). We write $s_F(t_f) = T + v(t_f + T)$, where T is the interval in the unperturbed situation and expand (76) to first order in A_1. This yields a shift

$$v(t) = -\frac{\Delta h(t)}{\eta'|_T + h_0'|_T} \tag{82}$$

due to the perturbation $\Delta h(t) = J_0A_1|\tilde{\kappa}(T,\tilde{\omega})|\exp\{i\tilde{\omega}t - i\alpha\}$. With $\eta'|_T$ we have denoted the derivative $\frac{d}{ds}\eta(s)$ evaluated at time $s = T$ and, similarly, $h_0'|_T$ is defined as $\frac{d}{ds}h_0(s)$, also evaluated at $s = T$.

Linearization of (75) yields

$$A(t) - A(t-T) = -\frac{d}{dt}v(t)A_0 \tag{83}$$

and with the ansatz (79), (80) we find

$$1 - \exp\{-i\tilde{\omega}T\} = i\tilde{\omega}\frac{J_0A_0\tilde{\kappa}(T,\tilde{\omega})}{\eta'|_T + h_0'|_T}. \tag{84}$$

In most realistic models, the right-hand side of (84) is of small absolute value, i.e.,

$$b \equiv \left|\frac{\tilde{\omega}J_0A_0\tilde{\kappa}(T,\tilde{\omega})}{\eta'|_T + h_0'|_T}\right| \ll 1. \tag{85}$$

In order to keep the left-hand side small as well, we have to require

$$\tilde{\omega} = \omega_n + \nu_n + i\lambda_n, \tag{86}$$

with $\omega_n = n2\pi/T$ and $|\nu_n| \ll |\omega_n|$. The leading terms in an expansion of (84) are

$$(\nu_n + i\lambda_n)T = \frac{J_0A_0}{\eta'|_T + h_0'|_T}\omega_n\tilde{\kappa}(T,\omega_n) \equiv b_ne^{-i\alpha_n}, \tag{87}$$

where $b_n \in \mathbf{R}$ is an amplitude defined by (87) and α_n is the phase of the Fourier transform (81) evaluated at $s = T$ and $\tilde{\omega} = \omega_n$. Equation (87) is our final expression which allows a straightforward discussion of the stationary (incoherent) state.

First, the requirement $\lambda_n = 0$ yields the bifurcation condition $\alpha_n = 0$ and $\nu_n = b_n/T$ or $\alpha_n = \pi$ and $\nu_n = -b_n/T$. Furthermore, an oscillation grows maximally, if $\nu_n = 0$ and $\alpha_n = \pi/2$. In this case we have $\lambda_n^{max} = -b_n/T$. These arguments are summarized in the following criterion: The stationary state is *unstable* with respect to an oscillation with frequency $\omega \approx \omega_n$, if $0 \le \alpha_n \le \pi$.

Let us apply these general results to three specific examples. First we consider the response function $\kappa(s,s') = \epsilon(s')$ with $\epsilon(s')$ given by (24). This is the kernel that has been used in the standard version of the spike response

model; cf. Sec. I C 4. Taking the Fourier transform yields

$$|\tilde{\epsilon}(\omega)| = (1+\omega^2\tau_s^2)^{-1},$$
$$\alpha = \omega\Delta^{\rm ax} + 2\arctan(\omega\tau_s). \tag{88}$$

Changing the delay $\Delta^{\rm ax}$ we can adjust the phase α while keeping $|\tilde{\epsilon}(\omega)|$ fixed. Note that even a small change in $\Delta^{\rm ax}$ shifts the higher modes ($n \gg 1$) from stable to unstable regimes and vice versa. In general, for finite τ_s we expect that at least one of the modes is unstable, whatever the delay $\Delta^{\rm ax}$ [59]. Note also that $|\tilde{\epsilon}(\omega)| \to (\omega\tau_s)^{-2}$ for $\omega \to \infty$. Thus higher modes decay rather slowly if we are in a stable regime. For inhibitory interaction, we take $\kappa(s,s') = -\epsilon(s')$ with $\epsilon(s')$ given by (24). The only difference is a change of the phase α by π in (88). As a consequence, the stability pattern is inverted, that is, modes which grow ($\lambda_n < 0$) in the case of excitatory interaction decrease for inhibitory coupling and vice versa.

Abbott and van Vreeswijk [62] have studied the stability of asynchronous states in a network of three-step integrate-and-fire neurons. They find that the asynchronous state is stable if the time constants $\tau_2 = \tau_3$ are sufficiently large, but higher modes decay only slowly. In order to understand their results, let us consider as our second example the response function (40). This corresponds to the special case of a three-step integrate-and-fire model with $\tau_1 = \tau_2 = \tau_3 \equiv \tau$. The Fourier transform is

$$\tilde{\kappa}(T,\omega) = \frac{1-e^{-T/\tau}e^{-i\omega T}}{(1+i\omega\tau)^3}. \tag{89}$$

As before, we consider the case $\omega = \omega_n = 2\pi/T$ with $\omega_n\tau \gg 1$. The first two terms in an expansion in powers of $(\omega_n\tau)^{-1}$ are

$$\tilde{\kappa}(T,\omega_n) = \frac{1-e^{-T/\tau}}{(\omega_n\tau)^3}\left(i - \frac{3}{\omega_n\tau} + \cdots\right). \tag{90}$$

The leading term is purely imaginary and of positive absolute value. Comparison with (87) yields $\lambda_n > 0$. According to our definition of λ this is equivalent to the statement that oscillatory modes with $\omega_n\tau \gg 1$ decay. In particular, for τ sufficiently large, all oscillatory modes decay and the stationary state (incoherent activity) is stable; cf. [62]. Note, however, that $\lambda_n \ll 1$ for fast modes. Thus higher harmonics decay only slowly. Also, as before even a tiny delay makes one or more of the modes unstable. Thus zero delay is a singular case.

