Network 4 (1993) 259–284. Printed in the UK

Mean-field analysis of neuronal spike dynamics

Alessandro Treves†

Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK

Received 5 October 1992, in final form 19 February 1993

Abstract. I consider a mean-field description of the dynamics of interacting integrate-and-fire neuron-like units. The basic dynamical variables are the membrane potential of each (point-like) 'cell' and the conductance associated with each synaptic connection, both of which evolve discontinuosly in time. In addition, an *intrinsic potassium conductance*, also evolving discontinuosly in time, can be associated to each cell in order to model firing frequency adaptation in real neurons. The mean-field theory is exact if the units can be grouped into N_C classes, each comprising infinitely many identical, and identically coupled, units; and can be used as an approximation if, instead, a class comprises few or just one unit. The formalism yields both the stationary asynchronous solutions and the transients leading to those solutions. The full spectrum of time-constants for the transients associated with one particular steady state is given by a single equation, imposing the vanishing of the determinant of an $N_C \times N_C$ matrix. In the case of an associative memory, this equation can be manipulated into a simple form, using standard replica methods. An analysis of the spectrum indicates that the major role in determining the transients time constants is played by the effective decay times of postsynaptic currents, which can be quite short. This suggests that local recurrent neocortical circuits may produce a very rapid dynamics, consistent with such circuits participating in the rapid course of information processing, evidenced by new experimental data recorded in primate temporal cortex.

1. Introduction

Many analytical studies of the collective properties of neural networks have been based on very crude representations of actual single neuron behaviour by means of simplified processing 'units'. The specific choice of the units has been largely dictated by the nature of the properties to be investigated. Thus, in network models of associative memory [1, 2, 3], in which information is supposedly carried with temporally coarse resolution by rates of emission of action potentials, units have been chosen in which a single output variable (often even binary) represented some short-time average of the neuron's spiking rate. The focus then has been on studying the steady-state behaviour, i.e. the attractor structure [4], with the dynamics either neglected altogether or artificially defined (for example, in terms of 'updatings') merely in order to fully specify the model. In those network models [5, 6] of early visual processing, instead, which purport to illustrate the hypothesis [7, 8, 9] that the partial synchronization of emission times might serve to bind across the visual field features pertaining to single objects, units have been generally considered that have an intrinsically oscillatory probability of emitting spikes.

† Present address: SISSA—Biophysics, via Beirut 2-4, 34013 Trieste, Italy.

0954-898X/93/030259+26$07.50

Much more realistic descriptions of the relevant biophysical processes have been used in modelling the dynamics of just one or a few interconnected neurons. In trying to understand the collective properties of large networks, though, attempts to get closer to biological situations have been confined to simulation studies, with all their obvious limitations. Yet the need to obtain an analytical grasp of the *dynamics* of large networks becomes urgent, as new experimental data [10] is beginning to throw light on the temporal course of information processing, e.g. in neocortical association areas of primates.

This paper aims to show that it is possible to study analytically aspects of the dynamics of large interconnected networks even when neurons (and synapses) are represented more realistically than usually attempted. The two main ingredients, included in the present treatment, are an integrate-and-fire description of the time evolution of single cells membrane potentials, and, crucially, a description of synaptic interactions in terms of the dynamics of post-synaptic conductances. Moreover, frequency adaptation, a salient feature of neocortical pyramidal cells, which affects significantly their temporal response patterns, is produced in the present model by an intrinsic potassium conductance driven by a cell's own action potentials. While these are still very simple approximations [11] (they form the basis of the simplest neurophysiologically inspired simulation studies [12] and hardware implementations [13]) they are important steps towards addressing dynamical issues raised by experimental data, and which lie well beyond the scope of previous treatments. For example, over what *time scales* can a local firing pattern—elicited by an incoming stimulus—reach a steady-state distribution? Which of the biophysical parameters at the single unit level contributes to set those time scales? Are those time scales consistent with the hypothesis that the steady-state distribution is utilised in further stages of cortical processing?

Analysing a realistic form of dynamics will also be necessary in order to study the *basins of attraction* of stable attractor states [14]. Moreover, it will allow a discussion of some simple models which have been put forward in the context of *synchronous oscillations* [15, 16], although the description proposed here may not be complete enough (e.g. by not including bursting [17]) to discuss the phenomenon itself.

This paper focusses on establishing the relationship between the time scales at the single unit and at the collective level. The complex question of the stability of the steady-state distributions is addressed in another paper [18], while a more detailed discussion of the time course of neocortical processing will also be given elsewhere [19]. The paper is organized as follows: the model and the notation used are defined in section 2. The mean-field analysis is introduced using the toy case of just two classes of identical cells. This is done first for cells that do not display adaptation, in section 3, followed by a discussion of the spectrum, and by a proof that periodic solutions are impossible in this case. The analysis is extended to cells that do adapt in section 4. A brief comment is made on the situation in which *all* cells are identical and identically coupled. Section 5 deals with more interesting cases, including that of an associative memory, and presents an example of the spectrum which includes the effects of adaptation. A brief discussion relates the analysis to experimental results in the section 6.

2. Basic model

2.1. Single cells and synapses

In the simplest type of integrate-and-fire description, the one adopted here, the geometry of the neuron is reduced to a point, and the time evolution of the membrane potential V_i of cell i during an interspike interval follows the RC equation

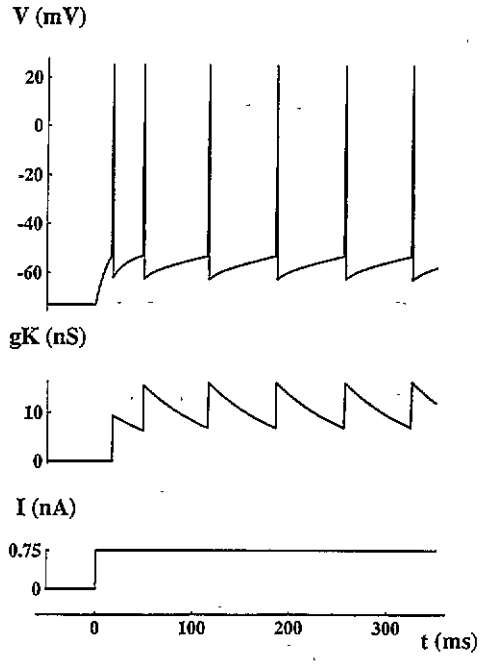

Figure 1. Model *in vitro* behaviour. $V(t)$ and $g^{\mathrm{K}}(t)$ in response to a step of injected current $I(t)$ of 0.75 nA. The relevant parameters are $V^0 = -73$ mV, $V^{\mathrm{thr}} = -53$ mV, $V^{\mathrm{ahp}} = -63$ mV, $V^{\mathrm{K}} = -85$ mV, $C = 0.375$ nF, $g^0 = 25$ nS, $\Delta g^{\mathrm{K}} = 9.375$ nS, $\tau^{\mathrm{K}} = 80$ ms. Spikes are pasted by hand at emission times for clarity of presentation.

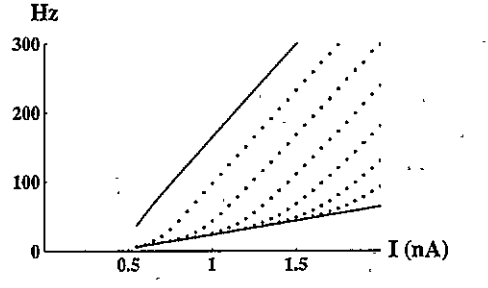

Figure 2. Current-to-frequency transduction. The top solid curve is the firing frequency in the absence of adaptation, $\Delta g^{\mathrm{K}} = 0$. The dotted curves are the instantaneous frequencies computed as the inverse of the ith interspike interval (top to bottom, $i = 1, \ldots, 6$). The bottom solid curve is the adapted firing curve ($i \to \infty$). Parameters as in figure 1.

$$C_i \frac{\mathrm{d}V_i(t)}{\mathrm{d}t} = g_i^0(V_i^0 - V_i(t)) + g_i^{\mathrm{K}}(t)(V_i^{\mathrm{K}} - V_i(t)) + \sum_{\alpha} g^{\alpha}(t)(V^{\alpha} - V_i(t)) + I_i(t) \qquad (1)$$

where C_i is the capacity of the cell membrane, g_i^0 its passive conductance, V_i^0 the resting potential, g_i^{K} an active potassium conductance producing firing rate adaptation, V_i^{K} the corresponding equilibrium potential, and g^{α} and V^{α} the conductance and equilibrium potential of each input synapse α. $I_i(t)$ is an external current injected into the cell, considered here solely in order to illustrate, in figures 1 and 2, the way the model would mimic *in vitro* data on the response characteristics of neocortical cells.

As V_i reaches a threshold value V_i^{thr}, the cell emits a spike†. Unlike the Hodgkin–Huxley model, the integrate-and-fire model does not describe the cascade of membrane conductance changes that accompany an action potential. All these changes, instead, are supposed to be extremely rapid, and to leave the cell repolarized at a reset potential V_i^{ahp}, from which the evolution resumes as in equation (1). One may introduce an absolute refractory period lasting 1–2 ms, during which the cell cannot emit further spikes no matter how strongly stimulated. Here, however, the absolute refractory period is set, for simplicity, to zero,

† Here V_i^{thr} is taken to be constant. In other cases [12] it is assumed to vary in time, with an evolution similar to that of g_i^{K}, equation 3, and then it contributes to firing frequency adaptation.

although this obviously makes infinite spiking rates possible. In all situations of interest to us, in fact, the actual rates will be determined by other mechanisms at much lower levels than the inverse of the absolute refractory period, which will therefore be influential.

Some of the synaptic contacts onto a neuron are taken to be made by the axons of other neurons in the network, and some by external afferents, and in both cases the relative conductance is assumed to open instantaneously by a fixed amount following every incoming spike, and to relax exponentially to its closed state. Thus if t_{k,j_α} is the time of emission of the kth spike by the neuron j_α presynaptic to synapse α

$$\frac{dg^\alpha(t)}{dt} = -\frac{g^\alpha(t)}{\tau^\alpha} + \Delta g^\alpha \sum_k \delta(t - \Delta t - t_{k,j_\alpha}) \tag{2}$$

where τ^α is a synaptic time constant and Δt summarizes axonal, synaptic and dendritic delays. A similar dynamics governs the potassium conductance, the only intrinsic and time-varying conductance included in the model†

$$\frac{dg_i^K(t)}{dt} = -\frac{g_i^K(t)}{\tau^K} + \Delta g_i^K \sum_k \delta(t - t_{k,i}) \tag{3}$$

i.e. when cell i fires, g_i^K is increased by Δg_i^K, and then it decays exponentially with time constant τ^K. This type of conductance dynamics is meant as a representation of the summed dynamics of many individual channels, and in the synaptic case it is but an approximation of the more accurate α-function representation [21], valid in the limit in which the activation of the conductance is much faster than its inactivation. The analysis below can in fact be easily extended to conductances that follow an α-function dynamics [18].

This level of description is obviously inadequate to reproduce complex neuronal behaviours, such as those related to the cell geometrical structure or to calcium dynamics, e.g. active dendritic processing or bursting. It is, on the other hand, sufficient in order to reproduce the basic features of neuronal transduction, its graded nature and, in the case of regular-firing pyramidal cells, the adaptation in frequency following the first few spikes. The behaviour observed *in vitro* to a step current injected into the cell body is modelled, here, by setting all $g^\alpha \equiv 0$ (figure 1). The resulting current-to-frequency transduction curves, illustrating the effect of adaptation, are shown in figure 2.

