Threshold setting and mean firing rates in networks of integrate-and-fire neurons

Hamish Meffin and William G. Gibson[1]
School of Mathematics and Statistics
The University of Sydney
New South Wales 2006
Australia

July, 1998

[1]To whom correspondence should be addressed

Abstract

The time-course and stability of the firing rates in a network of interconnected spiking integrate-and-fire neurons is analyzed for the simplified case where the conduction of the neuron's membrane is taken to be zero. The treatment is valid for networks of arbitrary size, no symmetry requirements are placed on the inter-neuronal connections and constant external inputs are allowed. An equation is given that explicitly relates the mean firing rates to the neuronal threshold, the synaptic response functions and the external inputs. For purely excitatory networks the Perron-Frobenius theorem for non-negative matrices allows the existence and stability of the steady states to be related to the eigenvalues and eigenvectors of the matrix of synaptic connections. The addition of inhibition allows more complex behaviour which is studied using a mixture of theory and computer simulations. In contrast to the excitatory case, the final state now depends on the initial state, thus allowing the possibility of memory storage. In addition, such a network can exhibit multistable and oscillatory behaviours.

1 Introduction

An ongoing challenge in neural network research is to incorporate more biological realism without complicating the model to the point where analytic results cannot be obtained. The simplest networks use on-off (McCulloch-Pitts) neurons as their basic elements, and many important and interesting results have been obtained from the analysis of such networks, especially in the study of associative memory storage and recall. More sophisticated neuronal models incorporate ion channels (Hodgkin-Huxley type) as well as dendritic structure, but then very little progress can be made analytically. (Single neuron models have been reviewed by Segev [16].) Between these two extremes is the so-called "integrate-and-fire" model neuron (sometimes called the Lapicque model), based on an equivalent-circuit representation of the electrical properties of a point neuron [22] [18].

The integrate-and-fire model divides the neuron's behaviour into two distinct regimes: a prolonged period of linear integration (summation of inputs) and then a sudden firing. Whereas in the Hodgkin-Huxley model both these processes are described by a single set of differential equations, in the integrate-and-fire model the differential equation describes only the sub-threshold integration phase; the neuron's membrane potential starts from zero and increases or decreases according to the input to the neuron; when it reaches a certain set threshold θ the neuron fires an instantaneous "spike", the potential is reset to zero and the summation process recommences. The spike is then transmitted to other neurons which are post-synaptically connected to the firing neuron, where it causes a change in their membrane potentials. Thus the integrate-and-fire neuron is realistic in that it communicates via spikes, which approximate the action potentials of real neurons.

This paper uses the integrate-and-fire model neuron. However, in order to obtain a number of analytic results the model has been further simplified by setting the conductance of the cell membrane to zero, thus giving a "leakless" or "leak-free" neuron. The model neuron now sums all its afferent currents and the membrane potential increases (monotonically in the case of purely excitatory inputs) until the threshold is reached, when the neuron fires and the potential is reset to its resting state. In the absence of input the membrane potential would remain constant, whereas in the normal "leaky" case it would decay exponentially with a time constant which is the product of the resistance and capacitance of the membrane. If this time constant is long compared to the mean inter-spike interval for the firing of the neuron, then neglecting the leak is a reasonable approximation.

Mathematical analysis of networks of integrate-and-fire neurons is difficult. This is particularly true if a description of precise firing times is sought. In particular, studies of the synchronisation and phase-locking properties of these networks have typically succeeded by focusing on limiting cases. Thus papers have dealt with networks that are homogeneous [14] [9] [23], that involve just two neurons [3] [7] [9] [23], use neurons that communicate instantaneously [14] [3] [7] [10], or ignore the leaky nature of neurons [3]. This is not to say that these works are not valuable as a step towards understanding the dynamics more generally; they certainly reveal effects or properties that may be tested in more realistic settings, by simulation if necessary.

Such a simplified approach, but with still some biological realism, has been taken by Amit and Tsodyks [1] [2] in their work on associative memory. They start by arguing that under conditions prevalent in the neo-cortex a conversion from spikes to rates is possible. Then, using a network with many neurons and high connectivity, they argue that the mean firing rates of their integrate-and-fire model are the same as those of a graded response model with a certain gain function. They caution that the validity of the conversion may be corrupted if there is synchronisation of the spike times. Gerstner and Van Hemmen [8] extended this result by proving that, under conditions where there are a finite number of subnetworks each containing many identical neurons, their class of spiking models have stable states in which the mean firing rates are the same as those of an equivalent graded response model. They also caution that other states can exist, for example synchronous ones and give instances where this can happen, Other authors have extended this approach to associative memory, for example Treves [20] and Burkitt [5]. The fundamental idea in all of these papers is to work in regimes where the mean firing rate of a neuron is given by mean firing rates of all presynaptic neurons; the detailed timing of spikes is unimportant. In the present work, it is shown that for the leakless integrate-and-fire model a conversion from spikes to rates in the above sense is always possible, regardless of the size, connectivity or symmetry of the network.

However, this is not the main point of the paper. Rather, a central concern is the investigation of the effect of choice of threshold θ on the maintenance of network stability and the organisation of attractors in the network. This leads to an exploration of the way in which the behaviour of the network is related to the set of external inputs, the threshold and the initial state, with an emphasis on statements that can be made for a broad class of networks. This generic behaviour is still improperly understood.

The plan of the paper is as follows. In §2 the integrate-and-fire neuron is described, and the way in which it operates in a network is specified. In the first instance, we consider purely excitatory networks (§3); that is, networks in which all the synaptic weights are non-negative and external inputs are also excitatory. In this case, we can derive a simple equation approximating the time course of the the firing rates of the neurons in the network. With the help of some known properties of non-negative matrices, we are able to give a complete account of the behaviour of the network. These results apply to networks of any size, but for simplicity illustrations are usually given for two-neuron systems.

The addition of inhibition (§4), whether by way of synaptic connections or from external sources, considerably complicates the picture and a complete account of all possible behaviours cannot be given. We focus on a class of networks whose synapses have identical time courses (but not magnitudes) for their response and have synaptic weight matrices with dominant eigenvalues. Particular attention is paid to elucidating the way in which external input influences the network and to the type of bifurcation structure it introduces in association with the threshold.

2 The Network

2.1 Integrate-and-fire neurons

We consider a network of N interconnected neurons in which the ith neuron is described by the equation

$$\frac{dV_i}{dt} = -\frac{V_i}{\tau} + I_i, \quad V_i < \theta, \tag{1}$$

where V_i represents the internal state of the neuron (and may be identified with the membrane potential), τ is the membrane time constant (and is thus a measure of the time taken for the potential to decay due to the "leaky" nature of the cell membrane) and I_i is the input into the neuron (and may be thought of as a suitably scaled current). This equation describes the sub-threshold behaviour of the neuron: provided V_i remains less that some pre-assigned threshold θ its time course is found by solving equation (1). Whenever V_i reaches θ the neuron "fires", sending out an instantaneous potential spike and at the same instant V_i is reset to zero. The summation process, governed by equation (1), then begins again.

The current $I_i \equiv I_i(t)$ into the neuron at time t can be split as

$$I_i(t) = I_i^{int}(t) + I_i^{ext}(t), \tag{2}$$

where $I_i^{int}(t)$ is the current received from the firing of the other neurons in the network and $I_i^{ext}(t)$ is the current received from sources external to the network.