As a third and final example, let us consider the Mirollo-Strogatz model [56] in the limit of weak coupling. The Fourier transform of (44) is

$$\tilde{\kappa}(T,\omega) = \frac{1}{\tau}\int_{\Delta^{\rm ax}}^{T+\Delta^{\rm ax}} ds' \, \frac{\gamma(T-s'+\Delta^{\rm ax})}{\gamma(T)} e^{-i\omega s'}, \tag{91}$$

where $\Delta^{\rm ax}$ is some transmission delay. We specifiy γ and take $\gamma(t) = \tau\exp(t/\tau)$, which corresponds to the one-step integrate-and-fire model. This yields

$$\tilde{\kappa}(T,\omega) = \frac{1-e^{-(i\omega+\tau^{-1})T}}{1+i\omega\tau}e^{-i\omega\Delta^{\rm ax}}. \tag{92}$$

As before, we assume $\omega\tau \gg 1$ and evaluate at $\omega_n = n2\pi/T$. If we expand in $(\omega_n\tau)^{-1}$, we find

$$\tilde{\kappa}(T,\omega) = \frac{1-e^{-T/\tau-1}}{\omega_n\tau}\left(-i + \frac{1}{\omega_n\tau} + \cdots\right)e^{-i\omega_n\Delta^{\rm ax}}. \tag{93}$$

Let us discuss the case of $\Delta^{\rm ax} = 0$. In contrast to (90), the leading term in (93) is imaginary, but negative. Thus all modes with $\omega_n\tau > 1$ grow exponentially. As before, a delay $\Delta^{\rm ax} > 0$ changes the bifurcation pattern. For example, in the case of $\Delta^{\rm ax} = T/2$, all oscillatory modes with uneven n and $\omega_n\tau > 1$ decay, but the stationary state remains unstable with respect to modes $\omega_n\tau > 1$ with $n = 2, 4, \ldots$.

Summarizing the above considerations, we arrive at the following results.

(i) The stationary state of the noiseless system is almost always unstable and oscillations can build up.

(ii) In the regime where linear stability analysis holds, oscillations have a frequency $\omega_n \approx n2\pi/T$ where T is the interspike interval in the stationary state.

(iii) Whether an oscillation with frequency ω_n builds up or decays is determined by the phase α_n which can be adjusted via the delay $\Delta^{\rm ax}$ of the synaptic response.

(iv) The growth factor λ_n in the case of optimally adjusted phase α_n is determined by the coupling strength J_0 and the Fourier transform $|\tilde{\kappa}(T,\tilde{\omega})|$ of the synaptic response function. Since $\lim_{\omega\to\infty}|\tilde{\kappa}(T,\tilde{\omega})| = 0$, higher harmonics grow more slowly than the first mode ($n = 1$).

Finally, let us generalize our results to the case of finite noise level. We take $\kappa(s,s') = \epsilon(s')$ and expand Eq. (53) to first order in A_1. The requirement $\lambda = 0$ yields the bifurcation condition

$$1 - \tilde{D}_h(\omega) = h_0|\tilde{\epsilon}(\omega)|e^{-i\alpha}\frac{d}{dh}\tilde{D}_h(\omega), \tag{94}$$

where $\tilde{D}_h(\omega) = \int_0^\infty D_h(t|t-s)e^{-i\omega s}ds$ is the Fourier transform of the interval distribution in the unperturbed case. We note that not only the mean interspike interval $\langle s\rangle$ but also all higher moments of the interval distribution are important. For low noise level, we can approximate the interval distribution by a Gaussian with mean $\langle s\rangle$ and finite width $\langle\Delta s^2\rangle \ll \langle s\rangle^2$. In the case of noise, the Fourier transform $|\tilde{D}_h(\omega)|$ is smaller than 1 for all ω and the left-hand side of (94) cannot vanish. A glance at the right-hand side of (94) shows that oscillations can build up only if $|\tilde{\epsilon}(\omega)|$ is sufficiently large. Since $|\tilde{\epsilon}(\omega)| \to 0$ for $\omega \to \infty$, high frequency oscillations ($n \gg 1$) are suppressed. This leads to additional conclusions concerning noisy systems.

(iv) The stability of stationary states depends not only on the mean interspike interval $\langle s\rangle$, but on the Fourier transform of the full interval distribution.

(v) Fast oscillations (higher harmonics) are suppressed by noise.

2. *Oscillations and locking*

Let us consider a noiseless homogeneous system of spiking neurons during a collective oscillation. To be specific, we require that all neurons fire synchronously with period $T_{\rm osc}$, i.e.,

$$A(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_{\rm osc}). \tag{95}$$

The oscillation period is found self-consistently from the threshold condition (76)

$$T_{\rm osc} = \inf\left\{ s \,|\, \eta(s) + J_0 \sum_{n=0}^{\infty} \kappa(s, s + nT_{\rm osc}) = \theta \right\}. \tag{96}$$

To be specific, let us consider $\kappa(s,s') = \epsilon(s')$. In the case of $\theta = 0$ and $\sum_{n=2}^{\infty} \epsilon(nT_{\rm osc}) \ll \epsilon(T_{\rm osc})$ we can give a simple graphical interpretation of this result. The first intersection of the synaptic response $J_0\epsilon(s)$ with the negative refractory function $-\eta(s)$ yields the oscillation period $T_{\rm osc}$; cf. Fig. 4.

In passing we note that semicollective oscillations with roughly two or three times the basic frequency are also possible; cf. [128]. Consider a population which has spontaneously separated into two or, generally, k, subpopulations of equal strength. Such a separation could be induced by instabilities of the stationary state towards higher harmonics; cf. the preceding paragraph. Whereas every single neuron fires with period $T_{\rm osc}^{(k)}$, the collective activity now oscillates with period $T_{\rm osc}^{(k)}/k$. Thus we have $A(t) = \sum_{n=-\infty}^{\infty}(1/k)\delta(t - nT_{\rm osc}^{(k)}/k)$ and the threshold condition yields

$$T_{\rm osc}^{(k)} = \inf\left\{ s \,|\, \eta(s) + \frac{J_0}{k} \sum_{n=0}^{\infty} \kappa\left(s, s + \frac{nT_{\rm osc}}{k}\right) = \theta \right\}. \tag{97}$$

As before, the oscillation period is determined self-consistently. For $k = 1$ we are back to (96).

The solutions of (96) or (97) yield the period of possible collective excitation. To check the stability of such an oscillation, we go back to the case of a finite number of neurons with connections J_{ij}. We set $J_0 = \sum_{j=1}^{N} J_{ij}$.