2.2. *Networks*

Equations (1), (2) and (3), together with the prescription concerning spikes, form a closed system, once the architecture of the network, the incoming afferents and all the various parameters are specified. The cell parameters should reflect the class to which each cell belongs [22], with the cell classes chosen in a way appropriate to the particular brain structure being modeled, and the specific dynamical issue under consideration. In neocortex, for example, one may want to begin by separating out $GABA_A$ and $GABA_B$ inhibitory cells [23], layer 2/3 and layer 5 pyramidal cells, among the latter the ones with short and long apical dendrite [24], and so on, and then try to reproduce, maybe with finer subdivisions, experimental data concerning both cell and synapse parameters and connection probabilities. Here we start by illustrating the method, and the results it can produce, with a toy case, in which there are only two classes of neurons: excitatory and inhibitory. Further, the various parameters are taken to be identical within each class, and the simplest connectivity scheme is considered, in which all cells are connected, and the 'strength' of each synapse—i.e., the

† Cf the multitude of intrinsic conductances observed in real neurons, e.g. in the hippocampus [20].

conductance Δg^α—depends only the classes of the pre- and post-synaptic cell. Later, we shall see how the same methods can be extended to cases in which there is variability in cell and synaptic parameters, or there are more than just two classes of cells‡.

Let the subscripts $F, G, \ldots$ index cell classes—as a start, they will only take the values E (excitatory) or I (inhibitory). Whenever, as in the case of synaptic parameters, it is necessary to specify also the *pre*-synaptic cell, the same letters will be used as superscripts. Each cell class will be characterized by a reset membrane potential V_F^{ahp} and a threshold potential V_F^{thr}. It is convenient to express all characteristic potentials in terms of the variable

$$x = \frac{V - V_F^{\text{ahp}}}{V_F^{\text{thr}} - V_F^{\text{ahp}}} \tag{4}$$

measuring the excursion of the membrane potential between spikes. For example, the resting membrane potential for excitatory cells will be, in these units, x_E^0. The equilibrium potential characterizing excitatory synapses onto inhibitory neurons will be x_I^E, and so on.

It is also convenient to parametrize the degree to which the potassium conductance is open as

$$y = \frac{g_i^{\text{K}}}{\Delta g_i^{\text{K}}}. \tag{5}$$

The remaining synaptic parameters, in the simplest network considered here, are the set of time constants $\{\tau_G^F\}$ and the set of conductances $\{\Delta g_G^F\}$. For example, τ_E^I and Δg_E^I characterize any synapse from an inhibitory to an excitatory neuron. It helps to reduce all conductance parameters, that in fact act as 'couplings', to the dimension of a frequency, by dividing them by the capacities C_F. Thus ω_F^0 will denote the inverse of the time constant for passive membrane leakage in cells of class F, i.e. g_F^0/C_F. Similarly, the quantities $N_G \Delta g_F^G / C_F$, which measure, in units of frequency, the total synaptic strengths of cells of class G (of which there are N_G) onto a cell of class F, will be denoted as ω_F^G.

The simple assumptions concerning the connectivity and the synaptic efficacies are reflected in the fact that the inputs to any cell in the network are determined by globally defined quantities, namely the mean fields. These measure, as a function of time, the effective fraction of synaptic conductances (in units of Δg) opened on the membrane of any cell of a given class (say, F) by the action of all presynaptic cells of another given class (G):

$$z_F^G(t) = \frac{1}{N_G} \sum_{\alpha \in G} \frac{g^\alpha(t)}{\Delta g_F^G}. \tag{6}$$

The external inputs, denoted as $S_F(t)$, are taken, again for the sake of simplicity, to be excitatory only.

In summary, the single-neuron dynamics, for cells of class F, is described by the equations

$$\dot{x}_i(t) = \omega_F^0(x_F^0 - x_i(t)) + \omega_F^{\text{K}}(x_F^{\text{K}} - x_i(t))y_i(t) + \omega_F^E(x_F^E - x_i(t))z_F^E(t) + \omega_F^I(x_F^I - x_i(t))z_F^I(t) + \omega_F^S(x_F^E - x_i(t))S_F(t) \tag{7}$$

and

$$\dot{y}_i(t) = -\frac{y_i}{\tau^{\text{K}}} \tag{8}$$

‡ The toy case is chosen to have two classes of cells in order to avoid the pathologies of the yet simpler case, in which all cells are identical and identically coupled, cf subsections 3.4 and 4.3.

in between spikes, with the supplementary prescription that as $x_i(t)$ reaches 1, it is reset to 0, $y_i(t)$ is increased by 1, and a new spike is emitted by cell i, i.e. $t_{k,i} = t$.

3. Dynamics, with no adaptation

Consider first the case $\omega_F^K \equiv 0$, so that the y-dynamics becomes irrelevant.

Write equation (7) as

$$\dot{x}_i(t) = A_F(t) - x_i(t)B_F(t), \tag{9}$$

by using the notation for the rise rates of the membrane potential

$$\begin{aligned} A_F(t) &= x_F^0\omega_F^0 + x_F^E\omega_F^E z_F^E(t) + x_F^I\omega_F^I z_F^I(t) + x_F^E\omega_F^S S_F(t) \\ B_F(t) &= \omega_F^0 + \omega_F^E z_F^E(t) + \omega_F^I z_F^I(t) + \omega_F^S S_F(t). \end{aligned} \tag{10}$$

Then introduce the densities

$$\rho_F(x,t) = \frac{1}{N_F}\sum_i \delta(x - x_i(t)) \tag{11}$$

where N_F is the number of cells of class F.

In terms of these quantities, the dynamic system can be written

$$\begin{aligned} \dot{\rho}_F(x,t) &= B_F(t)\rho_F(x,t) - [A_F(t) - xB_F(t)]\frac{\partial}{\partial x}\rho_F(x,t) \\ &\quad +\delta(x)A_F(t)\left[\rho_F(0^+,t) - \rho_F(0^-,t)\right] \\ &\quad -\delta(x-1)[A_F(t) - B_F(t)W]\rho_F(1^-,t)) \\ \dot{z}_G^F(t) &= -\frac{1}{\tau_G^F}z_G^F(t) + [A_F(t-\Delta t) - B_F(t-\Delta t)]\,\rho_F(1^-,t-\Delta t). \end{aligned} \tag{12}$$

with the 'boundary' conditions

$$[A_F(t) - B_F(t)]\rho_F(1^-,t) = A_F(t)\left[\rho_F(0^+,t) - \rho_F(0^-,t)\right] \tag{13}$$

which correspond to the spike prescription. The second set of equations (12) is derived from equations (2),(6) by noting that $\delta(t) = \dot{x}(t)\delta(x - x(t))$, and it brings out the role of the firing rates $\nu_F(t) \equiv [A_F(t) - B_F(t)]\rho_F(1^-,t))$ in driving the conductance dynamics.

3.1. Stationary solutions

If the $S(t)$'s are constant in time, one has stationary solutions in the form

$$\begin{aligned} \rho_F^0(x) &= \Theta(x)\frac{\nu_F^0}{A_F^0 - xB_F^0}\Theta(1-x) \\ z_G^{F0} &= \tau_G^F\nu_F^0 \end{aligned} \tag{14}$$

($\Theta(x)$ being the Heaviside function) where the firing rates of each class of cells are

$$\nu_F^0 \equiv (T_F^0)^{-1} = B_F^0\left[\ln\frac{A_F^0}{A_F^0 - B_F^0}\right]^{-1}. \tag{15}$$

i.e. the values of ν_E^0 and ν_I^0 are given by the system of equations above (for $F = E, I$), with A_F^0 and B_F^0 expressed back in terms of ν_E^0 and ν_I^0 according to equations (10),(14). This system yields finite, non-zero values for the amounts of excitation and inhibition, within an ample region in the space of the parameters ω, x and τ. ν_E^0 and ν_I^0 tend to infinity (and the approximation of neglecting the absolute refractory period loses meaning) beyond a certain hypersurface in the above space. The hypersurface is given by the system

$$\begin{aligned}
\frac{\nu_E^0}{\nu_I^0} \equiv r &= (\omega_E^E \tau_E^E r + \omega_E^I \tau_E^I)\left[\ln\left(\frac{x_E^E \omega_E^E \tau_E^E r + x_E^I \omega_E^I \tau_E^I}{(x_E^E - 1)\omega_E^E \tau_E^E r + (x_E^I - 1)\omega_E^I \tau_E^I}\right)\right]^{-1} \\
1 &= (\omega_I^E \tau_I^E r + \omega_I^I \tau_I^I)\left[\ln\left(\frac{x_I^E \omega_I^E \tau_I^E r + x_I^I \omega_I^I \tau_I^I}{(x_I^E - 1)\omega_I^E \tau_I^E r + (x_I^I - 1)\omega_I^I \tau_I^I}\right)\right]^{-1}
\end{aligned} \tag{16}$$

and it is independent of S^0. Conversely, the equations determining the hypersurfaces beyond which either $\nu_E^0 = 0$ or $\nu_I^0 = 0$ are, respectively, $A_E^0 = B_E^0$ and $A_I^0 = B_I^0$, and they do depend on the stimuli S^0. Obviously, if for example $\nu_I^0 = 0$, the dynamical system reduces to a homogeneous population of firing excitatory cells.

3.2. Transients

Having found a stationary solution, the dynamics of the transients (keeping the same symbols for the subtracted quantities, e.g. $A_E(t) \to A_E^0 + A_E(t)$) is governed by

$$\begin{aligned}
\dot{\rho}_F(x,t) = {} & B_F^0 \rho_F(x,t) - [A_F^0 - x B_F^0]\frac{\partial}{\partial x}\rho_F(x,t) \\
& + B_F(t)\rho_F^0(x) - [A_F(t) - x B_F(t)]\frac{\partial}{\partial x}\rho_F^0(x) \\
& + B_F(t)\rho_F(x,t) - [A_F(t) - x B_F(t)]\frac{\partial}{\partial x}\rho_F(x,t) \\
\dot{z}_G^F(t) = {} & -\frac{1}{\tau_G^F} z_G^F(t) + [A_F^0 - B_F^0]\rho_F(1^-, t - \Delta t) \\
& + [A_F(t - \Delta t) - B_F(t - \Delta t)]\,\rho_F^0(1^-) \\
& + [A_F(t - \Delta t) - B_F(t - \Delta t)]\,\rho_F(1^-, t - \Delta t)
\end{aligned} \tag{17}$$

with the constraint that

$$\int_{-\infty}^{1} \mathrm{d}x\, \rho_F(x,t) = 0 \tag{18}$$

and that the discontinuities at $x = 0, 1$ satisfy

$$\begin{aligned}
&[A_F^0 - B_F^0 + A_F(t) - B_F(t)]\,(\rho_F^0(1^-) + \rho_F(1^-, t)) \\
&\qquad = \left[A_F^0 + A_F(t)\right]\left[\rho_F^0(0^+) + \rho_F(0^+, t)\right].
\end{aligned} \tag{19}$$

In the last equation we have assumed that after a finite time the density distribution for any firing class of cells has support on the interval (0,1), even if $V^0 < V^{\mathrm{ahp}}$.

One may now linearize the system. The part linear in the transients (neglecting the last line of each of equations (17)) describes a collection of modes, each characterised by a time dependence $\exp(\lambda t)$, for the transient parts of the dynamical quantities considered, i.e. the $N_C \times N_C$ (in our toy case just 4) adimensional conductances z and the N_C (here,

2) density functions $\rho(x)$. The spectrum of all possible (inverse) time constants λ of such collective modes is in general given by solving what is essentially an eigenvalue equation (but note the additional dependence on λ introduced through the time delay Δt). Even in its linearized version the dynamical system may appear complicated to analyse, as part of the variables are really functions of x as well as t. The number of effective variables can be reduced, however, to just N_C simple scalar quantities: a single manageable determinant equation then yields all possible λ's, and the only trace of the reduction procedure is in additional λ dependences of the corresponding matrix. The procedure, which does not entail any approximation but rather stepwise solutions, is shown here in the following.

Consider any particular complex value λ. Keeping the full set of variables z is in fact redundant, because they can be expressed in terms of the firing rates $\nu_F(t)$ of each class

$$\dot{z}_G^F(t) = -\frac{1}{\tau_G^F} z_G^F(t) + \nu_F(t - \Delta t) \tag{20}$$

and this holds for both the unsubtracted and the transient quantities, so that

$$\nu_F(t - \Delta t) = (\lambda + 1/\tau_E^F) z_E^F(t) = (\lambda + 1/\tau_I^F) z_I^F(t) \tag{21}$$

where each transient field $z_G^F(t)$ diverges if $-\lambda^{-1}$ equals its corresponding synaptic time constant τ_G^F, a fact which will have important implications later on.