Suppose neuron j fires at time $t = 0$ and sends a voltage spike to neuron i. The various processes occurring at the synaptic connection between j and i convert this into a current into neuron i, the conversion being governed by a synaptic response function $p_{ij}(t)$; that is, $p_{ij}(t)$ is the current into neuron i at time t due to the firing of neuron j at time $t = 0$. The current $I_i^{int}(t)$ is the input to neuron i summed over all (spatial) inputs and over all firing history; thus if t_j^k is the time at which neuron j fired its kth spike then

$$I_i^{int}(t) = \sum_{j=1}^{N} \sum_{k=1,2,\ldots} p_{ij}(t - t_j^k) + p_i^0(t), \tag{3}$$

where $p_i^0(t)$ is due to all synaptic inputs before $t = 0$. (Alternatively, $p_i^0(t)$ may be thought of as resulting from a one-off pulse that neuron i receives at $t = 0$.) In the following, we shall generally suppress $p_i^0(t)$ since it can be absorbed into $I_i^{ext}(t)$.

A simple form for the response function is a decaying exponential:

$$p_{ij}(t) = \begin{cases} 0, & t < 0 \\ \alpha J_{ij} e^{-\alpha t}, & t > 0, \end{cases} \tag{4}$$

where $\alpha > 0$ and J_{ij} are constants; this has no delay time and starts instantly at $t = 0$ from the value $p_{ij}(0) = \alpha J_{ij}$. A more realistic form would incorporate some time delay τ_0 such that $p_{ij}(0) = 0$ for $t < \tau_0$ and then would have a smooth time course, better represented by a sum of exponentials or an α-function (see, for example, [24]). However, for convenience we will use the simple form of equation (4) in the illustrative examples given below; the formal proofs hold for more general response functions, including the possibility of completely different response functions for different synapses. Unless stated otherwise, the only requirements are that $p_{ij}(t) = 0$ for $t < 0$, $p_{ij}(t) \geq 0$ for all $t \geq 0$ (excitatory) or $p_{ij}(t) \leq 0$ for all $t \geq 0$ (inhibitory) and that the function $h(t) = \int_t^\infty |p_{ij}(t')|dt'$ is integrable on $[0, \infty)$.

The area under the curve $p_{ij}(t)$ is a measure of the total charge transferred to neuron i due to the firing of neuron j and is the synaptic efficacy J_{ij}:

$$J_{ij} = \int_0^\infty p_{ij}(t)dt. \tag{5}$$

This will be positive for an excitatory synapse, negative for an inhibitory one, and zero if there is no connection from j to i.

In this paper, the model is further simplified by taking the neurons to be non-leaky. Thus $\tau = \infty$ and equation (1) reduces to

$$\frac{dV_i}{dt} = I_i, \quad V_i < \theta. \tag{6}$$

According to Knight [12], if $I_i >> \theta/\tau$ then the firing rate of the leaky model is well approximated by that of the leakless model with the current reduced by an amount $\theta/2\tau$; the approximation is already accurate to about 4 percent when $I_i = 2\theta/\tau$. Networks of leakless neurons are still capable of surprisingly complex behaviour, as will be seen below.

3 Purely Excitatory Networks

In this section we study the behaviour of networks that contain only excitatory synaptic connections ($J_{ij} \geq 0$) and whose external inputs are also excitatory ($I_i^{ext} \geq 0$). We study this scenario, not because it is of immediate biological relevance, but because some of the fundamental aspects of the general dynamics are explained most clearly in this case.

The excitatory assumptions mean that the membrane potential $V_i(t)$ of each neuron is a monotonic increasing function of time, except that it is abruptly reset to zero each time the neuron fires. The internal input current $I_i^{int}(t)$ will fluctuate in a manner dependent on the firing pattern of the other neurons and the particular form used for the synaptic response function.

The behaviour of a typical neuron under purely excitatory inputs is illustrated in figure 1, where the response function of equation (4) has been used. The current (figure

1(a)) jumps and then decays exponentially with time-constant $1/\alpha$ (=1) between inputs. The potential (figure 1(b)) increases monotonically until the threshold θ (=10) is reached, at which point it is reset to zero. (Note that only the sub-threshold behaviour of V is shown in figure 1(b); the firing of the neuron may be thought of as an instantaneous spike which occurs each time V reaches θ.)

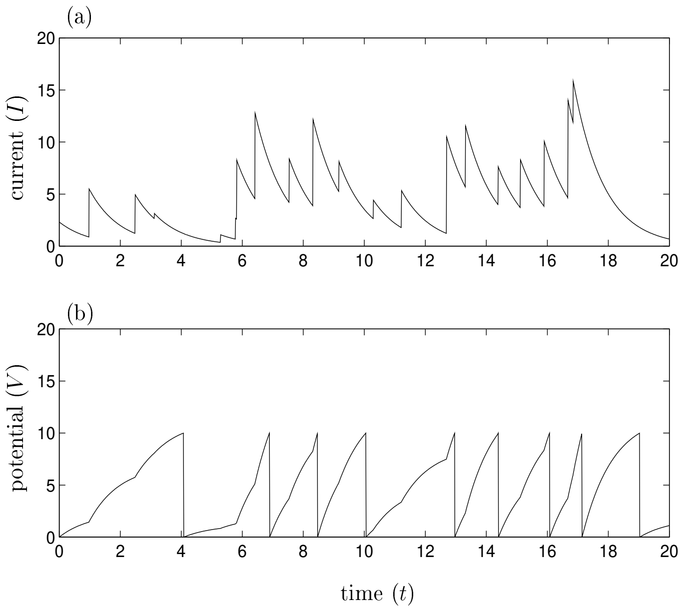

Figure 1: Typical behaviour of a leak-free integrate-and-fire neuron receiving a sequence of excitatory impulses of different strengths at random times. The calculation uses the single-exponential response function given by equation (4) with $\alpha = 1$ and a threshold $\theta = 10$. (a) The time-course of the current I, calculated using equation (3); at each received impulse I increases abruptly by an amount αJ_{ij} and then decays exponentially with time constant $1/\alpha$. (b) The resulting membrane potential V, calculated using equation (6); V increases monotonically until the threshold θ is reached, whereupon it is reset to zero. Each time V reaches θ the neuron is assumed to fire, sending out an instantaneous impulse representing an action potential. These impulses can be thought of as spikes occurring at the peaks of V, but the graph shows only the sub-threshold behaviour of the neuron.

The picture does not alter much if equation (4) is replaced by a more biologically realistic response function, for example a double exponential [24], particularly as the rise time is typically much shorter than the decay time. Furthermore, our fundamental result below, which describes the firing rates of the network, is still valid for any reasonable choice of $p_{ij}(t)$. For example, shifting $p_{ij}(t)$ along the t-axis allows arbitrary time delays.