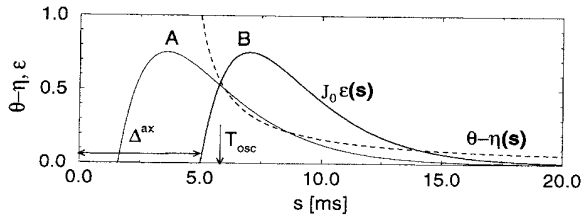

FIG. 4. Stability of a coherent oscillation. The intersection of the scaled response $J_0\,\epsilon(s)$ (solid line) with the effective threshold (dashed line) $\theta - \eta(s)$ yields the oscillation period $T_{\rm osc}$. The oscillation is stable, if the slope of $\epsilon(s)$ is *positive* at the intersection, as in case B (transmission delay $\Delta^{ax} = 5$ ms). It is unstable in case $A(\Delta^{ax} = 2$ ms).

For the sake of simplicity we restrict our arguments to the fully collective case ($k = 1$) and assume that during the past all neurons have fired synchronously at $t^n = -nT_{\rm osc}$ for $n \geq 1$. Around $t = 0$, however, when the next firing is due, the neurons are subject to a perturbation that causes a small shift in the firing time $t_i^0 = \delta_i^0$ with $|\delta_i^0| \ll T_{\rm osc}$. The next firing time $t_i^f = T_{\rm osc} + \delta_i^1$ can be calculated from the threshold condition (76). If we linearize around $t = T_{\rm osc}$ and use the fact that in the unperturbed case $\theta = \eta(T_{\rm osc}) + J_0 \sum_{n=1}^{\infty} \kappa(T_{\rm osc}, nT_{\rm osc})$, then we find

$$\delta_i^1 = \frac{\eta'}{\eta' + h'}\delta_i^0 + \frac{1}{\eta' + h'} \sum_{j=1}^{N} J_{ij}[c_0\theta(\delta_i^1 - \delta_j^1)\delta_j^1 + c_1\delta_j^0], \tag{98}$$

with

$$\eta' = \frac{d}{ds}\eta|_{T_{\rm osc}} + J_0 \sum_{n=1}^{\infty} \frac{d}{ds}\kappa(s, nT_{\rm osc})|_{T_{\rm osc}}, \tag{99}$$

$$h' = J_0 \left\{ \epsilon'|_{0+} + \sum_{n=1}^{\infty} \frac{d}{ds'}\kappa(T_{\rm osc}, s')|_{nT_{\rm osc}} \right\} \tag{100}$$

$$\epsilon'|_{0+} = J_0^{-1} \sum_{j=1}^{N} J_{ij}\theta(\delta_i^1 - \delta_j^1)c_0\,. \tag{101}$$

Furthermore we have $c_0 = \frac{d}{ds'}\kappa(T_{\rm osc}, s')|_{0+}$ and $c_1 = \frac{d}{ds'}\kappa(T_{\rm osc}, s')|_{T_{\rm osc}}$. The notation 0+ indicates a limit from above. The θ functions arise since, in the case of delayless interaction, the synaptic potential of the postsynaptic neuron may not be differentiable at $s' = 0$, the time of a presynaptic firing.

Note that c_0 and $\epsilon'|_{0+}$ always vanish, if the synaptic response includes a delay. In this case, the derivative of $\kappa(T_{\rm osc}, s')$ taken at $s' = T_{\rm osc}$ usually dominates the sum in (100). It follows that we have $c_1 \approx h'/J_0$. We use this approximation and return to the homogeneous network ($J_{ij} = J_0/N$). From (98) we find

$$N^{-1} \sum_{i=1}^{N} \delta_i^1 = N^{-1} \sum_{i=1}^{N} \delta_i^0 = \langle\delta\rangle \tag{102}$$

and

$$\delta_i^1 - \langle\delta\rangle = \frac{\eta'}{\eta' + h'}(\delta_i^0 - \langle\delta\rangle). \tag{103}$$

As a summary of the above arguments, we state the following results.

(i) A homogeneous shift in the firing times of all neurons $|\langle\delta\rangle| > 0$ is not reduced. This does not matter, however, since all neurons keep firing synchronously.

(ii) An inhomogeneous shift $\delta_i^0 \neq \langle\delta\rangle$ is reduced only if $(\eta'/\eta' + h') < 1$. Since for all standard models $\eta' > 0$ this requires $h' > 0$. In other words, the firing should occur at a point s where the synaptic potential is rising.

For an intuitive understanding of (ii), let us consider the response function $\kappa(s,s') = \epsilon(s')$ with $\epsilon(s')$ given by (24). We assume a finite delay Δ^{ax} and a time constant τ_s which is sufficiently short to allow fast rise and decay.

In this case, $\epsilon'|_{T_{\rm osc}}$ is the dominant term in (100) and (ii) is equivalent to the condition that the firing should occur at a point where $\frac{d}{ds'}\epsilon(s') > 0$. In a network of neurons which have spiked coherently at $t = 0$, all neurons are refractory during some time thereafter. The effective threshold decays as $\theta - \eta(t)$. After a time $\Delta^{\rm ax}$, neurons start to experience the postsynaptic potential of the firing at $t = 0$. If both the delay $\Delta^{\rm ax}$ and the rise time τ_s of the postsynaptic potential are short, then the maximum of $\epsilon(s')$ is reached, before the effective threshold is low enough to allow the next spike; cf. Fig. 4, case *A*. Thus firing will occur later in the regime where $\frac{d}{ds'}\epsilon(s') < 0$. We apply (ii) and find that the collective oscillation is unstable. On the other hand, if the delay is long enough, firing occurs while the postsynaptic potential is still rising and a collective oscillation is stable; cf. Fig. 4, case *B* and [58,59]. Similarly, if the rise time is sufficiently long, locking is possible. In this case, however, we have to take into account that the postsynaptic potential decays only slowly and use the full condition $h' = J_0 \sum_{n=1}^{\infty} \epsilon'|_{nT_{\rm osc}} > 0$.

With a delayed *inhibitory* interaction, $\kappa(s, s') = -\epsilon(s')$ with $\epsilon(s')$ given by (24), the above locking pattern is slightly different. The inhibition reaches its maximal strength after $s_m = \Delta^{\rm ax} + \tau_s$. With a realistic set of parameters, e.g., $\Delta^{\rm ax} = 2$ ms, $\tau_s = 2$ ms, the maximum is reached faster than the neuron recovers from refractoriness. Thus spiking can occur only later while inhibition decreases. But then, $\frac{d}{ds'}\kappa(s, s') = -\frac{d}{ds'}\epsilon(s') > 0$ for $s' > s_m$. Thus, in the case of inhibitory coupling, locking is stable for a wide range of different delays.