Moreover, the linearized equations for the $\dot{\rho}_F(x, t)$'s are solved by transient distributions of the form

$$\rho_F(x,t) = \frac{\rho_F^0(x)}{B_F^0 + \lambda}[A_F^0 B_F(t) - B_F^0 A_F(t)] \times \left\{ [A_F^0 - xB_F^0]^{-1} + \frac{\lambda}{B_F^0}[A_F^0 - xB_F^0]^{\lambda/B_F^0} \frac{[A_F^0 - B_F^0]^{-1} - [A_F^0]^{-1}}{[A_F^0 - B_F^0]^{\lambda/B_F^0} - [A_F^0]^{\lambda/B_F^0}} \right\} \tag{22}$$

except when $\lambda = -B_F^0$, in which case

$$\rho_F(x,t) = \frac{\rho_F^0(x)}{B_F^0} \frac{[A_F^0 B_F(t) - B_F^0 A_F(t)]}{[A_F^0 - xB_F^0]} \times \left\{ 1 - \ln[A_F^0 - xB_F^0] + \frac{A_F^0 \ln[A_F^0 - B_F^0] - [A_F^0 - B_F^0] \ln A_F^0}{B_F^0} \right\} \tag{23}$$

and when $\lambda = 0$, in which case

$$\rho_F(x,t) = \frac{\rho_F^0(x)}{B_F^0}[A_F^0 B_F(t) - B_F^0 A_F(t)] \times \left\{ [A_F^0 - xB_F^0]^{-1} + \frac{[A_F^0 - B_F^0]^{-1} - [A_F^0]^{-1}}{\ln[A_F^0 - B_F^0] - \ln A_F^0} \right\}. \tag{24}$$

These distributions satisfy the required constraints (to first order).

One is left at this point with the system of N_C equations

$$\begin{aligned} \nu_F(t) &= \left[A_F^0 - B_F^0\right] \rho_F(1^-, t) + [A_F(t) - B_F(t)]\ \rho_F^0(1^-) \\ &= \frac{\nu_F^0}{B_F^0 + \lambda} \left\{ B_F(t) + \lambda \frac{B_F(t) + A_F(t)[e^{-(\lambda + B_F^0)T_F^0} - 1]}{[A_F^0 - B_F^0][e^{-\lambda T_F^0} - 1]} \right\} \end{aligned} \tag{25}$$

where the $A_F(t)$'s and $B_F(t)$'s are given by equations (10) in terms of the $z_F^G(t)$'s, and those in turn by equations (21) in terms of the $\nu_G(t)$'s.

Defining

$$Q_F^G(\lambda) = \frac{\nu_F^0 \omega_F^G}{(\lambda + 1/\tau_F^G)(B_F^0 + \lambda)} \left\{ 1 + \lambda \frac{1 + x_F^G[\mathrm{e}^{-(\lambda + B_F^0)T_F^0} - 1]}{[A_F^0 - B_F^0][\mathrm{e}^{-\lambda T_F^0} - 1]} \right\} \tag{26}$$

the complex equation that sets the possible λ's reads

$$\left| \mathbf{Q}(\lambda) - \mathrm{e}^{\lambda \Delta t} \mathbf{1} \right| = 0. \tag{27}$$

In our $N_C = 2$ toy case the equation is just

$$[Q_E^E(\lambda) - \mathrm{e}^{\lambda \Delta t}][Q_I^I(\lambda) - \mathrm{e}^{\lambda \Delta t}] - Q_E^I(\lambda) Q_I^E(\lambda) = 0. \tag{28}$$

The transients decay, and the static distribution is stable, if no solution exist with $\mathrm{Re}(\lambda) \geq 0$. For instabilities that appear continuously, the onset is at $\mathrm{Re}(\lambda) = 0$. Note that even if at the same time $\mathrm{Im}(\lambda) \to 0$ the limit is regular, and one may use equation (27) with

$$Q_F^G(0) = \frac{\nu_F^0 \omega_F^G \tau_F^G}{B_F^0} \left\{ 1 - \frac{1 + x_F^G[\mathrm{e}^{-B_F^0 T_F^0} - 1]}{[A_F^0 - B_F^0] T_F^0} \right\}. \tag{29}$$

3.3. The spectrum

A numerical solution of equation (27) yields a rich spectrum, whose main features can be appreciated in figures 3 and 4. In order to understand how the different zeros of equation (27) relate to the several parameters present in the model, within a particular regime, it is obviously important to choose the parameters as realistically as possible. An attempt to be inside a neocortically meaningful regime is made in the examples given, with the limitation that pyramidal cell firing can hardly be represented realistically in the absence of adaptation.

It is helpful to distinguish a gross and a fine structure in the spectrum (a hyperfine structure will be introduced later by the presence of adaptation). The *gross* structure, extending in the kHz range (and thus associated with the shortest time constants), is the one portrayed in figure 3. A non-zero delay Δt produces a characteristic set of zeros (figure 3(*a*)); some of these may have a positive real part and thus make the corresponding stationary solution unstable. The zeros with a negative real part occur with an approximate $\Delta\lambda \approx \mathrm{i}(2\pi)/(N_C \Delta t)$ periodicity along the imaginary axis, as determined by the changes in sign of the two $\exp(\lambda \Delta t)$ terms. Their negative real part is determined, instead, by the competition between the above terms and those containing Q, and thus scales so as to satisfy $\mathrm{Re}(\lambda) \approx (1/\Delta t) \ln[-\mathrm{Re}(\lambda)(A_F^0 - B_F^0)/(\nu_F^0 \omega_F^G x_F^G)]$. This set of zeros is removed by setting $\Delta t = 0$ (figure 3(*b*)). There may or may not remain additional zeros with $|\mathrm{Re}(\lambda)|$ large.

The *fine* structure of the spectrum, arising from the spiking nature of the dynamics, is more of interest to this paper. The relevant λ-scales are set by the inverse of the conductance time constants and by the firing frequencies characterising the stationary solutions, and thus cover the 10–300 Hz range suggested here as crucial to understanding information processing in certain real neural systems. The position of these set of zeros can be related to the structure of the Q-matrix. The matrix element $Q_F^G(\lambda)$ has poles at $\lambda = -1/\tau_F^G$ and on the imaginary axis at $\lambda = \pm \mathrm{i} 2\pi n \nu_F^0$ (while we have seen it to be regular at $\lambda = -B_F^0$ and $\lambda = 0$). The poles on the imaginary axis produce in the $\mathrm{Re}(\lambda) < 0$ half-plane corresponding zeros at approximately the same $\mathrm{Im}(\lambda)$ values (figure 4). The $\mathrm{Re}(\lambda)$ values of the zeros are determined by a sharp variation in the magnitude of the matrix elements Q_F^G, and thus in the

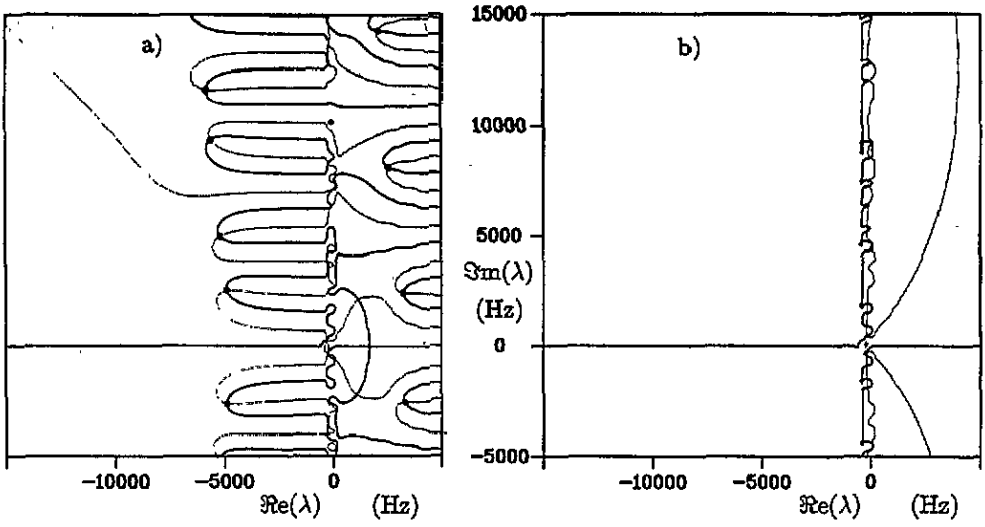

Figure 3. Gross structure of the spectrum. The solid and shaded curves represent the vanishing of, respectively, the imaginary and real part of equation (27). Zeros are indicated by full circles and poles by masked intersections. The features close to the imaginary axis are artifacts of the numerical program (see details in figure 4). (*a*) $\Delta t = 1$ ms and (*b*) $\Delta t = 0$ ms. Excitatory parameters: $V_E^0 = -73$ mV, $V_E^{\rm thr} = -53$ mV, $V_E^{\rm abp} = -63$ mV, $V_E^E = 0$ mV, $V_E^I = -75$ mV, $C_E/g_E^0 \equiv 1/\omega_E^0 = 15$ ms, $\omega_E^0 = 10\omega_E^E/3 = 3\omega_E^I/2 = \omega_E^S/2$, $\tau_E^E = \tau_I^E =$ ms, $S^E = \tau_E^E \times 50$ Hz. Inhibitory parameters: $V_I^0 = -70$ mV, $V_I^{\rm thr} = -55$ mV, $V_I^{\rm abp} = -70$ mV, $V_I^E = 0$ mV, $V_I^I = -75$ mV, $C_I/g_I^0 \equiv 1/\omega_I^0 = 5$ ms, $\omega_I^0 = \omega_I^E = 4\omega_I^I = 2\omega_I^S$, $\tau_E^I = \tau_I^I = 20$ ms, $S^I = \tau_I^E \times 50$ Hz. Stationary solution with rates $\nu_E^0 = 40.6$ Hz, $\nu_I^0 = 30.4$ Hz, unstable in (*a*) and—contrary to the expectations from the naive mean-field equations—stable in (*b*).

example of figure 4 lie close to the $\mathrm{Re}(\lambda) = -1/\tau^E$ and $\mathrm{Re}(\lambda) = -1/\tau^I$ lines where the real parts of the prefactors of Q_F^G change sign. This result, whose generality is yet unclear, has important implications: the *fine structure* of the spectrum, possibly away from pathological parameter regimes, *would consist of N_C approximately armonic series of oscillatory modes, with decay constants determined by the conductance inactivation time constants.*

It is interesting to compare the spectrum obtained here with that of a 'naive' mean-field theory, the one described by the system†

$$\begin{aligned} \dot{z}^E(t) &= -\frac{1}{\tau^E} z^E(t) + \nu_E^0(z^E(t), z^I(t)) \\ \dot{z}^I(t) &= -\frac{1}{\tau^I} z^I(t) + \nu_I^0(z^E(t), z^I(t)) \end{aligned} \tag{30}$$

where $\nu_F^0(z^E(t), z^I(t))$ is simply the stationary input–output relation, extended to be a function of time-varying arguments [25]. In this simplified system, the dynamics of individual spikes is neglected, by performing some ill-defined sort of temporal average alongside the spatial average of the mean-field treatment; it has been argued [26], though, that the reduction to the simplified system preserves some of the interesting dynamical features. The stability of stationary solutions of such naive mean-field equations [27] has been discussed repeatedly in the literature, and recently a more realistic form of $\nu_E^0(z^E(t), z^I(t))$ has been suggested, that ensures inhibitory control of excitatory firing

† Synaptic time constants are taken here to depend only on the presynaptic cell class.