3.1 Equation for the firing rates

We shall now derive an expression for the time dependence of the firing rates of the neurons. If $r_i(t)$ is the number of times that neuron i has fired by time t and $T_i(t)$ is the time of the latest firing before t, then it is a consequence of equation (6) that

$$r_i(t) = \frac{1}{\theta}\int_0^{T_i(t)} dt' I_i(t'). \tag{7}$$

Alternatively, we could define a smoothed version of the count as

$$\overline{r}_i(t) = \frac{1}{\theta}\int_0^{t} dt' I_i(t') = r_i(t) + \frac{V_i(t)}{\theta}, \tag{8}$$

where the right-most term is limited to the range from 0 to 1 since the neuron receives only excitatory input. From this, we define the "instantaneous firing rate"

$$\nu_i(t) = \frac{d\,\overline{r}_i(t)}{dt} = \frac{I_i(t)}{\theta}. \tag{9}$$

Working with this firing rate, we find by substituting equations (2) and (3) into (9) that

$$\begin{aligned}\theta\nu_i(t) &= I_i^{ext}(t) + \sum_{j=1}^{N}\sum_{k=1,2,\ldots} p_{ij}(t-t_j^k) &(10)\\ &= I_i^{ext}(t) + \sum_{j=1}^{N}\int_0^t dr_i(t')p_{ij}(t-t'). &(11)\end{aligned}$$

We now approximate the integral in the last term by replacing the measure $dr_i(t)$ with its smoothed version $d\overline{r}_i(t)$; over any interspike interval they both contribute the same measure to their respective integrals and our approximation will be accurate so long as the interspike interval remains small compared to the decay time of the response function $p_{ij}(t)$ (plus time delays) and the external current $I_i^{ext}(t)$. The approximation also requires that any small errors are not compounded by exponential divergence. (For an example of where this can occur, see Figure 2b below.) Using $d\overline{r}_i(t) = \nu_i(t)dt$ we have a single closed equation for the firing rate:

$$\theta\nu_i(t) = I_i^{ext}(t) + \sum_{j=1}^{N}\int_0^t dt' p_{ij}(t-t')\nu_j(t'). \tag{12}$$

This equation may be solved using Laplace transforms. Denoting the transform of a function $f(t)$ by $\hat{f}(s)$ and using the convolution property, equation (12) becomes

$$(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))\hat{\boldsymbol{\nu}}(s) = \hat{\boldsymbol{I}}^{ext}(s) \tag{13}$$

where $\boldsymbol{\theta}$ is the diagonal matrix with entries θ, $\hat{\boldsymbol{\nu}} = [\hat{\nu}_1, \ldots, \hat{\nu}_N]'$, $\hat{\boldsymbol{I}}^{ext} = [\hat{I}_1^{ext}, \ldots, \hat{I}_N^{ext}]'$ and $\hat{\boldsymbol{P}} = [\hat{p}_{ij}]$. Solving this for $\hat{\boldsymbol{\nu}}(s)$ and applying the inverse transform gives the full time-dependent solution for firing rates as

$$\boldsymbol{\nu}(t) = \frac{1}{2\pi i}\int_c ds\, \mathrm{e}^{st}(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))^{-1}\hat{\boldsymbol{I}}^{ext}(s), \tag{14}$$

where the contour c is a vertical line in the complex plane to the right of all singularities of the integrand.

3.2 Steady state solution under constant excitatory input

To proceed further we must make some assumptions about the external input, $\boldsymbol{I}^{ext}(t)$, the simplest of which is to assume that it is constant in time and excitatory. We may then study steady state solutions of equation (14).

We wish to identify the singularities of the integrand in equation (14). For simplicity we assume that the response functions are such that their Laplace transforms are meromorphic: that is, there are finitely many singularities all of which are poles. This is not a restrictive assumption since response functions commonly used in the literature have this property. It then follows that $\hat{\boldsymbol{\nu}}(s)$ is meromorphic, so all the singularities of the integrand are poles (see Appendix B) and the residue theorem may be applied to evaluate the contribution to the integral from each of them. One pole occurs at $s = 0$, since the Laplace transform of the constant function $\boldsymbol{I}^{ext}(t) = \boldsymbol{I}^{ext}$ is $\hat{\boldsymbol{I}}^{ext}(s) = \boldsymbol{I}^{ext}/s$. All other poles arise from the factor $(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))^{-1}$; they are typically the zeros of the function $D(s) = \det(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))$ (see Appendix B for a discussion of this). Assuming that the threshold is such that all the poles ξ are simple, it follows that each of them contributes a term with time dependence $\mathrm{e}^{\xi t}$, so that firing rates have the form

$$\boldsymbol{\nu}(t) = (\boldsymbol{\theta} - \boldsymbol{J})^{-1}\boldsymbol{I}^{ext} + \sum_{\xi : D(\xi)=0} \boldsymbol{c}(\xi)\mathrm{e}^{\xi t}, \tag{15}$$

where the coefficients $\boldsymbol{c}(\xi) = \lim_{s\to\xi} s^{-1}(s-\xi)(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))^{-1}\hat{\boldsymbol{I}}^{ext}$. We have used the fact that the term evaluated at the pole $s = 0$ involves $\hat{\boldsymbol{P}}(0) = \boldsymbol{J}$, where $\boldsymbol{J}$ is the matrix of synaptic weights. From this we see that there is a unique steady state $(\boldsymbol{\theta} - \boldsymbol{J})^{-1}\boldsymbol{I}^{ext}$. It can be verified that the steady state must have these firing rates, without resorting to the approximation we used in going from equation (11) to equation(12) (see Appendix A). However, in order to study the time dependent aspects of the solution, which appear in the remaining terms of equation (15), it is convenient to use this approximation. These remaining terms show that the steady state is stable if and only if every zero of $D(s)$ has negative real part.

Under the purely excitatory assumptions of this section we can take this stability condition further, by relating the stability directly to the so-called Perron-Frobenius eigenvalue (P-F eigenvalue), Θ, of the synaptic matrix $\boldsymbol{J}$. This matrix is non-negative in the sense that all of its elements are non-negative and hence, according to the Perron-Frobenius theorem, there exists a real, positive, eigenvalue Θ of $\boldsymbol{J}$ which is larger than the magnitude of any of the other eigenvalues [17]. The theorem requires the additional assumption of the irreducibility of $\boldsymbol{J}$, which may be interpreted here to mean that there are no sub-networks that do not receive input from the remainder of the network, which is a perfectly reasonable assumption. The P-F eigenvalue plays a critical role because it can be shown that the steady state is stable if and only if $\theta > \Theta$ (see Appendix B).

In addition to the stability of the steady state solution, we should check that the firing rates are non-negative. Under the assumption that the external inputs $\boldsymbol{I}^{ext}$ are excitatory and the assumptions of the Perron-Frobenius theorem, an extension to this theorem [17] tells us that the steady state $\boldsymbol{\nu} = (\boldsymbol{\theta} - \boldsymbol{J})^{-1}\boldsymbol{I}^{ext}$ is non-negative if and only if $\theta > \Theta$: the same condition as for stability.

These results allow us to completely characterise the dynamics in this purely excitatory case: there is a unique steady state $\boldsymbol{\nu} = (\boldsymbol{\theta} - \boldsymbol{J})^{-1}\boldsymbol{I}^{ext}$ which is stable and physical if and only if $\theta > \Theta$; below this threshold value the firing rates become unbounded in exponential excitatory run-away. These behaviours are illustrated in figure 2, which depicts the results of simulations conducted on a two-neuron network with synaptic matrix

$$\boldsymbol{J} = \begin{pmatrix} 4 & 1 \\ 2 & 5 \end{pmatrix}, \tag{16}$$

and external input $\boldsymbol{I}^{ext} = [5, 5]$. The P-F eigenvalue is $\Theta = 6$. The figure shows the simulated firing rates $\boldsymbol{\nu}(t)$, as defined by equation (9) (jagged line); super-imposed on this is the theoretically predicted firing rates calculated from equation (14) (smooth line). Figure 2(a) shows the relaxation to the steady state $\boldsymbol{\nu} = [3.75, 6.25]$ when $\theta = 7.0 > \Theta$; the agreement between theory and simulation is evident. In figure 2(b) an unbounded state is shown for which $\theta = 5.7 < \Theta$; the discrepancy between theory and simulation is due to the compounding of small errors due to exponential divergence.