Finally, let us consider the case with no delay ($\Delta^{\rm ax} = 0$). For the sake of simplicity we restrict our arguments to the situation where only neuron i is perturbed at $t \approx 0$ and $\delta_j^0 = 0$ for $j \neq i$. In this case we find the simple relation

$$\delta_i^1 = \frac{\eta'}{\eta' + h'} \delta_i^0 \tag{104}$$

but now we have to consider the contribution $\epsilon'|_{0+}$ as well. To be specific, let us consider the situation that $h' > 0$ if $\delta_i^1 > 0$, but $h' < 0$ if $\delta_i^1 < 0$. This is the standard situation if we assume fast excitatory interaction. In particular, it applies to the Mirollo-Strogatz model, (17) and (18), or the one-step integrate-and-fire model, (10) and (11). It follows from (104) that a neuron which is late ($\delta_i^0 > 0$) is pulled back into the collective oscillation whereas a neuron which is early ($\delta_i^0 < 0$) fires even earlier the next time and drifts away. With our local stability analysis we cannot predict the time course of the phase shift during the drifting process. Drifting may be fast or slow or even stop at some point. It *must* slow down and stop, however, if after n steps $\delta_i^n \approx -T_{\rm osc}$. In other words, when the neuron is early by a full period, it is finally back in the collective state. We summarize the discussion through the following statement:

(iii) In the case of no delay and a sharply rising synaptic response function the system is only semistable. Neurons which fire too early will drift away. The drift stops if they eventually are ahead by a full period.

Let us apply this result to the two-step integrate-and-fire model defined by (10) and (12). The response kernel has been given in (39). The derivative $\frac{d}{ds'}\kappa(T_{\rm osc}, s')|_{nT_{\rm osc}}$ is negative, but of small absolute value if $T_{\rm osc} \gg \tau_1, \tau_2$. Thus h' in (100) is negative for those neurons which are early ($\delta_i^0 < 0$). For neurons which are late, however, the dominant contribution to h' is $\epsilon'|_{0+} \propto \frac{d}{ds'}\kappa(T_{\rm osc}, s')|_{0+} = (\tau_1\tau_2)^{-1} > 0$. Thus, according to (104), all neurons that are late by a delay $\delta_i^0 > 0$ experience a locking "force" which successively reduces the delay, whereas neurons which are early ($\delta_i^0 < 0$) slowly drift away.

This explains the effect discussed by Tsodyks *et al.* [60]. In their work they have studied a fully connected network of two-step integrate-and-fire neurons with delayless excitatory interaction. They introduced a small inhomogeneity by an external current $I_i^{\rm ext}$ which is kept constant in time but varies from neuron to neuron. Thus some neurons are systematically early, others are late. They find that only the group of "late" neurons is bound in the collective oscillation whereas the early neurons drift.

E. Model networks with noise

Real neurons are noisy, partly due to synaptic noise, partly due to inherent firing noise, but also due to the large "rest" of the brain which is out of experimental control. In Secs. I C 4 and III C, we have discussed how noise is included in the spike response model. Here we turn to pulse-coupled phase models or, equivalently, to integrate-and-fire models.

The standard way of introducing noise into integrate-and-fire neurons (16) is by an additive noise term $\sigma\zeta_i(t)$ with mean $\langle\zeta_i(t)\rangle = 0$ and variance $\sigma^2\langle\zeta_i(t)\zeta_j(t')\rangle = \sigma^2\delta_{ij}\delta(t - t')$. Thus we have

$$\frac{d}{dt}h_i = G(h_i) + I_i(t) + \sigma\zeta_i(t). \tag{105}$$

In the following we restrict our discussion to the standard integrate-and-fire model (10) with $G(h_i) = -h_i/\tau_1$. This allows us to separate the noise induced shift $\xi_i(t)$ in the membrane potential from its average time course $\bar{h}_i(t)$, i.e., $h_i(t) = \xi_i(t) + \bar{h}_i(t)$ where $\bar{h}_i(t)$ is a solution to the noiseless problem, Eq. (10). We focus on the change of the shift $\xi_i(t)$ and find

$$\frac{d}{dt}\xi_i(t) = -a\xi_i(t) + \sigma\zeta_i(t), \tag{106}$$

with $a = 1/\tau_1$. This is a standard Ornstein-Uhlenbeck process.

Let us discuss the influence of noise. Neuron i fires if $h_i > \theta$ or, equivalently, if $\xi_i(t) > \theta - \bar{h}_i(t)$. Let us now consider all neurons in pool $L(\mathbf{x})$ which have fired at $t = t^f$. They all experience the same mean potential $\bar{h}(t)$ for $t > t^f$. In principle, it should be possible to calculate the survivor function $S_h(t^f + s|t^f)$ and interval distribution $D_h(t^f + s|t^f)$ for arbitrary $\bar{h}(t)$. In the case of $\bar{h}(t) \equiv h_0$ the latter is given by the distribution of first passage times as calculated for the Ornstein-Uhlenbeck process. The only complication in the general case is that the distance $\theta - \bar{h}(t)$ that has to be overcome by

the diffusive process depends on the time t. We make no attempt to derive an expression in the general case (see [108,99] for a review of diffusion type neuron models). Rather we suggest two different approaches to an *ad hoc* approximation.

(1) In the first approach we start with the distribution of passage times $D_{h_0}(t^f + s|t^f)$ in the case of constant membrane potential $\overline{h}(t) \equiv h_0$. The first moment of the distribution yields the mean first passage time $\tau(h_0 - \theta)$. At this point of our considerations we replace the Ornstein-Uhlenbeck process by a Poisson process with instantaneous rate $\tau^{-1}(h_0 - \theta)$. This process has the same mean, but a different distribution of passage times, i.e., a simple exponential. In a further step we use the expression $\tau^{-1}(h-\theta)$ even in the case of time dependent h. This leads us to the survivor function $S_h(t^f + s|t^f) = \exp\{-\int_0^s \tau^{-1}[h(t^f+s')]ds'\}$ and we are back to the spike response model with noise. As a rough approximation to the mean first passage time in a diffusive process we can take $\tau(h-\theta) \propto (h-\theta)$ which is valid if h is near threshold.