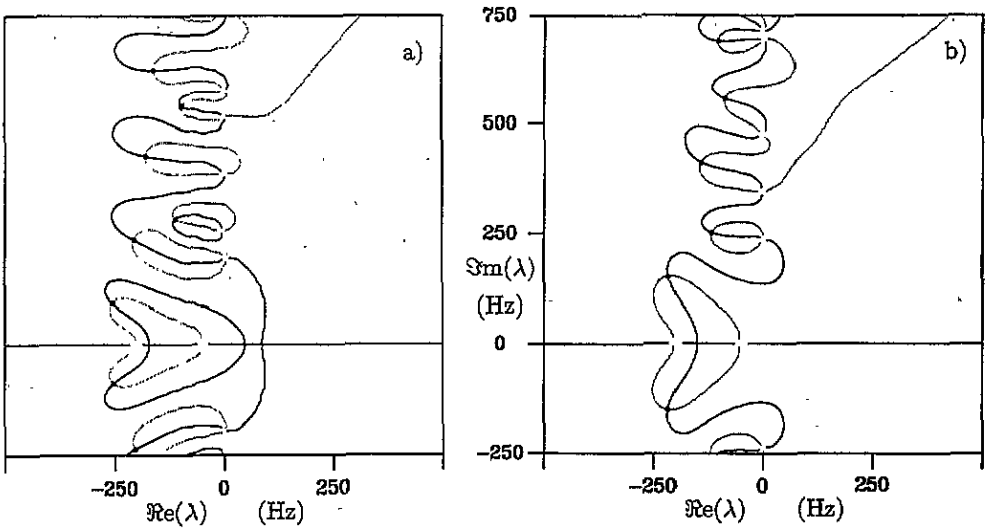

Figure 4. Fine structure of the spectrum. Notation as in figure 3. (*a*) Parameters as in figure 3 and (*b*) intraclass couplings modified to satisfy $\omega_E^E = \omega_E^0$, $\omega_I^I = 10\omega_I^0$, with in both cases $\Delta t = 0$. In (*b*) the stationary solution is again unstable (instability outside the graph) with rates $\nu_E^0 = 38.0$ Hz, $\nu_I^0 = 55.3$ Hz. Note how the stationary rates set the imaginary part of the zeros, while the real part is determined by the inverse conductance inactivation constants.

rates [28]. As the simplified equations follow the dynamics of the feedback fields, while the firing frequencies are described as adjusting themselves instantaneously to the correct level, these equations could expected to yield approximately the same zeros as (and thus be a valid approximation of) the correct mean-field theory under two conditions: (i) each class of cells must be able to adjust its frequency very rapidly by firing at a high rate, and thus must be well above threshold, $A_F^0 \gg B_F^0$ and (ii) the corresponding stationary state must be stable. In fact, it is easy to see that the naive spectrum is given by an equation analogous to equation (28), where the matrix $\mathbf{Q}(\lambda)$ is substituted by a simplified matrix $\tilde{\mathbf{Q}}(\lambda)$. Using equations (30), (15), (10) and (29) one finds

$$\begin{aligned}\tilde{Q}_F^G(\lambda) &= \frac{1}{\lambda + 1/\tau_F^G}\frac{\partial \nu_F^0}{\partial z^G(t)} \\ &= \frac{1/\tau_F^G}{\lambda + 1/\tau_F^G} Q_F^G(0)\end{aligned} \tag{31}$$

that is, the only λ-dependence left is the one in the prefactor. Obviously most of the spectrum is lost in the reduction to the naive theory, and therefore also the stability of the stationary solutions cannot be discussed in the simplified framework, as it requires knowledge of the full spectrum. Again the complex question of the stability will not be treated in this paper, but aspects of it are considered elsewhere [18].

3.4. Absence of periodic solutions

Consider, now, solutions with $A_E(t)$, $B_E(t)$, $A_I(t)$, $B_I(t)$ periodic with periods τ_E^S, τ_I^S respectively, as driven by periodic external stimuli $S_E(t)$, $S_I(t)$.

Take an *arbitrary* density distribution $\rho_F^0(x_0)$ at a given time t_0, continuous for $0 \leq x_0 < 1$, vanishing outside this interval, normalized and satisfying the boundary condition

of equation (13) at $t = t_0$; then the equations for $\dot{\rho}_F(x, t)$ are solved by the simple evolution

$$\rho_F(x(x_0, t), t) = \rho_F^0(x_0)(R_F(t_1))^{-\Theta(t-t_1)} \exp\left(\int_{t_0}^{t} dt' \, B_F(t')\right) \tag{32}$$

where, apart from the exponential modulation, the density just translates along the membrane potential axis according to

$$x(x_0, t) = x_0 \exp\left(-\int_{t_0}^{t} dt' \, B_F(t')\right) + \int_{t_0}^{t} dt' \left[A_F(t') \exp\left(-\int_{t'}^{t} dt'' \, B_F(t'')\right)\right] \tag{33}$$

until it reaches $x = 1$ at $t = t_1(x_0)$, and after that onwards:

$$x(x_0, t) = \int_{t_1}^{t} dt' \left[A_F(t') \exp\left(-\int_{t'}^{t} dt'' \, B_F(t'')\right)\right] \tag{34}$$

until it reaches again x_0 at time $t_0 + (\nu_F(x_0, t_0))^{-1}$. The reset factor

$$R_F(t) \equiv \frac{A_F(t)}{A_F(t) - B_F(t)} \tag{35}$$

ensures that the boundary condition is satisfied for $t > t_0$.

The most general distribution $\rho_F^0(x_0)$ generates a density $\rho_F(x(x_0, t), t)$ periodic in time with period τ_F^S *only* if $(\nu_F(x_0, t_0))^{-1} \equiv \tau_F^S$, and if in addition

$$R_F(t) = \exp \int_{t}^{t+\tau_F^S} dt' \, B_F(t') \equiv R_F^0 \quad \text{is constant.} \tag{36}$$

A periodic $\rho_F(x(x_0, t), t)$ would in turn generate functions $z_G^F(t)$ periodic with the same periods, and this leads to contradictions on two grounds: (i) the two frequencies, τ_E^S and τ_I^S, would be mixed up back in the $A_E(t)$, $B_E(t)$, $A_I(t)$, $B_I(t)$; (ii) even if $\tau_E^S = \tau_I^S = \tau^S$, this would imply $1/\tau^S = \nu_E = \nu_I$, violating the equations that ν_E and ν_I must satisfy, and that can be derived e.g. from equation (36), by Fourier expanding all the relevant quantities.

The particular distribution given by

$$\rho_F^0(x_0) = \frac{\nu_F \exp - \int_{t_0}^{t_1} dt' \, B_F(t')}{A_F(t_1(x_0)) - B_F(t_1(x_0))} \tag{37}$$

makes

$$[A_F(t) - B_F(t)] \, \rho_F(1^-, t) = \nu_F \tag{38}$$

a constant, and thus it would seem to produce constant fields $z_G^F(t)$, offering a way to avoid the above contradictions. Such a distribution does not, however, satisfy the boundary condition, because that would require that

$$\frac{\nu_F}{A_F(t_0) - B_F(t_0)} \frac{A_F(t_0) - B_F(t_0)}{A_F(t_0)} = \frac{\nu_F \exp\left[-\int_{t_0}^{t_0+(\nu_F)^{-1}} dt' \, B_F(t')\right]}{A_F(t_0 + (\nu_F)^{-1}) - B_F(t_0 + (\nu_F)^{-1})} \tag{39}$$

which is incompatible with the time course of $A_F(t)$, $B_F(t)$ being determined, in its non constant components, solely by $S_F(t)$.

Therefore, provided more than one class of cells is included, there exist no stationary or periodic responses, in the long time limit, to a periodic input: the behaviour of the system is more complex, possibly quasiperiodic or chaotic.

4. Dynamics with adaptation

Consider now the more general situation with $\omega_F^K \neq 0$.

It is now convenient to write equation (7) as

$$\dot{x}_i(t) = A_F(t) - x_i(t)B_F(t) + y_i(t)C_F - x_i(t)y_i(t)D_F \tag{40}$$

by introducing the additional notation

$$C_F = x_F^K \omega_F^K \qquad D_F = \omega_F^K. \tag{41}$$

Instead of $\rho_F(x, t)$, one should now consider a density $\pi_F(x, y, t)$. If, however, the initial conditions imply that

$$\pi_F(x, y, t_0) = \sigma_F(y, t_0)\delta(x - x_0(y))$$

then it is easy to see that x will remain linked to y throughout the dynamical evolution. With this restriction on the initial conditions (which will be irrelevant anyway later on, in extending the formalism to the case $N_C \to \infty$), we consider the evolution of the densities

$$\sigma_F(y, t) = \frac{1}{N_F}\sum_i \delta(y - y_i(t)). \tag{42}$$

In other words, the dynamics of the single cell is followed by monitoring the potassium conductance rather than the membrane potential, which is now considered, just for analytical convenience, a variable dependent on $y_i(t)$ (and t). The relationship between x and y is obtained by integrating

$$\frac{dx}{dy} = -\frac{\tau^K}{y}[A(t) - xB(t) + yC - xyD] \tag{43}$$

between spikes. The additional restriction that $0 < x < 1$, valid after a finite time, shall also be applied, for simplicity, straight away.

In terms of $\sigma_F(y, t)$, the dynamic system becomes

$$\begin{aligned}
\tau_F^K \dot{\sigma}_F(y, t) = {} & \sigma_F(y, t) + y\frac{\partial}{\partial y}\sigma_F(y, t) \\
& + \delta(y - \bar{y}_F(t) - 1)\left[\bar{y}_F(t) + 1 + \tau_F^K \dot{\bar{y}}_F(t)\right]\sigma_F(\bar{y}_F(t) + 1, t) \\
& - \delta(y - \bar{y}_F(t))\left[\bar{y}_F(t) + \tau_F^K \dot{\bar{y}}_F(t)\right]\sigma_F(\bar{y}_F(t), t) \\
\dot{z}_G^F(t) = {} & -\frac{1}{\tau_G^F} z_G^F(t) + \left[\frac{\bar{y}_F(t - \Delta t)}{\tau_F^K} + \dot{\bar{y}}_F(t - \Delta t)\right]\sigma_F(\bar{y}_F(t - \Delta t), t - \Delta t)
\end{aligned} \tag{44}$$

with the boundary conditions

$$\left[\bar{y}_F(t) + 1 + \tau_F^K \dot{\bar{y}}_F(t)\right]\sigma_F(\bar{y}_F(t) + 1, t) = \left[\bar{y}_F(t) + \tau_F^K \dot{\bar{y}}_F(t)\right]\sigma_F(\bar{y}_F(t), t) \tag{45}$$

where $\bar{y}_F(t)$ is the y-value at which cells of class F emit a spike, depending on previous history, as defined implicitly, integrating equation (43), by

$$\begin{aligned}
1 = {} & \int_{\bar{y}_F(t)}^{1+\bar{y}_F(t-T_F(t))} \tau_F^K \, dy \left[\frac{A_F(t - \tau_F^K \ln(y/\bar{y}_F))}{y} + C_F\right] \\
& \times \exp\left(-\int_{\bar{y}_F}^{y} \tau_F^K \, dy'\right)\left[\frac{B_F(t - \tau_F^K \ln(y'/\bar{y}_F))}{y'} + D_F\right]
\end{aligned} \tag{46}$$

while the instantaneous period $T_F(t)$ satisfies

$$T_F(t) = \tau_F^K \ln \frac{\bar{y}_F(t - T_F(t)) + 1}{\bar{y}_F(t)} \tag{47}$$

The simpler form of the dynamic system stems from the simplicity of the y-dynamics, while the complex x-dynamics remains hidden in $\bar{y}_F(t)$.

4.1. Stationary solutions

If the $S(t)$'s are constant in time, one has stationary solutions in the form

$$\sigma_F^0(y) = \frac{\nu_F^0 \tau_F^K}{y} \qquad z_G^{F0} = \tau_G^F \nu_F^0 \tag{48}$$

where the firing rates of each class of cells are

$$(\nu_F^0)^{-1} = T_F^0 = \tau_F^K \ln \frac{\bar{y}_F^0 + 1}{\bar{y}_F^0}. \tag{49}$$

and

$$1 = \int_{\bar{y}_F^0}^{1+\bar{y}_F^0} \tau_F^K \, dy \left[\frac{A_F^0}{y} + C_F \right] \left(\frac{\bar{y}_F^0}{y} \right)^{B_F^0 \tau_F^K} e^{-D_F \tau_F^K (y - \bar{y}_F^0)}. \tag{50}$$

The values of ν_E^0 and ν_I^0 can be obtained by solving the equations numerically having chosen values for the parameters.