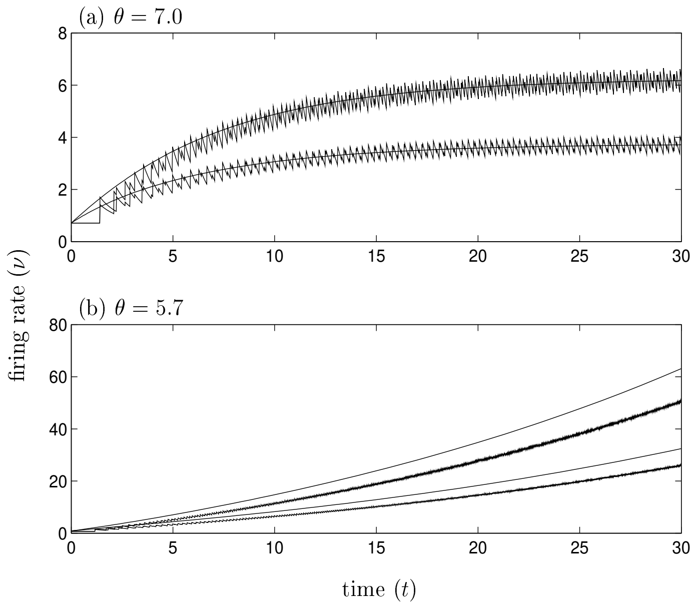

Figure 2: Behaviour of a two-neuron network with external excitatory input. The synaptic matrix is given by equation (16), for which the P-F eigenvalue is $\Theta = 6$ and the second eigenvalue is 3; the external current is $\boldsymbol{I}^{ext} = [5.0, 5.0]$. The jagged lines line show the firing rates $\boldsymbol{\nu}(t)$ obtained from simulation, as defined by equation (9); super-imposed over these are smooth lines showing the theoretically predicted firing rates calculated from equation (14). (a) Relaxation to the steady state, $\boldsymbol{\nu} = [3.75, 6.25]$, when $\theta = 7.0 > \Theta$; the agreement between theory and simulation is evident. (b) For $\theta = 5.7 < \Theta$ the firing rates become unbounded; the discrepancy between theory and simulation is due to the compounding of small errors due to exponential divergence.

4 Networks with Inhibition

The introduction of inhibitory synapses and external inhibitory inputs significantly complicates the neat picture given above for the purely excitatory case. This is not unwelcome: the excitatory case is too simple to be of interest as a neural network, as its dynamical behaviour is limited to a single steady-state attractor. Inhibition allows the possibility of multistable and oscillatory dynamics in the network, but there are also similarities with the excitatory case.

4.1 Equation for the firing rates for the steady states

If a neuron receives a net inhibitory current for an indefinite period of time it will cease firing. This process partitions the neurons into those that are firing ("foreground") and those that not ("background"). The previous analysis leading to equation (15) will continue to apply to the foreground neurons provided that the membrane potential $V_i(t)$ is still of order θ in equation (8). This will occur, for example, in a steady state when the neuron is receiving a steady net excitatory current since then $V_i(t)$ is monotone increasing from 0 to θ.

We shall assume a constant external current $\boldsymbol{I}^{ext}$ and study steady states of the network. Provided the firing rates of the network remain close enough to the steady state to be non-negative we have, for each of the 2^N partitions of the N neurons into foreground and background, an equation analogous to equation (15) for the time course of the firing rates of foreground neurons:

$$\boldsymbol{\nu}_f(t) = (\boldsymbol{\theta} - \boldsymbol{J}_f)^{-1}\boldsymbol{I}_f^{ext} + \sum_{\xi:D_f(\xi)=0} \boldsymbol{c}_f(\xi)\mathrm{e}^{\xi t}. \tag{17}$$

That is, instead of one equation we now have 2^N. The subscript f means that the quantity pertains to a foreground neuron; for example $\boldsymbol{J}_f$ is the matrix of synaptic weights between foreground neurons, $D_f(s) = \det(\boldsymbol{\theta} - \hat{\boldsymbol{P}}_f(s))$ and $\boldsymbol{c}_f(\xi) = \lim_{s\to\xi} s^{-1}(s-\xi)(\boldsymbol{\theta} - \hat{\boldsymbol{P}}_f(s))^{-1}\hat{\boldsymbol{I}}_f^{ext}$. We must check that in the steady state the foreground firing rates,

$$\boldsymbol{\nu}_f = (\boldsymbol{\theta} - \boldsymbol{J}_f)^{-1}\boldsymbol{I}_f^{ext}, \tag{18}$$

are positive and also that afferent currents into background neurons,

$$\boldsymbol{I}_b = \boldsymbol{J}_b\boldsymbol{\nu}_f + \boldsymbol{I}_b^{ext}, \tag{19}$$

are negative ($\boldsymbol{J}_b$ is the matrix of synaptic weights from foreground to background neurons and $\boldsymbol{I}_b^{ext}$ is the external current into the background neurons). Equation (17) indicates that the steady state is stable if and only if every zero of $D_f(s)$ has negative real part.

4.2 Conditions for the stability of firing rates

To proceed further requires additional assumptions, because we do not have an equivalent to the Perron-Frobenius theorem to inform us about the non-negativity of the solutions, nor can such general statements about the stability be made as in the excitatory case. In order to reveal the fundamental structure of the steady state attractors as cleanly as possible we make the following assumptions. First, for all partitions, the matrix $\boldsymbol{J}_f$ has a real, positive eigenvalue, Θ_f, greater than or equal to the magnitude of any of its other eigenvalues; this will be called the dominant eigenvalue of $\boldsymbol{J}_f$ and is analogous to the P-F eigenvalue of the excitatory case. Second, all synaptic response functions have the same time course; that is $p_{ij}(t) = J_{ij} g(t)$ where $\int_0^\infty dt\, g(t) = 1$. Under these conditions it can be shown that the steady state, $\boldsymbol{\nu}_f = (\boldsymbol{\theta} - \boldsymbol{J})^{-1} \boldsymbol{I}^{ext}$ and $\boldsymbol{\nu}_b = \boldsymbol{0}$, is stable if and only if $\theta > \Theta_f$ (see Appendix B). The second condition can be relaxed somewhat, so that the response functions need only have approximately the same time course, in the sense that it can be shifted to the left or right by some amount less than ϵ. (The choice of ϵ is discussed in Appendix B.)

However, larger time shifts can lead to unstable behaviour, as the following example shows. Consider a two-neuron network with the symmetric matrix of synaptic response functions

$$\boldsymbol{P}(t) = \begin{pmatrix} 4E_6(t) & 3E_{\tau_1}(t) \\ 3E_{\tau_1}(t) & -4E_{\tau_2}(t) \end{pmatrix}, \tag{20}$$

where the function $J_{ij} E_\tau(t)$ is the decaying exponential of equation (4) with $\alpha = 1$, but shifted to the right along the t-axis by τ so as to introduce a time delay of this amount. The quantities τ_1 and τ_2 represent time delays that can be varied independently. Figure 3 shows the results of numerical calculations, based on an analysis of the zeros of $D_f(s) = \det(\boldsymbol{\theta} - \hat{\boldsymbol{P}}_f(s))$ (see the theory surrounding equation (15)), to find the minimum number Φ_f, as a function of τ_1 and τ_2, such that the steady state is stable if and only if $\theta > \Phi_f$. It shows an area, including the case of equal delays, $\tau_1 = \tau_2 = 6$ (that is, identical time courses), for which $\Phi_f = \Theta_f = 5$. However, there are choices of delays for which $\Phi_f > \Theta_f$. Simulations show that for these delays a choice of threshold from $\theta \in (\Theta_f, \Phi_f)$ can in some cases lead to oscillatory dynamics. This result is at odds with the dynamics of networks of graded response neurons with symmetric matrices for which the existence of a Lyapanov function prohibits oscillatory behaviour [11].