(2) A different approximation is the following. As before we consider neurons in pool $L(\mathbf{x})$ which have fired their last spike at t^f. Due to noise, the membrane potential of the neurons varies. More precisely, the Ornstein-Uhlenbeck process produces a Gaussian distribution of the membrane potential of different neurons around the mean value $\overline{h}(t)$ with a width $\langle \xi^2 \rangle = (\sigma^2/2a)(1 - e^{-2as})$ where $s = t - t^f$ is the time that has passed since the last firing. If the mean $\overline{h}(t)$ approaches threshold, some of the neurons actually pass threshold and fire. Usually this would disturb the Gaussian distribution. If, however, the neurons are driven by a strong input, then $\overline{h}(t)$ crosses the threshold with finite "speed" $\frac{d}{dt}\overline{h}|_{t_\theta} > 0$ where t_θ is the time at which $\overline{h}$ reaches the threshold, that is, $\overline{h}(t_\theta) = \theta$. The total time that the distribution needs to cross threshold is roughly $\Delta t = \langle \xi^2 \rangle^{1/2} / \frac{d}{dt}\overline{h}|_\theta$. Of course, the distribution is disturbed as soon as some of the neurons fire. But as long as diffusion is slow, i.e., $\Delta t\, \sigma^2 \ll \langle \xi^2 \rangle$, there is no time for rearrangements. Consequently the threshold process basically "scans" a Gaussian distribution. As a result we have a Gaussian interval distribution with mean $\langle s \rangle = \inf\{s \,|\, \overline{h}(t^f + s) = \theta\}$ and width Δt. Given $D_{\overline{h}}(t^f + s|t^f)$, the survivor function can be found by integration of (33), i.e., $S_{\overline{h}}(t^f + s|t^f) = 1 - \int_0^s D_{\overline{h}}(t^f + s'|t^f)ds'$. Thus both approximation schemes yield the quantities needed for a solution of (53). If the neurons in are driven by strong input, then the synaptic potential crosses the threshold at finite "speed" and the second scheme leads to a good approximation. If, on the other hand, input is very low such that firing is driven mainly by noise, then the first scheme should be preferred.

To summarize, there are different ways of dealing with noise. In the case of the spike response model, noise is modeled by an instantaneous "escape rate" leading to a generalized Poisson process and an explicit expression for the survivor function can be derived. In the case of integrate-and-fire models, fluctuations are modeled by an additive noise term in the differential equation which leads to an Ornstein-Uhlenbeck process. An approximative evaluation of the survivor function is still possible, but network conditions must be known in order to find a good approximation scheme.

F. The model of Wilson and Cowan

The approach of Wilson and Cowan [96] is restricted to neurons with absolute refractory period only. In order to relate their approach to the spike response model, we take the refractory function

$$\eta(s) = \begin{cases} -\infty & \text{for } 0 < s \leq \gamma^{\text{refr}} \\ 0 & \text{for } s > \gamma^{\text{refr}} \end{cases} \tag{107}$$

and a synaptic response function or the form $\kappa(s,s') = \epsilon(s')$ with arbitrary $\epsilon(s')$. Also we take a firing probability $P_F(h;\delta t) = \tau^{-1}[h]\delta t$ with no explicit dependence upon the time s since the last postsynaptic spike. These assumptions cause a considerable simplification of the dynamic equation (71) since $\tau^{-1}[h(\mathbf{x},s,t)] = 0$ for $0 \leq s \leq \gamma^{\text{refr}}$ and $\tau^{-1}[h(\mathbf{x},s,t)] = \tau^{-1}[h^{\text{syn}}(\mathbf{x},t)]$ if $s > \gamma^{\text{refr}}$. Equation (71) yields

$$A(\mathbf{x},t) = \tau^{-1}[h^{\text{syn}}(\mathbf{x},t)] \left\{ 1 - \int_0^{\gamma^{\text{refr}}} A(\mathbf{x}, t-s)ds \right\}, \tag{108}$$

where we have used the normalization (52) and $h^{\text{syn}}(\mathbf{x},t) = \sum_{\mathbf{y}} J(\mathbf{x},\mathbf{y}) \int_0^\infty \epsilon(s')A(\mathbf{y}, t-s')ds'$. This is formally identical to the integral equation of Wilson and Cowan, if we take the time step Δt to be infinitesimally small. In order to compare (108) with (7), we have to add to the synaptic potential an external contribution $h^{\text{ext}}(\mathbf{x},t) = \int_0^\infty \epsilon(s')P(t-s')$ and identify $\tau^{-1}[h]$ with $S_{\mathbf{x}}[h]$. Note, however, that this implies a shift in the interpretation of τ^{-1} versus $S_{\mathbf{x}}$. In the original approach of Wilson and Cowan, the function $S_{\mathbf{x}}$ is induced by a variation of parameters, in particular different firing thresholds within a population of neurons. Thus it is a "quenched" disorder that gives rise to additional correlations which have to be neglected; cf. [96]. In our interpretation, all neurons are identical and $\tau^{-1}[h]$ is caused by local noise in each neuron. Thus the problem of additional correlations is avoided.

Note also that $S_{\mathbf{x}}$ is bounded in the interval $0 \leq S_{\mathbf{x}} \leq 1$ whereas $\tau^{-1}[h]$ has no upper bound. This is, however, an effect which is due to a difference between discrete versus continuous time dynamics. For the above identification of the two models we have assumed an infinitesimal time step $\Delta t \to 0$. For a discrete dynamics, the firing probability is given by (28) and is bounded between 0 and 1; cf. Sec. I C 4.

Our discussion of the spike response model in the preceding sections shows that we can get analytical insight into the collective behavior directly from the integral equation. A transformation to a differential equation which requires coarse graining of time is not necessary. To summarize this section, the integral equations of Wilson and Cowan are a special case of the pool equations

in the spike response model and correspond to neurons with absolute refractory period and discrete time steps.

IV. CONCLUSIONS

(1) We have introduced a model of a neuron which generalizes previous formulations of the spike response model and contains various versions of the integrate-and-fire model as special cases. In fact, it turns out to be the most general one-dimensional model with linear input and renewal properties.

(i) It follows from our considerations that in the limit of weak coupling, all spiking neuron models can be classified by their synaptic response kernel $\kappa(s, s')$ where s and s' denote the time that has passed since the last postsynaptic and presynaptic spike, respectively. Thus different models of spiking neurons can be systematically related to each other. We have used this procedure to compare various versions of the integrate-and-fire model, the Mirollo-Strogatz model, and the standard formulation of the spike response model.