The limiting hypersurface beyond which E and I tend to infinity is given now by

$$\frac{\nu_E^0}{\nu_I^0} \equiv r = \frac{\omega_E^E \tau_E^E r + \omega_E^I \tau_E^I + \omega_E^K \tau_E^K r}{\ln \widetilde{R}_E^0(r)} \qquad 1 = \frac{\omega_I^E \tau_I^E r + \omega_I^I \tau_I^I + \omega_I^K \tau_I^K}{\ln \widetilde{R}_I^0(r)} \tag{51}$$

with e.g.

$$\widetilde{R}_I^0(r) = \frac{x_I^E \omega_I^E \tau_I^E r + x_I^I \omega_I^I \tau_I^I + x_I^K \omega_I^K \tau_I^K}{(x_I^E - 1)\omega_I^E \tau_I^E r + (x_I^I - 1)\omega_I^I \tau_I^I + (x_I^K - 1)\omega_I^K \tau_I^K}. \tag{52}$$

The S-dependent planes beyond which either $\nu_E^0 = 0$ or $\nu_I^0 = 0$ are given, as for $\omega_F^K = 0$, by the conditions $A_E^0 = B_E^0$ and $A_I^0 = B_I^0$. The C and D terms are irrelevant, because $y_i \simeq 0$ most of the time when cell i fires at very low frequency.

4.2. Transients

The dynamics of the transients around a stable stationary solution can be analyzed along the same lines as previously. The only difference is that among the transient quantities (denoted, again, by keeping the same symbols, i.e. $\sigma_F(y, t) \to \sigma_F^0(y) + \sigma_F(y, t)$) one has to include now the fluctuation in the potassium conductance at spike emission, $\bar{y}_F(t)$. In expanding to first order in the transients the dynamic system equations (44), one finds that away from the singular points on the y-axis, the (subtracted) densities follow the simple equations of motion

$$\tau_F^K \dot{\sigma}_F(y, t) = \sigma_F(y, t) + y \frac{\partial}{\partial y} \sigma_F(y, t) \tag{53}$$

while for the feedback fields one has

$$\dot{z}_G^F(t) = -\frac{1}{\tau_G^F} z_G^F(t) + \frac{\bar{y}_F^0}{\tau_F^K} \sigma_F(\bar{y}_F^0, t - \Delta t) + \left[\frac{\bar{y}_F(t-\Delta t)}{\tau_F^K} + \dot{\bar{y}}_F(t-\Delta t) \right] \sigma_F^0(\bar{y}_F^0) + \frac{\bar{y}_F^0}{\tau_F^K} \left[\sigma_F^0(\bar{y}_F^0 + \bar{y}_F(t-\Delta t)) - \sigma_F^0(\bar{y}_F^0) \right]. \tag{54}$$

The simplicity of equation (53), due to the simple relaxation dynamics of the potassium conductance, is compounded by the presence of discontinuities at $y = \bar{y}_F^0$, $\bar{y}_F^0 + \bar{y}_F(t)$, $\bar{y}_F^0 + 1$, $\bar{y}_F^0 + \bar{y}_F(t) + 1$, and by the necessity to include among the deviations from the equilibrium densities, i.e. $\sigma_F(y, t)$, alongside a term of transient amplitude on the interval $(\bar{y}_F^0 + \bar{y}_F(t), \bar{y}_F^0 + \bar{y}_F(t) + 1)$, a finite term $-\sigma_F^0(y)\text{sign}(\bar{y}_F(t))$ on the transient intervals $(\bar{y}_F^0, \bar{y}_F^0 + \bar{y}_F(t))$ and $(\bar{y}_F^0 + 1, \bar{y}_F^0 + \bar{y}_F(t) + 1)$ at the boundaries. Taken together these three contributions satisfy

$$\int \mathrm{d}y\, \sigma_F(y, t) = 0. \tag{55}$$

Considering again transients with a time dependence $\exp(\lambda t)$, one finds solutions for the deviations in the densities of the type

$$\sigma_F(y, t) = \frac{\lambda \tau_F^K \sigma_F^0(y) y^{\lambda \tau_F^K}}{(\bar{y}_F^0 + 1)^{\lambda \tau_F^K} - (\bar{y}_F^0)^{\lambda \tau_F^K}} \left[\frac{1}{\bar{y}_F^0} - \frac{1}{\bar{y}_F^0 + 1} \right] \bar{y}_F(t) \tag{56}$$

on top of the finite terms just mentioned. $\bar{y}_F(t)$ can in turn be related to the transient rise rates A, B through equation (46)

$$\bar{y}_F(t) = \int_{\bar{y}_F^0}^{1+\bar{y}_F^0} \mathrm{d}y \left(\frac{\bar{y}_F^0}{y} \right)^{(\lambda + B_F^0)\tau_F^K + 1} \mathrm{e}^{-D_F \tau_F^K (y - \bar{y}_F^0)} \times \left\{ A_F(t) + B_F(t) \frac{[A_F^0 + y C_F]}{\lambda} \left[1 - \left(\frac{y}{\bar{y}_F^0} \right)^{\lambda \tau_F^K} \right] \right\} \times \left\{ A_F^0 - B_F^0 + \bar{y}_F^0 (C_F - D_F) - \mathrm{e}^{-T_F^0 (\lambda + B_F^0 + 1/\tau_F^K) - D_F \tau_F^K} [A_F^0 + C_F (1 + \bar{y}_F^0)] \right\}^{-1}. \tag{57}$$

These equations yield a relation between the transient amplitudes of the fields $F_G(t)$, finally resulting, with the same derivations as above, into an equation for λ formally identical to equation (27), but with $Q_F^G(\lambda)$ now defined as

$$Q_F^G(\lambda) = \frac{\nu_F^0 \omega_F^G}{(\lambda + 1/\tau_F^G)} \left[1 + \frac{1}{(1 + \bar{y}_F^0)(\mathrm{e}^{\lambda T_F^0} - 1)} \right] \times \left\{ \int_{\bar{y}_F^0}^{1+\bar{y}_F^0} \tau_F^K \mathrm{d}y \left[\frac{A_F^0}{y} + C_F + \frac{\lambda x_F^G}{y} \right] \left(\frac{\bar{y}_F^0}{y} \right)^{(\lambda + B_F^0)\tau_F^K} \mathrm{e}^{-D_F \tau_F^K (y - \bar{y}_F^0)} - 1 \right\} \times \left\{ A_F^0 - B_F^0 + \bar{y}_F^0 (C_F - D_F) - \mathrm{e}^{-T_F^0 (\lambda + B_F^0 + 1/\tau_F^K) - D_F \tau_F^K} [A_F^0 + C_F (1 + \bar{y}_F^0)] \right\}^{-1}. \tag{58}$$

It can be checked that $Q_F^G(\lambda)$ reduces to the previous definition of equation (26) as $C_F, D_F \to 0$.

4.3. *The peculiar case of a homogeneous network of identical excitatory cells*

If *all* cells are identical, excitatory and fully interconnected with synapses of uniform strength, and if the net receives uniform external stimuli constant in time, then—and only in this artificially simple case—it may go into a mode of periodic oscillations. In fact one can write down, in this particular case, long-time solutions of the dynamics corresponding to fully synchronized spiking by all the cells. One can then find the limiting conditions (phase boundaries, or instability lines) beyond which such synchronized solutions destabilize into static solutions, or vice versa the limit beyond which static solutions destabilize and begin to synchronize. Note that there may exist different periodic solutions, e.g. corresponding to several groups comprising an equal number of cells, each firing synchronously at equally spaced intervals [29], but it is likely that the fully synchronous solutions be the most robust of all these, and it may be convenient to consider them first.

In studying a net of identical excitatory cells one may determine how the available long-term solutions of its dynamics vary with the parameters, e.g. by singling out as independent variables ω^E, measuring the coupling, and ω^{K}, proportional to the size of the potassium conductance (with a single cell class subscripts may be dropped). For ω^E above a given $\omega^E_{\max}$ the firing rate in the static, asynchronous solution goes to infinity, invalidating the approximation of neglecting the absolute refractory period. $\omega^E_{\max}$ can be easily found from equation (16) (by eliminating inhibitory terms) as the solution of

$$1 = \frac{1}{\tau^E \omega^E_{\max} + \tau^{\mathrm{K}} \omega^{\mathrm{K}}} \ln \frac{x^E \tau^E \omega^E_{\max} + x^{\mathrm{K}} \tau^{\mathrm{K}} \omega^{\mathrm{K}}}{(x^E - 1)\tau^E \omega^E_{\max} + (x^{\mathrm{K}} - 1)\tau^{\mathrm{K}} \omega^{\mathrm{K}}}. \tag{59}$$

The asynchronous solution, however, destabilizes already when the first transient eigenvalue appears, with $\mathrm{Re}(\lambda) > 0$. If $\mathrm{Im}(\lambda) = 0$, this would occur at the transition line $Q^E(0) - 1 = 0$, or in terms of the parameters of the static solution

$$\frac{(\nu^0)^2 \omega^E \tau^E}{1 + \bar{y}_F^0} \int_{\bar{y}^0}^{1+\bar{y}^0} \tau^{\mathrm{K}} \, \mathrm{d}y \left[\frac{x^E}{y} - \left(\frac{A^0}{y} + C \right) \tau^{\mathrm{K}} \ln \frac{y}{\bar{y}^0} \right] \left(\frac{\bar{y}^0}{y} \right)^{B^0 \tau^{\mathrm{K}}} \mathrm{e}^{-D\tau^{\mathrm{K}}(y - \bar{y}^0)}$$
$$\times \left\{ A_F^0 - B_F^0 + \bar{y}_F^0 (C_F - D_F) \right.$$
$$\left. - \mathrm{e}^{-T_F^0(\lambda + B_F^0 + 1/\tau_F^{\mathrm{K}}) - D_F \tau_F^{\mathrm{K}}} [A_F^0 + C_F(1 + \bar{y}_F^0)] \right\}^{-1} = 1. \tag{60}$$

However, both the nature of the instability (and hence how it arises in λ-space) and whether it produces a flow towards a stable synchronous solution or towards something else, represent in general extremely complex issues, which may have very different answers in different parameter regimes. Therefore these issued will not be pursued here any further, but will be addressed in another paper [18].

5. Extensions to more interesting cases

5.1. *Variability within classes*

A slightly more complex case, which may be considered next, is that of a network in which excitation and inhibition can still be described exactly by just two macroscopic fields, but there is now variability within each class of cells, both in their intrinsic parameters and in

the way they are affected by the macroscopic fields. In other words, the rise rates of the membrane potential for each cell can still be expressed in terms of the fields $z_G^F(t)$, but each cell is now characterized by its own set of 'coupling' parameters ω_i. One deals with such a case by introducing ω-dependent densities $\rho_F(\omega, x, t)$ such that

$$1 = \int d\omega\, \rho_F(\omega) = \int d\omega \left[\int dx\, \rho_F(\omega, x, t) \right] \tag{61}$$

if we take, for simplicity, the situation with no adaptation, $\omega^K = 0$. The above methods can easily be extended to cover this case, and, for example, the static solutions are now given by the equations

$$\rho_F^0(\omega, x) = \frac{\nu_F^0(\omega)}{A_F^0(\omega) - x B_F^0(\omega)} \rho_F(\omega) \qquad F_G^0 = \int d\omega\, \tau_G^F \nu_F^0(\omega) \rho_F(\omega) \tag{62}$$

where the firing rates of each class of cells are

$$\nu_F^0(\omega) = B_F^0(\omega) \left[\ln \frac{A_F^0(\omega)}{A_F^0(\omega) - B_F^0(\omega)} \right]^{-1}. \tag{63}$$

5.2. More than two classes

A more general and straightforward approach, however, is to consider directly the case in which there is a large number N_C of different classes of cells, each having its own independent synaptic coupling with each other class. Some of the parameters characterising different classes may of course be equal, for example the synaptic time constants may be common to many pre- and post-synaptic pairs of classes.