The above example is typical of networks with inhibitory synapses; indeed a set of (possibly extreme) delays may always be constructed such that $\Phi_f \geq \Theta_f$. This leads to a condition sufficient for stability, valid for any $\boldsymbol{J}_f$. Consider the matrix, $\boldsymbol{J}_f^{abs}$, constructed by reversing the sign of any inhibitory synapses in $\boldsymbol{J}_f$, and let the P-F

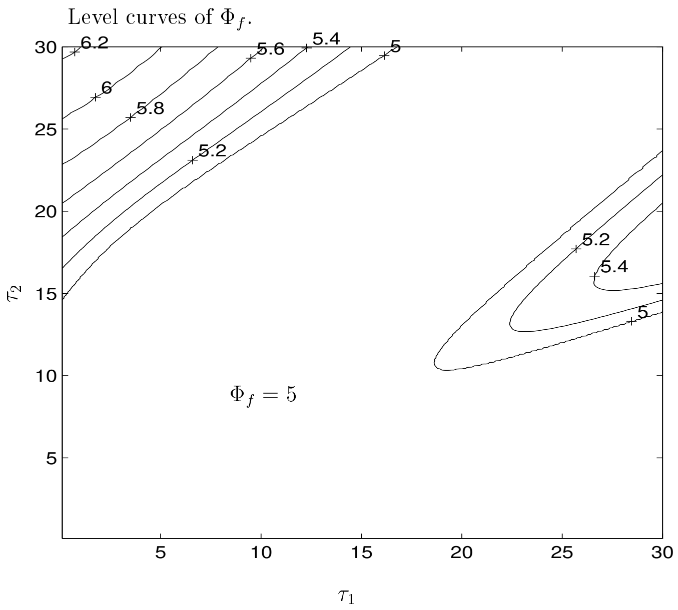

Figure 3: A contour plot for the stability parameter, Φ_f (see text), of a steady state as function of synaptic delays. The network is described by the matrix of synaptic response functions given in equation (20), and involves independent variable time delays τ_1 and τ_2. $\Phi_f = \Phi_f(\tau_1, \tau_2)$ is the minimum value that the threshold θ can have if the network state is to be stable. When all time delays are equal ($\tau_1 = \tau_2 = 6$) $\theta > \Phi_f = \Theta_f = 5$ where Φ_f is the P-F eigenvalue and the steady state is stable if and only if $\theta > \Phi_f$. There is a large region of time delays, centred about this point, for which this continues to hold - this corresponds to the flat region in the graph for which $\Theta_f = 5$. However, there are also regions of the graph for which $\Theta_f > 5$ and so a threshold θ larger than the P-F eigenvalue Φ_f is needed for stability.

eigenvalue of this matrix be Θ_f^{abs}; then the steady state is stable if $\theta > \Theta_f^{abs}$ (for proof, see Appendix B). A set of extreme delays may always be constructed so that this condition is also necessary. Further, the condition is necessary and independent of the choice of synaptic response function, precisely when $\boldsymbol{J}_f = \boldsymbol{D}\boldsymbol{J}_f^{abs}\boldsymbol{D}$ for some matrix $\boldsymbol{D}$ with ± 1's on the diagonal and 0's elsewhere (see Appendix B). In other words, the only synaptic matrices for which the choice of synaptic response functions can never affect the stability criterion are those belonging to the class of matrices $\{\boldsymbol{DND} : \boldsymbol{N} \geq \boldsymbol{0}\}$. For any other type of synaptic matrix, the response functions must be considered: if their time courses are significantly different then a calculation must be made to determine Φ_f by finding the largest real eigenvalue of $\hat{\boldsymbol{P}}_f(s)$ for $Re(s) > 0$.

4.3 Bifurcation structure of steady states

4.3.1 Dependence of bifurcation structure on external input

We now consider the problem of ensuring the positivity of the foreground firing rates and the negativity of the background afferent currents. A way of approaching the problem is to consider a particular partition of the neurons into foreground (f) and background (b) and then find the set of external currents leading to a valid solution with this partition. The process is then repeated for all other partitions (f, b). It is helpful to look at an example, so we shall consider the network with synaptic matrix given by

$$\boldsymbol{J} = \begin{pmatrix} 4 & -3 \\ -2 & 3 \end{pmatrix}, \tag{21}$$

with the choice of threshold $\theta = 5$. We proceed by looking at each of the four partitions in turn.

$\underline{f = \{1, 2\}, b = \phi}$

In this case $\boldsymbol{J}_f = \boldsymbol{J}$ and we look for steady states with both neurons firing. Since $\theta = 5$ is not an eigenvalue of $\boldsymbol{J}_f$ the inversion in equation (18) is valid and so

$$\boldsymbol{\nu}_f = (\boldsymbol{\theta} - \boldsymbol{J}_f)^{-1} \boldsymbol{I}_f^{ext}. \tag{22}$$

Using the requirement that $\boldsymbol{\nu}_f > \mathbf{0}$ results in the inequalities

$$\begin{aligned} -2I_1^{ext} + 3I_2^{ext} &> 0, \\ 2I_1^{ext} - I_2^{ext} &> 0, \end{aligned}$$

which define the set $\mathcal{I}_{12}$ of external currents leading to steady states with both neurons on. However, because the dominant eigenvalue of $\boldsymbol{J}_f$ is $\Theta_{12} = 6$, which is greater than the threshold $\theta = 5$, this state is not stable.

$\underline{f = \{1\}, b = \{2\}}$

For this partition $\boldsymbol{J}_f = 4$ and $\boldsymbol{J}_b = -2$, so the dominant eigenvalue is $\Theta_1 = 4$ which is less than the threshold, with the result that the steady state will be stable. When some of the neurons are in the background a second inequality is derived by substituting equation (18) into equation (19) and using the condition $\boldsymbol{I}_b \leq \mathbf{0}$ to give

$$\boldsymbol{J}_b(\boldsymbol{\theta} - \boldsymbol{J}_f)^{-1} \boldsymbol{I}_f^{ext} + \boldsymbol{I}_b^{ext} \leq \mathbf{0}. \tag{23}$$

This, together with the condition $\boldsymbol{\nu}_f > \mathbf{0}$, defines the set $\mathcal{I}_1$ of constant external currents leading to steady states with neuron 1 in the foreground and neuron 2 in the background:

$$
\begin{aligned}
I_1^{ext} &> 0, \\
-2I_1^{ext} + I_2^{ext} &\leq 0.
\end{aligned}
$$

The two other partitions into foreground and background are both stable and yield $\mathcal{I}_2$:

$$
\begin{aligned}
I_2^{ext} &> 0, \\
2I_1^{ext} - 3I_2^{ext} &\leq 0,
\end{aligned}
$$

$\mathcal{I}_\phi$:

$$
\begin{aligned}
I_1^{ext} &\leq 0, \\
I_2^{ext} &\leq 0.
\end{aligned}
$$

Using these four pairs of inequalities we can draw a diagram in the space of external currents showing the sets $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_\phi$ associated with respective stable steady states, which will collectively be called $\mathcal{I}$ (figure 4(a): $\mathcal{I}_{12}$ is the intersection of $\mathcal{I}_1$ and $\mathcal{I}_2$). This shows that for most choices of $\boldsymbol{I}^{ext}$ there is only one stable steady state, which may be the trivial one. The sets $\mathcal{I}_1$ and $\mathcal{I}_2$ overlap however ($\mathcal{I}_1 \cap \mathcal{I}_2 = \mathcal{I}_{12}$), representing a region of bistability where either neuron 1 is on and neuron 2 is off, or vice versa. The instability of states with both neurons on has resulted, not in unbounded behaviour, but rather in a region of bistability. An initial pulse of current determines to which of these two steady states the network converges (figure 4(e)).