(ii) Furthermore, we have analyzed the errors introduced by firing rate models. It turns out that a rate description is correct only in a stationary state of incoherent firing. A graded-response model is sensible if the membrane potential changes only slowly during a typical interspike interval. In particular, during collective oscillations a graded-reponse model is a bad approximation.

(2) All of the above network models can be understood and analyzed from the unifying point of view of a renewal theory with linear inputs. The dynamic evolution of the network is described by an integral equation for pool activities.

(3) Our approach shows that a systematic transition from single spiking neurons to populations of neurons is possible. The effective interaction between pool activities is highly nonlinear and involves several integrations over time. The integral equations of Wilson and Cowan [96] which apply to neurons with absolute refractory period only are contained as a special case.

(4) Finally, our approach allows a straightforward solution to the problem of the existence and stability of collective oscillations. Generally speaking, the shape and temporal relation between the two response functions, i.e., the synaptic response $\kappa(s, s')$ to incoming spikes and the internal response $\eta(s)$ to firing, determine the stability of incoherent, coherent, and partially coherent network states. In particular, the following results have been shown.

(i) The stability of incoherent states is determined by the Fourier transform of the synaptic response kernel $\kappa(s, s')$ with respect to the time s'. This general result unifies previous studies on a variety of different model systems [59,61,62].

(ii) The stationary state of a noiseless system is "almost always" unstable [59]. Instabilities may lead to collective oscillations of all neurons or to higher harmonics where the neurons spontaneously split into several subgroups. This is related to the phenomenon of phase clustering which can also be analyzed in the framework of return maps [129]. It is related to a result of Golomb *et al.* [128], but it is found here for *spiking* neurons.

(iii) Noise suprresses fast oscillations (higher harmonics) and stabilizes the system. This result, which has been found in simulation studies before [59], has been shown here analytically.

(iv) Collective oscillations with a period $T_{\rm osc}$ are stable only if firing occurs while the synaptic potential is increasing. The local stability analysis of collective states presented in Sec. III D generalizes earlier results which have been restricted to a specific form of a local perturbation [59,68].

(v) In particular, oscillations in networks with delay-free excitatory interactions are at most semistable. Excitatory interactions with long delays lead to stable oscillations. The same is true for models with inhibitory interaction and medium delays. This solves an often posed question, namely, whether excitatory or inhibitory interactions are more suitable to sustain collective oscillations. The answer is, it depends on the delay. Also it gives a direct explanation of the somewhat unexpected result of Tsodyks *et al.* [60], who found that a delayless network of integrate-and-fire neurons is only marginally stable—despite the exact result of Mirollo and Strogatz [56], who showed that a collective oscillation is the only attractive state in their (slightly different) model.

All of the models discussed in this paper are, of course, rather simple — far too simple to capture all phenomena of neuronal activity. Further steps towards a better theory of the brain are necessary and seem to be possible.

ACKNOWLEDGMENTS

Thanks are due to J. Leo van Hemmen, Andreas Herz, Richard Kempter, Raphael Ritz, and Stefan Wimbauer for stimulating discussions, helpful comments, and a critical reading of different versions of the manuscript. This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) under Grant Number He 1729/2-2.

[1] L. Lapique, J. Physiol. Pathol. Gen. **9**, 620 (1907); cited in H.C. Tuckwell, *Introduction to Theoretic Neurobiology* (Cambridge University Press, Cambridge, England, 1988).

[2] W. S. McCulloch and W. Pitts, Bull. Math. Biophys. **5**, 115 (1943).

[3] A. L. Hodgkin and A. F. Huxley, J. Physiol. (London) **117**, 500 (1952).

[4] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982).

[5] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **81**, 3088 (1984).

[6] *Neurocomputing: Foundations of Research*, edited by J. A. Anderson and E. Rosenfeld (MIT Press, Cam-