Again, it is easy to generalize the previous formulas. One should note that there are several possible static solutions, that correspond, in particular, to different combinations of classes of cells being quiescent in the long-time limit. Non-firing classes produce no fields, and therefore the corresponding equations drop out and are replaced by inequalities (ensuring that the membrane potentials for cells of those classes are kept below threshold), while the coupled system takes a different form with each combination of surviving equations. If N out of N_C classes comprise cells that fire in a particular static solution, the solution itself, in the general case of $\omega_F^K \neq 0$, is given by equations (48), with the firing rates determined by the system of $2 \times N$ equations

$$(\nu_F^0)^{-1} = T_F^0 = \tau_F^K \ln \frac{\bar{y}_F^0 + 1}{\bar{y}_F^0}. \tag{64}$$

and

$$1 = \int_{\bar{y}_F^0}^{1+\bar{y}_F^0} \tau_F^K dy \left[\frac{A_F^0}{y} + C_F \right] \left(\frac{\bar{y}_F^0}{y} \right)^{B_F^0 \tau_F^K} e^{-D_F \tau_F^K (y - \bar{y}_F^0)}. \tag{65}$$

Here the rise rates stand for

$$\begin{aligned} A_F^0 &= x_F^0 \omega_F^0 + \sum_{G=1}^{N} x_F^G \omega_F^G \tau_F^G \nu_G^0 + x_F^S \omega_F^S S_F^0 \\ B_F^0 &= \omega_F^0 + \sum_{G=1}^{N} \omega_F^G \tau_F^G \nu_G^0 + \omega_F^S S_F^0. \end{aligned} \tag{66}$$

Stable static solutions correspond to attractors, and the spectrum of time constants characterising the transient behaviour converging to the attractor is given again by the single equation

$$\left|\mathbf{Q}(\lambda) - e^{\lambda \Delta t}\mathbf{1}\right| = 0 \tag{67}$$

where the matrix elements of $\mathbf{Q}$ are given again by equations (58).

A particular case to which the above treatment can be applied, is when the number of classes N_C is not much larger than, or it is even equal to, the number of cells. In that case, the method is not exact (it would be only in the limit of very many cells per class), but may be regarded as a mean-field approximation, in which each cell is replaced by a continuous distribution of cells of equal characteristics and couplings, spread over different values of x (or y). Note that this is *not* equivalent to the approximation, discussed above, in which the discontinuous nature of the spiking dynamics is neglected in favour of a continuous description in terms of average rates. Here, every ingredient of the original integrate-and-fire dynamics is preserved, with its full dependence of the different time and coupling parameters. The present approximation is, rather, equivalent to assuming an effective indeterminacy in the establishment of initial conditions (namely, the value of the membrane potential) for each cell in the network; however, different evolution histories are not simply superimposed, because they interact with each other.

5.3. Associative memory

One type of network in which cells are grouped into a large number of classes (possibly as large as the number of cells) on the basis of the strengths of their mutual connections is an autoassociative memory. Consider for example the case of a network of N_C interconnected excitatory cells, of identical characteristics, with couplings given by the matrix

$$\omega_{ij} = k\frac{\omega_0}{N_C} + k\frac{\omega_1}{N_C}\sum_{\mu=1}^{p}\left(\frac{\eta_i^\mu}{a} - 1\right)\left(\frac{\eta_j^\mu}{a} - 1\right) \tag{68}$$

where k, ω_0 and ω_1 are given parameters, and the η_i^μ are normalized firing rates assigned, independently for each μ and i, from a common probability distribution P_η, such that $a = \langle\eta\rangle_\eta = \langle\eta^2\rangle_\eta$ (a is then a *sparse coding* parameter, see [30]). These connections endow the network with a set of attractor states correlated each with one of the p 'stored patterns' $\{\eta_i^\mu\}$. One could try to analyse the attractor structure emerging from the integrate-and-fire model. In this model, however, the resulting current-to-frequency transduction function characterizing single cells in the long-time limit is rather similar, as shown in figure 2, to a simple threshold-linear function; a similarity enhanced by the adaptation effect of the potassium conductance. One can expect, therefore, the attractor states to be close to those obtained with a static treatment based on the use of threshold-linear units, as I have presented elsewhere [31]. The new, dynamical, results that emerge from the present treatment, instead, are those concerning the transients and their time constants, and these are the ones described in the following.

The transient time constants are given by the the inverse of the values λ which satisfy

$$\Delta = \left|\mathbf{Q}(\lambda) - e^{\lambda \Delta t}\mathbf{1}\right| = 0 \tag{69}$$

with $\mathbf{Q}(\lambda)$ a $N \times N$ matrix of the form

$$Q_{ij}(\lambda) = \omega_{ij} q_j(\lambda) \tag{70}$$

which results from the assumptions that all cells (classes) have identical characteristics but for the couplings ω_{ij}. The form of $q_j(\lambda)$ can be extracted from equation (58). N is the number of active cells (quiescent ones drop out of the transient equations as well) and it is convenient to set $k = N_C/N$.

The zeros of the determinant could be obtained by expanding the inverse of the determinant (for fixed λ close but not equal to one of the critical values) into a Gaussian integral, evaluating it at the saddle point, and checking when it diverges as λ varies. One has to keep in mind, however, that the location of the zeros will depend on the (quenched) assignment of patterns $\{\eta_i^\mu\}$. In order to obtain a quenched average, based on the assumption that the extensive, self-averaging, quantity is $\ln \Delta$, one uses the standard trick of considering n-replicas and then letting $n \to 0$:

$$f \equiv -\frac{1}{N}\langle \ln \Delta \rangle_\eta = \lim_{n\to 0} \frac{2}{Nn}\left[\langle \Delta^{-n/2}\rangle_\eta - 1\right]. \tag{71}$$

A replica-symmetric evaluation of f is presented in the appendix. The result is

$$f = -\widetilde{\alpha} \ln \frac{\widetilde{\alpha} T_0 \omega_1}{r_0} + \widetilde{\alpha} - \frac{r_0}{T_0\omega_1} - \left\langle \ln(r_0 q_i(\lambda) - e^{\lambda \Delta t})\right\rangle + \mathcal{O}(1/N) \tag{72}$$

where the order parameter r_0 is determined by the saddle point equation

$$\widetilde{\alpha} = \frac{r_0}{T_0\omega_1} + \left\langle \frac{r_0 q_i(\lambda)}{r_0 q_i(\lambda) - e^{\lambda \Delta t}} \right\rangle \tag{73}$$

and where $\widetilde{\alpha} \equiv (p-1)/N$ must satisfy $\widetilde{\alpha} < \widetilde{\alpha}^*$ (see the appendix).

The zeros of the determinant correspond to the logarithmic singularities in equation (72). Note that there are N sets of potential singularities associated with the N conditions $r_0 q_i(\lambda) - e^{\lambda \Delta t} = 0$, yielding the decay rates of the transient activity of each cell, plus $p-1$ degenerate sets associated with the condition $\widetilde{\alpha} T_0 \omega_1 / r_0 = 0$, yielding the decay rates of transients correlations with patterns not being retrieved. The most interesting set of singularities, however, is the one associated with the transients in the correlation with the pattern being retrieved, and this has remained hidden in the $\mathcal{O}(1/N)$ corrections to equation (72). Going back to equation (86), one finds that this additional set of singularities is given by a vanishing argument of the logarithm in the last three lines of that equation, which reads

$$\left[\left(1 + \omega_0 \left\langle \frac{q_i(\lambda)}{r_0 q_i(\lambda) - e^{\lambda \Delta t}} \right\rangle\right)\right. \tag{74}$$

$$\times \left(1 + \omega_1 \left\langle \left(\frac{\eta^1}{a} - 1\right)^2 \frac{q_i(\lambda)}{r_0 q_i(\lambda) - e^{\lambda \Delta t}} \right\rangle\right) \tag{75}$$

$$\left. - \omega_0 \omega_1 \left\langle \left(\frac{\eta^1}{a} - 1\right) \frac{q_i(\lambda)}{r_0 q_i(\lambda) - e^{\lambda \Delta t}} \right\rangle^2 \right] = 0$$

in the replica-symmetric solution. This condition yields the λ-values which are probably most interesting from the point of view of the time course of information processing.

A simple case in which the calculation can be brought forward analytically, is that of a retrieval state in which all active cells fire at the same (normalized) rate $\eta = 1$ in the stored pattern, and also (neglecting fast and quenched sources of noise) receive the same input, and hence fire at the same rate, in the retrieved pattern, implying $q_i(\lambda) = q(\lambda)$ for all i. Then r_0 is given by

$$r_0 = \frac{1}{2}\left[\frac{e^{\lambda \Delta t}}{q(\lambda)} + (\widetilde{\alpha} - 1) T_0 \omega_1\right]\left\{1 - \sqrt{1 - \frac{4\widetilde{\alpha} e^{\lambda \Delta t} T_0 \omega_1 q(\lambda)}{[e^{\lambda \Delta t} + (\widetilde{\alpha} - 1) T_0 \omega_1 q(\lambda)]^2}}\right\} \tag{76}$$

(the correct solution of the quadratic equation being determined by the requirement that, in the $\tilde{\alpha} \to 0$ limit, $r_0 \propto \tilde{\alpha}$). Note also that averaging over active cells only $\langle(\eta/a - 1)\rangle = T_0$ (cf the appendix). Substituting for r_0, one finds that the first two types of singularity both correspond to the condition $q(\lambda) = 0$, that $\tilde{\alpha}^* = \{1 \pm [e^{\lambda\Delta t}/T_0\omega_1 q(\lambda)]^{1/2}\}^2$, and that the decay times of the transients in the overlap with the retrieved pattern are given by the equation, algebraically derived from equation (75)

$$q(\lambda)(\omega_0 + T_0^2\omega_1)[\omega_0 + (T_0 + \tilde{\alpha} - 1)T_0\omega_1] - [\omega_0 + T_0(T_0 - 1)\omega_1]e^{\lambda\Delta t} = 0. \tag{77}$$

Finally, as $\alpha \to 0$, the condition giving the last set of λ's reduces to

$$q(\lambda)(\omega_0 + T_0^2\omega_1) - e^{\lambda\Delta t} = 0. \tag{78}$$

5.4. The spectrum with adaptation

Adaptation has a marked effect on the λ-spectrum, and in order to illustrate it, it is convenient to use, rather than the equation for the network of excitatory and inhibitory cells, with its double sets of time scales, the formulas just derived for the associative memory net, where we have assumed all cells to share the same time parameters. These formulas have very limited applicability, because of the several simplificatory assumptions made. Thus, having neglected to include inhibitory cells that control the overall firing rate [31, 28], for example, renders a discussion of the *gross* structure of the spectrum, and its implications on the stability of stationary solution, rather academic. On the other hand, these formulas, in their simplicity, serve to bring out the features implicit in the new form of the Q-matrix (the one including adaptation, equation (58)) and in particular the emergence of a new set of very long transient time constant, which may be called the *hyperfine* structure of the spectrum.