4.3.2 Dependence of bifurcation structure on threshold

We may now proceed to examine the way in which threshold-setting changes the collection of sets $\mathcal{I}$ in our example, by considering this set through five descending threshold settings, $\theta = 7, 6, 5, 4$ and 3 (figure 5). This tells us how the existence, stability and multiplicity of steady states change with the threshold. It reveals that network behaviour passes from mono-stability, through to regions of bi-stability and finally to predominant regions of unbounded behaviour as the threshold passes from

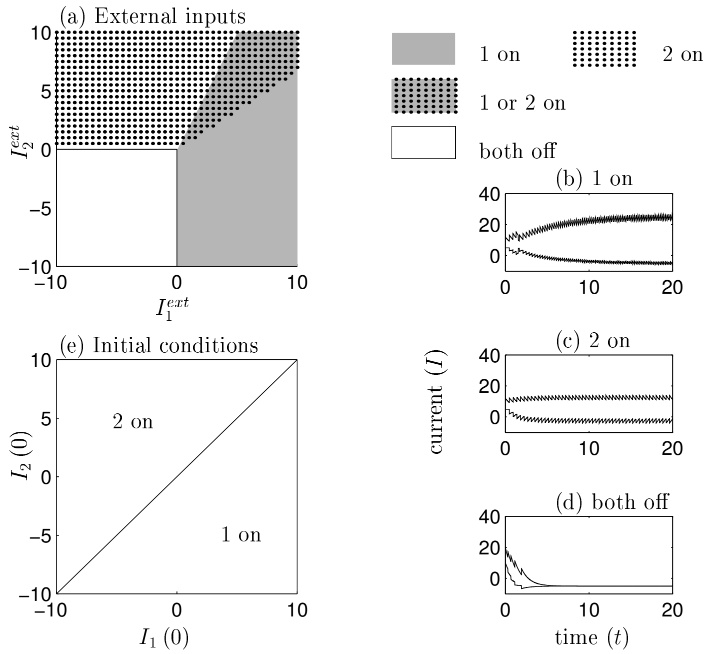

Figure 4: Behaviour of a two-neuron network with a mixture of excitatory and inhibitory synaptic connections, with external input. The synaptic matrix is given by equation (21) and the synaptic response function by equation (4) with $\alpha = 1$. The threshold is fixed at $\theta = 5$ and a range of external inputs is allowed. (a) shows the possible stable firing behaviours as a function of the external inputs, with regions corresponding to both neurons off (region $\mathcal{I}_\phi$, unshaded), only neuron 1 on (region $\mathcal{I}_1$, shaded) and only neuron 2 on (region $\mathcal{I}_2$, dotted). The intersection of the dotted and shaded regions is a region of bistability, where the final state of the network depends on the initial conditions $\boldsymbol{I}(0)$, as shown in (e). (b)-(d) show the time courses of the currents into neurons 1 and 2 for typical values of the external currents, corresponding to the three different behaviours.

high to low values. For $\theta = 7$ every $\boldsymbol{I}^{ext}$ is associated with a single stable steady state and all partitions are represented (figure 5(a)). A bifurcation occurs at $\theta = 6$, equal to the dominant eigenvalue of $\boldsymbol{J}$, when the set $\mathcal{I}_{12}$ loses stability and collapses to the half-line $I_2^{ext} = I_1^{ext} > 0$ (figure 5(b)). Bi-stability emerges as θ is decreased below this value, as shown for $\theta = 5$ (figure 5(c)), and continues until $\theta = 4 = \Theta_1$, at which point a second bifurcation occurs in which the set $\mathcal{I}_1$ loses stability and collapses to the line $I_1^{ext} = 0$ (figure 5(d)). This leaves a region in which none of the steady state equations (18) and (19) has a valid solution. It can be shown that whenever this is true, unbounded behaviour occurs: the steady state does not lose stability to more complicated but bounded dynamics (see appendix A). This region is again enlarged dramatically at $\theta = 3 = \Theta_2$, when the set $\mathcal{I}_2$ loses stability and collapses to the line $I_2^{ext} = 0$, leaving only one quadrant associated with a stable steady state, namely the trivial one (figure 5(e)).

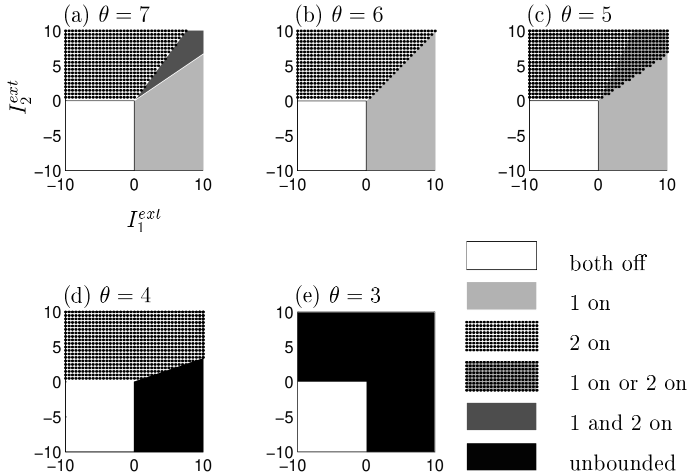

Figure 5: Behaviour of a two-neuron network with a mixture of excitatory and inhibitory synaptic connection, with external input. The network is the same as that for figure 4 and the behaviour as a function of the external inputs is shown for five descending threshold settings: (a) $\theta = 7.0$. The network exhibits monostable behaviour, with each $\boldsymbol{I}^{ext}$ giving rise to a single steady state with either both neurons on (region $\mathcal{I}_{12}$, dark shading), only neuron 1 on (region $\mathcal{I}_1$, light shading), only neuron 2 on (region $\mathcal{I}_2$, dotted) or both off (region $\mathcal{I}_\phi$, unshaded); (b) $\theta = 6 = \Theta_{12}$, the dominant eigenvalue of $\boldsymbol{J}$. The behaviour is still monostable, but the region $\mathcal{I}_{12}$ corresponding to both neurons firing has lost stability and collapsed to the half-line $I_1^{ext} = I_2^{ext} > 0$; (c) $\theta = 5$. The regions $\mathcal{I}_1$ and $\mathcal{I}_2$ overlap, giving a region of bistability (compare figure 4(a)); (d) $\theta = 4 = \Theta_1$, the dominant eigenvalue of $\boldsymbol{J}_1$. The region $\mathcal{I}_1$ has lost stability and collapsed to the line $I_1^{ext} = 0$, leaving a region corresponding to unbounded behaviour (shaded black); (e) $\theta = 3 = \Theta_2$, the dominant eigenvalue of $\boldsymbol{J}_2$. The region $\mathcal{I}_2$ has lost stability and collapsed to the line $I_2^{ext} = 0$,. The only stable state left is the trivial one, but most inputs lead to unbounded firing rates.