bridge, MA, 1988).
[7] *Brain Theory, Advances in Neuroscience*, Vol. 1, edited by G. L. Shaw and G. Palm (World Scientific, Singapore, 1988),
[8] P. Peretto, Biol. Cybern. **50**, 51 (1984).
[9] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. A **32**, 1007 (1985).
[10] D. J. Amit, *Modeling Brain Function* (Cambridge University Press, Cambridge, England, 1989).
[11] B. Müller and J. Reinhard, *Neural Networks: An Introduction* (Springer-Verlag, Berlin, 1991).
[12] *Models of Neural Networks*, edited by E. Domany, J. L. van Hemmen, and K. Schulten (Springer-Verlag, Berlin, 1991).
[13] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation* (Addison-Wesley, Redwood City, CA, 1991).
[14] P. Peretto, *An Introduction to the Modeling of Neural Networks* (Cambridge Universtity Press, Cambridge, England, 1992).
[15] D. M. MacKay and W. S. McCulloch, Bull. Math. Biophys. **14**, 127 (1952).
[16] H. B. Barlow, Kybernetik **2**, 1 (1963).
[17] D. H. Perkel, G. L. Gerstein, and G. P. Moore, Biophys. J. **7**, 391 (1967).
[18] R. B. Stein, Biophys. J. **7**, 797 (1967).
[19] R. B. Stein, Biophys. J. **7**, 37 (1967).
[20] R. Eckhorn *et al.*, Biol. Cybern. **22**, 49 (1976).
[21] L. M. Optican and B. J. Richmond, J. Neurophysiol. **57**, 162 (1987).
[22] R. R. de Ruyter van Stevenick and W. Bialek, Proc. R. Soc. B London, Ser. **234**, 379 (1988).
[23] W. Bialek, F. Rieke, R. R. de Ruyter van Stevenick, and D. Warland, Science **252**, 1854 (1991).
[24] T. J. Gawne, B. J. Richmond, and L. M. Optican, J. Neurophysiol. **66**, 379 (1991).
[25] D. H. Perkel, G. L. Gerstein, and G. P. Moore, Biophys. J. **7**, 419 (1967).
[26] G. L. Gerstein and D. H. Perkel, Biophys. J. **12**, 453 (1972).
[27] M. Abeles, *Local Cortical Circuits* (Springer-Verlag, Berlin, 1982).
[28] A. Aertsen, G. Gerstein, and P. Johannesma, in *Brain Theory*, edited by G. Palm and A. Aertsen (Springer-Verlag, Berlin, 1986), pp. 7–24.
[29] P. Johannesma *et al.*, in *Brain Theory*, edited by G. Palm and A. Aertsen (Springer-Verlag, Berlin, 1986), pp. 25–47.
[30] G. Palm, A. Aertsen, and G. L. Gerstein, Biol. Cybern. **59**, 1 (1988).
[31] J. J. Eggermont, *The Correlative Brain* (Springer, Berlin, 1990).
[32] J. Krüger, *Neuronal Cooperativity* (Springer, Berlin, 1991).
[33] C. von der Malsburg, MPI für Biophysikalische Chemie, Göttingen Internal Report No. 81-2, 1981 (unpublished).
[34] C. von der Malsburg and W. Schneider, Biol. Cybern. **54**, 29 (1986).
[35] C. von der Malsburg, in *Brain Theory*, edited by G. Palm and A. Aertsen (Springer-Verlag, Berlin, 1986), pp. 161–176.
[36] R. Eckhorn *et al.*, Biol. Cybern. **60**, 121 (1988).
[37] C. M. Gray and W. Singer, Proc. Natl. Acad. Sci. USA **86**, 1698 (1989).
[38] C. M. Gray, P. König, A. K. Engel, and W. Singer, Nature (London) **338**, 334 (1989).
[39] R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, Neural Comput. **2**, 293 (1990).
[40] A. K. Engel, P. König, and W. Singer, Proc. Natl. Acad. Sci. USA **88**, 9136 (1991).
[41] A. K. Engel, P. König, A. K. Kreiter, and W. Singer, Science **252**, 1177 (1991).
[42] W. Singer, in *Neural Cooperativity*, edited by J. Krüger (Springer, Berlin, 1991), pp. 165–183.
[43] A. K. Kreiter and W. Singer, Eur. J. Neurosci. **4**, 369 (1992).
[44] A. K. Engel *et al.*, Trends Neurosci. **15**, 218 (1992).
[45] R. Eckhorn *et al.*, Neuro Rep. **4**, 243 (1993).
[46] M. A. Cohen and S. Grossberg, IEEE Trans. Syst. Man Cybern. **13**, 815 (1983).
[47] D. Kleinfeld, Proc. Natl. Acad. Sci. USA **83**, 9469 (1986).
[48] H. Sompolinsky and I. Kanter, Phys. Rev. Lett. **57**, 2861 (1986).
[49] J. Buhmann and K. Schulten, Europhys. Lett. **4**, 1205 (1987).
[50] U. Riedel, R. Kühn, and J. L. van Hemmen, Phys. Rev. A **38**, 1105 (1988).
[51] A. V. M. Herz, B. Sulzer, R. Kühn, and J. L. van Hemmen, Europhys. Lett. **7**, 663 (1988).
[52] A. V. M. Herz, B. Sulzer, R. Kühn, and J. L. van Hemmen, Biol. Cybern. **60**, 457 (1989).
[53] D. Horn and M. Usher, Phys. Rev. A **40**, 1036 (1989).
[54] M. Kerszberg and A. Zippelius, Phys. Scr. **T33**, 54 (1990).
[55] M. Y. Choi, Phys. Rev. Lett. **61**, 2809 (1988).
[56] R. E. Mirollo and S. H. Strogatz, SIAM J. Appl. Math. **50**, 1645 (1990).
[57] Y. Kuramoto, Physica D **50**, 15 (1991).
[58] W. Gerstner and J. L. van Hemmen, Network **3**, 139 (1992).
[59] W. Gerstner and J. L. van Hemmen, Phys. Rev. Lett. **71**, 312 (1993).
[60] M. Tsodyks, I. Mitkov, and H. Sompolinsky, Phys. Rev. Lett. **71**, 1280 (1993).
[61] A. Treves, Network **4**, 259 (1993).
[62] L. F. Abbott and C. van Vreeswijk, Phys. Rev. E **48**, 1483 (1993).
[63] C. van Vreeswijk and L. F. Abbott, SIAM J. Appl. Math. **53**, 253 (1993).
[64] M. Usher, H. G. Schuster, and E. Niebur, Neural Comput. **5**, 570 (1993).
[65] J. Buhmann and K. Schulten, Biol. Cybern. **54**, 319 (1986).
[66] D. J. Amit and M. V. Tsodyks, Neural Comput. **2**, 259 (1991).
[67] W. Gerstner, in *Advances in Neural Information Processing Systems 3*, edited by R. P. Lippmann, J. E. Moody, and D. S. Touretzky (Morgan Kaufmann Publishers, San Mateo, CA, 1990), pp. 84–90.
[68] W. Gerstner, R. Ritz, and J. L. van Hemmen, Biol. Cybern. **68**, 363 (1993).
[69] W. Gerstner, R. Ritz, and J. L. van Hemmen, Biol. Cybern. **69**, 503 (1993).
[70] W. Gerstner, *Kodierung und Signalübertragung in Neuronalen Systemen: Assoziative Netzwerke mit stochastisch feuernden Neuronen*, Reihe Physik Vol. 15