Figure 5(a) shows, for one particular numerical example, the zeros of the $q(\lambda)$ factor. These have been related above to the transients in the overlap between the state of an autoassociative net and previously stored firing patterns *not* being retrieved. As in figure 4, there is a string of poles on the imaginary axis at $\lambda = \pm i2\pi n\nu^0$, $n \neq 0$, a pole on the real axis at $\lambda = -1/\tau^E$, and a set of zeros (the *fine* structure) approximately allineated with the string of poles in the $\mathrm{Im}(\lambda)$ dimension, and with a negative real component of the same order (the inverse conductance decay time) as the pole on the real axis. In addition, however, one observes a new set of poles and correspondingly, a new set of zeros, both occurring with the same periodicity $\Delta\lambda = i2\pi\nu^0$, and this time with one of each on the real axis as well. The zeros have real part $\mathrm{Re}(\lambda) = -1/\tau^K$ and correspond to the vanishing of the term in square brackets in the first line of equation (58), while the poles have real part

$$\mathrm{Re}(\lambda) = \nu^0\left[\ln\frac{A^0 + C(1 + \bar{y}^0)}{A^0 - B^0 + \bar{y}^0(C - D)} - D\tau^K\right] - B^0 - \frac{1}{\tau^K}$$

and correspond to the vanishing of the denominator (i.e. the last two lines) of equation (58). Such poles and zeros are superimposed at $\mathrm{Re}(\lambda) = -1/\tau^K$, and therefore cancel each other, in the absence of the potassium conductance, $\omega^K = 0$ (as can be checked by setting $C = D = 0$), but decouple and get farther apart as ω^K grows. As soon, then, as adaptation effect enter the game, of the type produced by the single (slow) potassium conductance considered here, a set of slow transients arise, including one non-oscillatory and an harmonic set of oscillatory modes.

One may consider now an equation of the form $q(\lambda)\tilde{\omega}^E - 1 = 0$, which has been related above to the transients in the overlap between the state of an autoassociative net and the

one firing pattern being retrieved, in the case of zero transmission delay† ($\widetilde{\omega}^E$ denotes the renormalized coupling resulting from the replica calculation). As can be seen in figure 5(b), the subtraction of the constant term, while obviously leaving the poles fixed, also preserves the essential structure the spectrum. One notes that the imaginary parts of both sets of zero, of the fine and hyperfine structure, are slightly reduced (preserving the basic periodicity), while the real parts are also reduced, and tend to approach the nearest pole, the faster the farther apart they are from the real axis. The single zero on the real axis, which was at $\lambda = -1/\tau^{\mathrm{K}}$ is now displaced, in the particular example in figure 5(b), into the positive real axis, making the stationary state considered unstable to a very slow purely exponential decay. In the example of figure 5(c), the parameter ω^{K} is reduced to 2/5 of what it was in figure 5(b). As a result of the reduced adaptation, the firing rate increases (and therefore the distance between the zeros in the imaginary dimension, roughly $2\pi\nu^0$, increases), and also the new set of poles is closer in the real dimension to the corresponding zeros of the hyperfine structure (onto which the poles would superimpose for $\omega^{\mathrm{K}} = 0$). Again, the zeros of the fine structure start off near the real axis with a real part close to that of the pole at $-1/\tau^E$, and tend to have a smaller (less negative) real part as their imaginary part increases. The single zero on the real axis remains, in this particular case, on the left (negative) side of the origin, leaving the solution stable to slow instabilities.

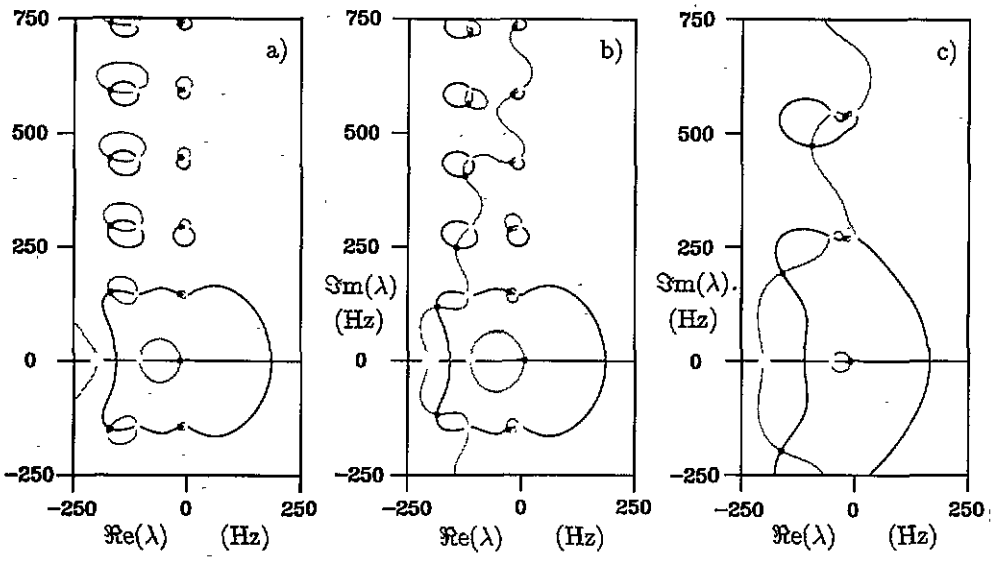

Figure 5. Fine and hyperfine structure of the spectrum, with adaptation effects included. (a), (b) $\tau^{\mathrm{K}} = 50$ ms and $\omega^{\mathrm{K}} = 12.5$ Hz; (c) $\omega^{\mathrm{K}} = 5$ Hz. Notation as in figure 3, with the shaded and solid curves curves giving the zeros of, respectively, the real and imaginary part of (a) $q(\lambda)$ and of (b), (c) $q(\lambda)\widetilde{\omega}^E - 1$, with $\widetilde{\omega}^E = 60$ Hz. Other voltage and coupling parameters as in figure 1 and figure 3 (excitatory cells) except for: (i) the inhibitory rate is fixed at $\nu_I^0 = 50$ Hz (ii) the excitatory coupling is set *a posteriori* at $1/\omega^E = 15$ ms ($\nu_E^0/50$ Hz). Stationary firing rates (a), (b) $\nu_E^0 = 23.5$ Hz and (c) $\nu_E^0 = 43.4$ Hz, reduced from $\nu_E^0 = 92.1$ Hz in the absence of adaptation. In (a) the zeros of the hyperfine structure have real part -20 Hz $\equiv -1/\tau^{\mathrm{K}}$.

† Having $\Delta t \neq 0$ would alter the stability of the stationary solution and the gross structure of the spectrum, but not the fine and hyperfine structure.

6. Implications for neocortical processing

An aim of this paper is to suggest a particular level of neuronal modelling as both amenable to analytical exploration *and* relevant to a discussion of real neuronal dynamics in the cortex. An obvious set of issues to which the method seems applicable is temporal encoding. In this direction, all the work has still to be done, but one result produced here may serve as a reminder. Whereas several interesting results on the problem of synchronicity have been obtained in extremely simplified conditions [15, 16, 5], these may easily not survive in a more realistic dynamics of the type introduced here; moreover, just including inhibitory alongside excitatory cells, or even a nonhomogeneous set of excitatory cells, is enough to prevent *any* type of synchronous periodic solution. The problem, then, becomes, in any minimally realistic context, one of characterising chaotic behaviour in a computationally meaningful way and with a reasonable dynamics, a programme that is still in its infancy [32].

In discussing transients towards stationary solutions, instead, therefore in situations in which it is a good approximation to consider those solutions as both relevant and stable, and external inputs as fixed in time after onset, the method already yields a way of relating the transient time scales to the biophysical parameters introduced in the model. Fast time parameters, such as axonal conduction times and synaptic delays, may have an important effect on stability but, within a stable solution, only result in extremely rapid transients—the gross structure of the spectrum. Conductance inactivation times have been shown to play the main role: those corresponding to synaptic conductances in determining the relatively fast transients of the fine structure, and that of the intrinsic (adapting) potassium conductance in determining the slow transients of the hyperfine structure. The very many other parameters present are crucial in determining the stationary firing rate of the cells, but have only a marginal influence on the transients—e.g. in setting, through the firing rate, the frequencies of the oscillatory modes.

In recent experiments in the inferotemporal cortex of primates [10], the temporal course of single cells reponses to complex visual stimuli (faces) was analysed, in particular with regard to the information content of the responses. One of the results was that most of the information content present in the full temporal train of responses was already expressed by the simple firing rate; further, most of the information in the firing rate could be extracted even by computing it over very short (20–50 ms) time intervals after response onset. In suggesting that the response relevant to further stages of cortical processing is attained very rapidly, these results raise the issue of whether there is enough time for local cortical circuits operating through recurrent (as opposed to feedforward) connections, to intervene (as the stimulus arrives locally) and affect the response. The analysis presented here indicates that recurrent processing may indeed very rapidly affect the response, because part of the transient modes, resulting from recurrent processing, are those whose time scales are determined by the rather rapid ($\approx$ 5–10 ms) excitatory conductance inactivation time†. Adaptation effects instead, which here have been modelled with a single conductance, but which in reality result from a complex set of conductances with time scales extending over hundreds of milliseconds result in transients which seem to be too slow. The question then becomes that of understanding how the information carried by the neuronal responses reflects the time course of different modes; a question that can be addressed in a formal model, by extending the dynamics of firing rates into considering the dynamics of information.

† A more complete discussion of the implications of the results presented here, and of their seemingly counterintuitive nature, will be given elsewhere [19].

Acknowledgments

The work reported here is part of a project pursued together with Larry Abbott, and originating from a visit in April 1991 to the Center for Complex Systems at Brandeis University, the hospitality of which is gratefully acknowledged. Our formal approach has been developed in parallel with a set of experiments on the temporal course of cortical cells responses, carried out in Edmund Rolls' lab in Oxford. My financial support was from an EC BRAIN grant to a collaboration organized by Edmund Rolls. Discussions with Ton Coolen, Daniel Amit and Misha Tsodyks have also been helpful.

Appendix

In evaluating equation (71), one notes that the values of λ may be found after changing the vectors upon which $\mathbf{Q}$ operates

$$\tilde{w}_i \to w_i \equiv \sqrt{q}_i(\lambda)\tilde{w}_i \tag{79}$$

so that $\mathbf{Q}$ is transformed into the symmetric matrix $\sqrt{q}_i(\lambda)\omega_{ij}\sqrt{q}_j(\lambda)$. $\sqrt{q}_i(\lambda)$ depends on the integrated inputs to unit i, through $A_i(t)$ and $B_i(t)$, and hence on the quenched assignment of the $\{\eta_i^\mu\}$. Considering however an attractor state correlated with a single pattern η^1, we shall neglect the residual dependence on the other $p-1$ patterns, and think of $\sqrt{q}_i(\lambda)$ as dependent only on η_i^1. Then, labelling $\gamma, \delta, \ldots$ the different replicas

$$\Delta^{-n/2} = \int \prod_{i\gamma} \frac{\mathrm{d}w_i^\gamma}{(2\pi)^{1/2}} \exp -\frac{1}{2}\sum_{ij\gamma} w_i^\gamma [\sqrt{q}_i(\lambda)\omega_{ij}\sqrt{q}_j(\lambda) - \mathrm{e}^{\lambda\Delta t}\delta_{ij}]w_j^\gamma. \tag{80}$$

Substituting for ω_{ij} and inserting the order parameters

$$u^\gamma = \frac{1}{N}\sum_i \sqrt{q}_i(\lambda)w_i^\gamma \qquad u_\mu^\gamma = \frac{1}{N}\sum_i (\frac{\eta_i^\mu}{a} - 1)\sqrt{q}_i(\lambda)w_i^\gamma \tag{81}$$

through δ-functions, one obtains

$$\begin{aligned}\Delta^{-n/2} = \int \prod_\gamma \left[\frac{\mathrm{d}u^\gamma \mathrm{d}t^\gamma}{2\pi N}\prod_\mu \frac{\mathrm{d}u_\mu^\gamma \mathrm{d}t_\mu^\gamma}{2\pi N}\prod_i \frac{\mathrm{d}w_i^\gamma}{(2\pi)^{1/2}}\right] \\ \times \exp\left\{-\frac{1}{2}\sum_\gamma\left[N\omega_0(u^\gamma)^2 + N\omega_1\sum_\mu (u_\mu^\gamma)^2 - \mathrm{e}^{\lambda\Delta t}\sum_i (w_i^\gamma)^2\right]\right\} \\ \times \exp\left\{\mathrm{i}\sum_\gamma\left[t^\gamma\left(Nu^\gamma - \sum_i \sqrt{q}_i(\lambda)w_i^\gamma\right)\right.\right. \\ \left.\left. + \sum_\mu t_\mu^\gamma\left(Nu_\mu^\gamma - \sum_i\left(\frac{\eta_i^\mu}{a} - 1\right)\sqrt{q}_i(\lambda)w_i^\gamma\right)\right]\right\}.\end{aligned} \tag{82}$$