In this example, the bifurcations were of two types. While both occurred at a dominant eigenvalue, the first is distinguished by a collapse to a half line ($\theta = 6 = \Theta_{12}$) of the set $\mathcal{I}_{12}$, followed by the emergence of a region of bi-stability as the threshold drops below this. The second is characterised by a collapse to a full line ($\theta = 4 = \Theta_1$ of the set $\mathcal{I}_1$ and also $\theta = 3 = \Theta_2$ of the set $\mathcal{I}_2$), followed by the emergence (or enlargement) of a region associated with unbounded behaviour. We conjecture that these observations may be generalised under the assumptions of this subsection to the following: Let $\theta = \Theta^L$ be the greatest dominant eigenvalue amongst partitions (f, b) such that there is a valid solution to the purely recurrent case

$$\boldsymbol{J}_f \boldsymbol{\nu}_f = \Theta^L \boldsymbol{\nu}_f \tag{24}$$

$$\boldsymbol{\nu}_f > \mathbf{0} \quad , \quad \boldsymbol{J}_b \boldsymbol{\nu}_f < \mathbf{0}, \tag{25}$$

and let Θ^M be the greatest dominant eigenvalue of all partitions. Then there is a threshold range (Θ^L, Θ^M) such that there are regions of external current giving rise to multi-stable steady states for $\theta \in (\Theta^L, \Theta^M)$. Further, for θ in this range, bifurcations occur at dominant eigenvalues and are characterised by an increase in the degree of multi-stability as θ drops below this value. For $\theta > \Theta^M$ there is a mono-stable steady state for every choice of $\boldsymbol{I}^{ext}$, and for $\theta < \Theta^L$ there will be unbounded behaviour for some regions of external current.

We have not been able to provide a complete proof of the above statement; it is based on partial proof, plus numerical calculations using 10,000 cases of randomly generated 3×3 and 4×4 matrices. In addition, our network simulations also support this statement. The conjecture is important because it identifies a threshold range in which desirable multistable behaviour occurs without the possibility of undesirable unbounded behaviour. It is then interesting to ask what the degree of multi-stability is, within this threshold range. This question is related to the capacity of an associative memory, which has been the subject of studies with numerous models using various realisations of the matrix $\boldsymbol{J}$ (for example [2], [5] and [20]).

In the most general case, where arbitrary synaptic response functions and synaptic matrices are allowed, we believe a similar bifurcation structure exists. It is complicated by stability conditions for steady states that may not be associated with dominant eigenvalues and the occurrence of oscillatory dynamics. We do not pursue these issues here.

5 Discussion and Conclusion

We have analysed the behaviour of a network of interconnected leak-free integrate-and-fire neurons, in the presence of external input $\boldsymbol{I}^{ext}$. The central result describes the time dependence of the firing rates leading to a steady state. The expression (17) is applicable to networks of arbitrary size and no restriction is placed on the form of the synaptic matrix (in contrast to most treatments which require $\boldsymbol{J}$ to be symmetric). On this basis, we have discussed the local and global stability of the network for various cases leading to a description of the attractor structure and bifurcations that occur under variation of the threshold parameter θ.

We first considered the case of a network with purely excitatory connections and input and found the behaviour to be very simple: a unique attractor, corresponding to a state with constant mean firing rates, exists and is stable if and only if the threshold, θ, is greater than the Perron-Frobenius eigenvalue, Θ, of the synaptic matrix $\boldsymbol{J}$; otherwise the behaviour is unbounded.

The addition of inhibition allows more complex behaviour. We argued that a similar equation to that of the excitatory case applies for the time course of the firing rates around each of the many steady states. Multistable behaviour is generally a possibility in this case: for a given θ and $\boldsymbol{I}^{ext}$, the network allows more than one non-trivial steady state and retrieval depends on the initial pulse $\boldsymbol{p}^0(t)$. By specialising to the case where every synaptic matrix $\boldsymbol{J}_f$ has a dominant eigenvalue and the time course of all synaptic response functions is the same, it was found that the steady state for a particular partition of the neurons into foreground and background is stable if and only if the threshold θ is greater than the dominant eigenvalue Θ_f for that partition. Bifurcations at these dominant eigenvalues were classified as one of two types: those for which a transition to multistability occurs and those for which unbounded behaviour ensues. These considerations led to the conjecture, supported in part by theory and in part by simulation, that there are two well defined but not necessarily distinct dominant eigenvalues, Θ^M and Θ^L, delimiting a range of threshold settings (Θ^L, Θ^M) in which there are choices of external currents that allow multi-stability but no currents that give rise to unbounded behaviour. Within this threshold range, this suggests a form of neural computation in which there are two types of input; the first type being steady input of the form $\boldsymbol{I}^{ext}$ and the second being pulse-like input of the form $\boldsymbol{p}^0(t)$. Then, for fixed $\boldsymbol{I}^{ext}$, the network behaves like a dynamical system with a landscape of fixed points and associated basins of attraction, with the initial conditions determining the final state of the system. Changing $\boldsymbol{I}^{ext}$ has the effect of switching to a different landscape and thus a different pattern of retrieval, while still utilising the same neurons with the same connection strengths. Put another way, the network may be given a pulse-like signal that corresponds to some stored memory, but unless this is reinforced by having the correct $\boldsymbol{I}^{ext}$ this memory cannot be retrieved.

A comparison between the excitatory case and the case where inhibitory synapses are included indicates that the primary role of inhibitory neurons, in this model, is to give additional structure to the solution space, thus allowing multistability. These neurons did not typically act as 'activity controllers' in the network, a function that was most often achieved by threshold setting. Only in the case where inhibition is such as to make all the dominant eigenvalues negative will the threshold have no bearing on the stability of the steady states.

The steady states of this model are the same as the fixed points of a network of graded-response neurons with governing equation

$$\frac{dI_i}{dt} = -\theta I_i + \sum_j \mathrm{J}_{ij}\nu_j + I_i^{ext} \tag{26}$$

and gain function $g(I_i) = \nu_i$ where

$$g(I) = \begin{cases} 0, & I \leq 0 \\ I & I > 0. \end{cases} \tag{27}$$

This describes a network of threshold-linear graded-response neurons [20] [19]. In the case we studied where all synapses have the same time course for the synaptic response function and all matrices $\boldsymbol{J}_f$ have a dominant eigenvalue, the stability of these fixed points/steady states is the same in both models (Appendix C). However, we gave an example to show that when different time courses are allowed, the stability criteria may be different. While it is possible to include time delays in the graded response model, the full time dependence of the synaptic response function cannot be incorporated. This can be crucial in determining the stability of the steady state, so the two models are not equivalent in this respect: the stability of the leakless integrate-and-fire model depends more subtly on a number of biologically relevant parameters. The techniques described in this paper may be useful for studying this effect for other (more realistic) gain functions by linearising around the steady state.

While the fixed points of the graded-response model are completely characterised by the mean firing rates of that state, in the leakless integrator model there is the possibility that there may be more than one attractor with different sequential firing times (for a fixed $\boldsymbol{I}^{ext}, \theta$ and $\boldsymbol{J}$) giving rise to the same mean rates. This has been demonstrated by the theoretical work of Hertz and Hopfield [10] for a particular version of the leakless integrate-and-fire model.