(Harri Deutsch Verlag, Frankfurt/Main, 1993).
[71] W. Gerstner and J. L. van Hemmen, in *Models of Neural Networks II*, edited by E. Domany, J. L. van Hemmen, and K. Schulten (Springer-Verlag, New York, 1994), Chap. 1.
[72] R. Ritz, W. Gerstner, and J. L. van Hemmen, Biol. Cybern. **71**, 349 (1994).
[73] H. U. Bauer and K. Pawelzik, Physica D **69**, 380 (1993).
[74] M. A. Wilson, U. S. Bhalla, J. D. Uhley, and J. M. Bower, in *Advances in Neural Information Processing Systems*, edited by D. Touretzky (Morgan Kaufmann Publishers, San Mateo, CA, 1989), pp. 485–492.
[75] W. M. Yamada, C. Koch, and P. R. Adams, in *Methods in Neuronal Modeling, from Synapses to Networks*, edited by C. Koch and I. Segev (MIT Press, Cambridge, MA, 1989), Chap. 4.
[76] R. D. Traub, R. K. S. Wong, R. Miles, and H. Michelson, J. Neurophysiol. **66**, 635 (1991).
[77] O. Ekeberg *et al.*, Biol. Cybern. **65**, 81 (1991).
[78] W. W. Lytton and T. J. Sejnowsky, J. Neurophysiol. **66**, 1059 (1991).
[79] A. M. Wilson and J. M. Bower, Neural Comput. **3**, 498 (1991).
[80] P. C. Bush and R. J. Douglas, Neural Comput. **3**, 19 (1991).
[81] C. Kurrer, B. Nieswand, and K. Schulten, in *Self-Organization, Emerging Properties and Learning*, edited by A. Babloyantz (Plenum Press, New York, 1990), pp. 81–85.
[82] P. König and T. B. Schillen, Neural Comput. **3**, 155 (1991).
[83] D. Hansel and H. Sompolinski, Phys. Rev. Lett. **68**, 718 (1992).
[84] C. von der Malsburg and J. Buhmann, Biol. Cybern. **67**, 233 (1992).
[85] J. Deppisch *et al.*, Network **4**, 243 (1993).
[86] L. F. Abbott, J. Phys. A **23**, 3835 (1990).
[87] J. Buhmann, Phys. Rev. A **40**, 4145 (1989).
[88] Y. Kuramoto, in *International Symposium on Mathematical Problems in Theoretical Physics*, edited by H. Araki (Springer-Verlag, Berlin, 1975), pp. 420–422.
[89] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Springer, Berlin, 1984), pp. 68–77.
[90] H. G. Schuster and P. Wagner, Biol. Cybern. **64**, 77 (1990).
[91] H. G. Schuster and P. Wagner, Biol. Cybern. **64**, 83 (1990).
[92] H. Sompolinsky, D. Golomb, and D. Kleinfeld, Proc. Natl. Acad. Sci. USA **87**, 7200 (1990).
[93] H. Sompolinsky, D. Golomb, and D. Kleinfeld, Phys. Rev. A **43**, 6990 (1991).
[94] E. Niebur *et al.*, in *Advances in Neural Information Processing Systems 3*, edited by R. P. Lippmann, J. E. Moody, and D. S. Touretzky (Morgan Kaufmann, San Mateo, CA, 1991), pp. 123–127.
[95] C. Kurrer and K. Schulten (unpublished).
[96] H. R. Wilson and J. D. Cowan, Biophys. J. **12**, 1 (1972).
[97] J. L. Feldman and J. D. Cowan, Biol. Cybern. **17**, 29 (1975).
[98] J. J. B. Jack, D. Noble, and R. W. Tsien, *Electric Current Flow in Excitable Cells* (Clarendon Press, Oxford, 1975).
[99] H. C. Tuckwell, *Introduction to Theoretic Neurobiology* (Cambridge University Press, Cambridge, England, 1988).
[100] D. Kernell and H. Sjöholm, Acta Physiol. Scand. **87**, 40 (1973).
[101] F. Baldissera and B. Gustafsson, Acta Physiol. Scand. **92**, 27 (1974).
[102] B. W. Connors and M. J. Gutnick, Trends Neurosci. **13**, 99 (1990).
[103] T. H. Brown and D. Johnston, J. Neurophysiol. **50**, 487 (1983).
[104] U. Ernst, K. Pawelzik, and T. Geisel (unpublished).
[105] D. R. Cox, *Renewal Theory* (Methuen, London, 1962).
[106] D. R. Brillinger, Biol. Cybern. **59**, 189 (1988).
[107] S. Rotter, *Wechselwirkende stochastische Punktprozesse als Modell für neuronale Aktivität im Neocortex der Säugetiere*, Reihe Physik Vol. 21 (Harri Deutsch, Frankfurt/Main, 1994).
[108] G. Sampath and S. K. Srinivasan, *Stochastic Models for Spike Trains of Single Neurons* (Springer, Berlin, 1977).
[109] A. V. Holden, *Models of the Stochastic Activity of Neurons* (Springer, Berlin, 1976).
[110] V. Braitenberg and A. Schütz, *Anatomy of the Cortex* (Springer-Verlag, Berlin, 1991).
[111] G. L. Gerstein, P. Bedenbrough, and A. M. J. H. Aertsen, IEEE Trans. Biomed. Eng. **36**, 4 (1989).
[112] S. Amari, Kybernetik **14**, 201 (1974).
[113] J. L. van Hemmen and R. Kühn, Phys. Rev. Lett. **57**, 913 (1986).
[114] J. L. van Hemmen, D. Grensing, A. Huber, and R. Kühn, J. Stat. Phys. **50**, 231 (1988).
[115] D. H. Hubel and T. N. Wiesel, J. Physiol. (London) **160**, 106 (1962).
[116] D. H. Hubel and T. N. Wiesel, Proc. R. Soc. London, Ser. B **198**, 1 (1977).
[117] V. Braitenberg, in *Brain Theory*, edited by G. Palm and A. Aertsen (Springer-Verlag, Berlin, 1986), pp. 81–96.
[118] G. Krone, H. Mallot, G. Palm, and A. Schütz, Proc. R. Soc. London, Ser. B **226**, 421 (1986).
[119] H. R. Wilson and J. D. Cowan, Kybernetik **13**, 55 (1973).
[120] J. L. van Hemmen, Phys. Rev. A **36**, 1959 (1988).
[121] D. Grensing and R. Kühn, J. Phys. A **19**, L1153 (1986).
[122] J. L. van Hemmen, D. Grensing, A. Huber, and R. Kühn, Z. Phys. B **65**, 53 (1986).
[123] W. von Seelen, Kybernetik **5**, 133 (1968).
[124] G. B. Ermentrout and J. D. Cowan, J. Math. Biol. **7**, 265 (1979).
[125] G. B. Ermentrout and J. D. Cowan, SIAM J. Appl. Math. **38**, 1 (1980).
[126] M. A. P. Idiart and L. F. Abbott, Network **4**, 285 (1993).
[127] W. Gerstner and J. L. van Hemmen, Biol. Cybern. **67**, 195 (1992).
[128] D. Golomb, D. Hansel, B. Shraiman, and H. Sompolinsky, Phys. Rev. A **45**, 3516 (1992).
[129] U. Ernst, K. Pawelzik, and T. Geisel, in *Proceedings of the ICANN'94*, edited by M. Marinaro and P. G. Morasso (Springer-Verlag, Berlin, 1994), pp. 1063–1065.