Performing a quenched average $\langle\cdot\rangle_\eta$ over the distribution P_η of stored patterns, leaving out one pattern correlated with the attractor state (labelled pattern 1), the last exponent in the third line is replaced by

$$\exp -\frac{T_0}{2}\sum_{\gamma\delta}\sum_{\mu>1} t_\mu^\gamma t_\mu^\delta \sum_i q_i(\lambda)w_i^\gamma w_i^\delta \tag{83}$$

where $T_0 \equiv (\langle \eta^2 \rangle_\eta - \langle \eta \rangle_\eta^2)/\langle \eta \rangle_\eta^2 = (1-a)/a$ [31]. Having introduced as a further order parameter the $n \times n$ matrix $\mathbf{s}$ of elements

$$s^{\gamma\delta} = \frac{1}{N} \sum_i q_i(\lambda) w_i^\gamma w_i^\delta, \tag{84}$$

one may perform the Gaussian integrals over u_μ, t_μ for $\mu > 1$, yielding

$$\begin{aligned}
\langle \Delta^{-n/2} \rangle_\eta = \int \prod_\gamma \left[\frac{\mathrm{d}u^\gamma \mathrm{d}t^\gamma}{2\pi N} \frac{\mathrm{d}u_1^\gamma \mathrm{d}t_1^\gamma}{2\pi N} \prod_i \frac{\mathrm{d}w_i^\gamma}{(2\pi)^{1/2}} \right] \prod_{\gamma\delta} \frac{\mathrm{d}s^{\gamma\delta} \mathrm{d}r^{\gamma\delta}}{2\pi N} \\
\times \exp\left\{ -\frac{N}{2} \sum_\gamma \left[\omega_0 (u^\gamma)^2 + \omega_1 (u_1^\gamma)^2 \right] \right\} \\
\times \exp\left\{ -\frac{1}{2} \left[(p-1) \mathrm{Tr}_\gamma \ln(\mathbf{1} + T_0 \omega_1 \mathbf{s}) - \mathrm{e}^{\lambda \Delta t} \sum_{i\gamma} (w_i^\gamma)^2 \right] \right\} \\
\times \exp\left\{ N \mathrm{i} \left[\sum_\gamma (t^\gamma u^\gamma + t_1^\gamma u_1^\gamma) + \sum_{\gamma\delta} s^{\gamma\delta} r^{\gamma\delta} \right] \right\} \\
\times \exp\left[\left[-\mathrm{i} \left\{ \sum_{i\gamma} \left[t^\gamma \sqrt{q_i}(\lambda) w_i^\gamma + t_1^\gamma \left(\frac{\eta_i^1}{a} - 1 \right) \sqrt{q_i}(\lambda) w_i^\gamma \right] \right.\right.\right. \\
\left.\left.\left. + \sum_{i\gamma\delta} r^{\gamma\delta} q_i(\lambda) w_i^\gamma w_i^\delta \right\} \right].
\end{aligned} \tag{85}$$

At this point, one can do the Gaussian w_i^γ integrals, followed by the Gaussian integrals over u^γ, u_1^γ, t^γ, t_1^γ. Rotating and rescaling variables $\mathrm{i}r^{\gamma\delta} \to r^{\gamma\delta}/2$ one has

$$\begin{aligned}
\langle \Delta^{-n/2} \rangle_\eta = \int \prod_{\gamma\delta} \frac{\mathrm{d}s^{\gamma\delta} \mathrm{d}r^{\gamma\delta}}{4\pi N} \exp\left[\left[-\frac{N}{2} \left\{ \tilde{\alpha} \mathrm{Tr}_\gamma \ln(\mathbf{1} + T_0 \omega_1 \mathbf{s}) - \sum_{\gamma\delta} s^{\gamma\delta} r^{\gamma\delta} \right.\right.\right. \\
\left.\left.\left. + \left\langle \mathrm{Tr}_\gamma \ln(\mathbf{r} q_i(\lambda) - \mathrm{e}^{\lambda \Delta t} \mathbf{1}) \right\rangle \right\} \right] \\
\times \exp\left\{ -\frac{1}{2} \mathrm{Tr}_\gamma \ln\left[\left(1 + \omega_0 \left\langle \frac{q_i(\lambda)}{\mathbf{r} q_i(\lambda) - \mathrm{e}^{\lambda \Delta t} \mathbf{1}} \right\rangle \right) \right.\right. \\
\times \left(1 + \omega_1 \left\langle \left(\frac{\eta^1}{a} - 1 \right)^2 \frac{q_i(\lambda)}{\mathbf{r} q_i(\lambda) - \mathrm{e}^{\lambda \Delta t} \mathbf{1}} \right\rangle \right) \\
\left.\left. - \omega_0 \omega_1 \left\langle \left(\frac{\eta^1}{a} - 1 \right) \frac{q_i(\lambda)}{\mathbf{r} q_i(\lambda) - \mathrm{e}^{\lambda \Delta t} \mathbf{1}} \right\rangle^2 \right] \right\}
\end{aligned} \tag{86}$$

where $\tilde{\alpha} \equiv (p-1)/N$ ($\tilde{\alpha}$ is used rather than α in order to keep in mind that N is the number of active cells (classes), *not* the total number) and $\langle \cdot \rangle$ denotes averages over the N cells. Note that Δ is a complex quantity, defined in general by analytic continuation, and that a certain number of rotations of variables in the complex plane may be required to evaluate the integrals for a given λ. Considering a replica-symmetric solution

$$\begin{aligned}
r^{\gamma\gamma} &= r_0 \qquad s^{\gamma\gamma} = s_0 \\
r^{\gamma\delta} &= r_1 \qquad s^{\gamma\delta} = s_1 \qquad (\gamma \neq \delta)
\end{aligned}$$

and taking the $n \to 0$ limit leads to

$$f = -\tilde{\alpha}\left[\ln(1 + T_0\omega_1(s_0 - s_1)) + \frac{T_0\omega_1 s_1}{1 + T_0\omega_1(s_0 - s_1)}\right] + r_0 s_0 - r_1 s_1 - \left\langle \ln((r_0 - r_1)q_i(\lambda) - e^{\lambda\Delta t}) + \frac{r_1 q_i(\lambda)}{(r_0 - r_1)q_i(\lambda) - e^{\lambda\Delta t}} \right\rangle \tag{87}$$

plus corrections of order $O(1/N)$. The saddle point equations for r_1, s_1 imply $r_1 = s_1 = 0$ for all $\tilde{\alpha}$ below a critical value $\tilde{\alpha}^*$, which in fact reveals that, in that $\tilde{\alpha}$-range, the replica treatment was redundant, as different replicas turn out to be uncorrelated. Eliminating s_0 as well leads to equation (72), where $\tilde{\alpha}^*$ is given by an analysis of the Jacobian as

$$\tilde{\alpha}^* = \left\langle \left(\frac{r_0(\tilde{\alpha}^*)q_i(\lambda)}{r_0(\tilde{\alpha}^*)q_i(\lambda) - e^{\lambda\Delta t}} \right)^2 \right\rangle. \tag{88}$$

The meaning of $\tilde{\alpha}^*$ is still unclear.

References

[1] Marr D 1971 Simple memory: a theory for archicortex *Phil. Trans. R. Soc.* B **262** 24–81

[2] Kohonen T 1977 *Associative Memory* (Berlin: Springer)

[3] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities *Proc. Natl Acad. Sci. USA* **79** 2554–8

[4] Amit D J 1989 *Modeling Brain Function* (Cambridge: Cambridge University Press)

[5] Sompolinsky H, Golomb D and Kleinfeld D 1990 Global processing of visual stimuli in a neural network of coupled oscillators *Proc. Natl Acad. Sci. USA* **87** 7200-4

[6] Sompolinsky H, Golomb D and Kleinfeld D 1991 Cooperative dynamics in visual processing *Phys. Rev.* A **43** 6990–7011

[7] von der Malsburg C and Schneider W 1986 A neural cocktail-party processor *Biol. Cybern.* **54** 29–40

[8] Gray C M, Konig P, Engel A K and Singer W 1989 Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties *Nature* **338** 334–7

[9] Eckhorn R, Bauer R, Jordan W, Brosch M, Kruse W, Munk M and Reitbaeck H J 1988 Coherent oscillations—a mechanism of feature linking in the visual cortex; multiple electrode and correlation analysis in the cat *Biol. Cybern.* **60** 121–30

[10] Tovee M J, Rolls E T, Treves A and Bellis R P 1993 Information encoding and the responses of single neurons in the primate temporal visual cortex *J. Neurophysiol.* in press

[11] Tuckwell H J 1988 *Introduction to Theoretical Neurobiology* (Cambridge: Cambridge University Press)

[12] McGregor R J 1987 *Neural and Brain Modelling* (New York: Academic)

[13] Mahowald M and Douglas R 1991 A silicon neuron *Nature* **354** 515–8

[14] Horner H, Bormann D, Frick M, Kinzelbach H and Schmidt A 1989 Transients and basins of attractions in neural network models *Z. Phys.* B **76** 381–98

[15] Mirollo R E and Strogatz S H 1990 Synchronization of pulse-coupled biological oscillators *SIAM J. Appl. Math.* **50** 1645–62

[16] Kuramoto Y 1991 Collective synchronization of pulse coupled oscillators and excitable units *Physica* **50D** 15–30

[17] Bush P C and Douglas R J 1991 Synchronization of bursting action potential discharge in a model of neocortical neurons *Neural Comput.* **3** 19–30

[18] Abbott L F and Van Vreeswijk C 1992 Asynchronous states in networks of pulse-coupled oscillators *Preprint* Brandeis University

[19] Treves A, Rolls E T and Tovee M 1992 On the time required for recurrent processing in the brain *Preprint* Oxford University

[20] Brown D A, Gähwiler B H, Griffith W H and Halliwell J V 1990 Membrane currents in hippocampal neurons *Prog. Brain Res.* **83** 141–60

[21] Jack J J B, Noble D and Tsien R W 1985 *Electric Current Flow in Excitable Cells* (Oxford: Clarendon)

[22] McCormick D A, Connors B W, Lighthall J W and Prince D A 1985 Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex *J. Neurophysiol.* **54** 782–806

[23] Connors B W, Malenka R C and Silva L R 1985 Two inhibitory postsynaptic potentials, and $GABA_A$ and $GABA_B$ receptor-mediated responses in neocortex of rat and cat *J. Physiol.* **406** 443–68
[24] Mason A and Larkman A 1990 Correlations between morphology and electrophysiology of pyramidal neurons in slices of rat visual cortex. II Electrophysiology *J. Neurosci.* **10** 1415
[25] Frolov A A and Medvedev A V 1986 Substantiation of the 'point approximation' for describing the total electrical activity of the brain with the use of a simulation model *Biophysics* **31** 332–6
[26] Amit D J and Tsodyks M V 1991 Quantitative study of attractor neural network retrieving at low spike rates: I. Substrate—spikes, rates and neuronal gain *Network* **2** 259–73
[27] Wilson H R and Cowan J D 1974 Excitatory and inhibitory interactions in localized populations of model neurons *Biophys. J.* **12** 1–24
[28] Abbott L F 1991 Realistic synaptic inputs for model neural networks *Network* **2** 245–58
[29] Van Vreeswijk C and Abbott L F 1991 Self-sustained firing in populations of integrate and fire neurons *Preprint* Brandeis University
[30] Treves A and Rolls E T 1991 What determines the capacity of autoassociative memories in the brain? *Network* **2** 371–97
[31] Treves A 1990 Graded-response neurons and information encodings in autoassociative memories *Phys. Rev.* A **42** 2418–30
[32] Hansel D and Sompolinsky H 1992 Synchronization and computation in a chaotic neural network *Phys. Rev. Lett.* **68** 718–21