Further work involves looking at specific forms of the synaptic matrix $\boldsymbol{J}$, for example one arising from a Hebbian learning rule, and investigating the ability of the network to act as an associative memory storage device. (Previous work on memory storage starting from a network of integrate-and-fire neurons includes [1] [2] [21] [4].) More investigation is needed in the case where $\boldsymbol{J}$ does not have a dominant eigenvalue or the synaptic response functions have different time courses: then oscillatory behaviour is possible. It is also desirable to extend the present analysis to the leaky case, as this is closer to the biological situation.

6 Acknowledgments

We thank Professor John Robinson for discussions during the early stages of this work. We also thank Dr Tony Burkitt for comments on a draft of the manuscript.

7 Appendix A

In this appendix we prove the result that if mean firing rates remain bounded at all times then

$$(\boldsymbol{\theta} - \boldsymbol{J}_f)\overline{\boldsymbol{\nu}}_f = \boldsymbol{I}_f^{ext}, \tag{28}$$

and

$$\boldsymbol{J}_b\overline{\boldsymbol{\nu}}_f + \boldsymbol{I}_b^{ext} \leq \mathbf{0}, \tag{29}$$

where the subscripts f and b denote foreground and background respectively, as has been explained in more detail in the body of the script. The mean firing rate of a neuron i over the period $[t', t'')$ is defined to be

$$\overline{\nu}_i(t', t'') = \frac{r_i(t', t'')}{t'' - t'}, \tag{30}$$

where $r_i(t', t'')$ is the number of times neuron i fires during the interval $[t', t'')$. In equations (28) and (29), $\overline{\boldsymbol{\nu}}_f \equiv \overline{\boldsymbol{\nu}}_f(0, \infty)$. For $\overline{\nu}_i(t', t'')$ to be meaningful we insist on some minimum sampling period $s > 0$ and require that $(t'' - t') \geq s$. The network is in a bounded state if there is some fixed quantity $\mu > 0$ such that

$$\overline{\nu}_i(t', t'') < \mu, \tag{31}$$

for all i and $(t'' - t') \geq s$.

There are two important points about this result. First, it allows us to verify the equations for the steady states, without resorting to approximation. Second, it is statement about the global stability of the network: if there are no positive solutions to equations (28) and (29) then the mean firing rates must become unbounded. That is, stable dynamics that are more complicated that a steady state, such as oscillatory or chaotic dynamics,are not possible.

To prove the result we calculate the mean current

$$\overline{I_i}(t', t'') = \frac{1}{t'' - t'} \int_{t'}^{t''} I_i(t)dt. \tag{32}$$

in two ways. First, we consider it as sum of the output of all neurons to obtain the relation

$$\overline{\boldsymbol{I}} = \boldsymbol{J}_b\overline{\boldsymbol{\nu}}_f + \boldsymbol{I}_b^{ext}, \tag{33}$$

where $\overline{\boldsymbol{I}} \equiv \overline{\boldsymbol{I}}(0,\infty)$. From equation (3),

$$\begin{aligned}
t\bar{I}_i^{int}(t) &= \sum_{j,k}\int_0^t p_{ij}(t'-t_j^k)\,dt' + \int_0^t p_{i0}(t) \\
&= \sum_{j,k}\int_0^\infty p_{ij}(t'-t_j^k)\,dt' - \sum_{j,k}\int_t^\infty p_{ij}(t'-t_j^k)\,dt' + \int_0^t p_{i0}(t) \\
&= \sum_j \mathrm{J}_{ij} r_j(0,t) - \sum_{j,k}\int_t^\infty p_{ij}(t'-t_j^k)\,dt' + \int_0^t p_{i0}(t).
\end{aligned}$$

The desired result follows upon showing that the second term in the last line is bounded as $t \to \infty$. For each j (in the following we drop the suffixes i, j)

$$\begin{aligned}
|\sum_{k=0}^{r(0,t)}\int_t^\infty p(t'-t^k)dt'| &\le \sum_{k=0}^{r(0,t)}\int_t^\infty |p(t'-t^k)|dt' \\
&= \sum_{k=0}^{r(0,t)}\int_{t-t^k}^\infty |p(t'')|dt'' \\
&= \sum_{k:t-t^k>s}\int_{t-t^k}^\infty |p(t'')|dt'' + \sum_{k:t-t^k\le s}\int_{t-t^k}^\infty |p(t'')|dt'' \\
&\le \sum_{k=0}^{r(0,t)} h(\frac{r(0,t)-k}{\mu}) + \mu s|J_{ij}|
\end{aligned}$$

The justification for the last step is as follows. In the first sum set $h(t) = \int_t^\infty |p(t'')|dt''$; then, since $(t-t^k) > s$, inequality (31) $\Rightarrow (t-t^k) > r(t,t^k)/\mu = (r(0,t)-k)/\mu$ and since $h(t)$ is positive and monotone decreasing we have $h(t-t^k) < h(r(0,t)-k)/\mu$. In the second sum there are $r(t-s,t)$ terms. For each term $\int_{t-t^k}^\infty |p(t'')|dt'' \le \int_0^\infty |p(t'')|dt'' = |J_{ij}|$ and inequality (31) $\Rightarrow r(t-s,t) < \mu s$. Thus in the limit as $t \to \infty$,

$$|\sum_{k=0}^{r(0,t)}\int_t^\infty p(t'-t^k)dt'| \;\le\; \sum_{m=0}^{\infty} h(m/\mu) + \mu s|J_{ij}| < \infty.$$

By the Poisson summation formula this sum is bounded because h is continuous and absolutely integrable. [25]

Now we consider the current as the input leading to firing of a neuron, and so derive a second equation for the mean current. Using equation (6)) it is straight forward to derive an equation analogous to equation (8):

$$\overline{I_i}(t',t'') = \theta\frac{r_i(t',t'')}{t''-t'} + \frac{V_i(t'')-V_i(t')}{t''-t'}. \tag{34}$$

One of two things can happen. The neuron may be in the foreground, in which case it continues to fire indefinitely and so there is an infinite subsequence of firing times at which $V_i(t)$ is bounded (the firing times t_i^k). Since $\bar{\nu}_i$ has a limit we have

$$\bar{\boldsymbol{I}}_f = \theta \overline{\boldsymbol{\nu}}_f, \tag{35}$$

which together with equation (32) gives equation (28). Alternatively, the neuron may be in the background, in which case the mean afferent current is negative, giving inequality (29).

8 Appendix B

This appendix contains the material pertaining to the stability analysis of the steady states.

8.1 Singularities of $\hat{\boldsymbol{\nu}}(s)$

The first point that requires justification is that $\hat{\boldsymbol{\nu}}(s)$ is meromorphic. This follows from our assumption that the Laplace transform $\hat{p}_{ij}(s)$ is meromorphic. Since $(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))^{-1}$ is a rational function of these functions, it is also meromorphic and hence $\hat{\boldsymbol{\nu}}(s)$ is too. We also need to clarify what we mean by "Any other poles arise from the factor $(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))^{-1}$: they are typically the zeros of the function $D(s) = \det((\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))$". First we show that any zero, ξ, of $D(s)$ is a pole of at least one $\hat{\nu}_i(s)$. The only time this would not be true is if the numerator of

$$(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))^{-1}\boldsymbol{I}^{ext} = \frac{\mathrm{adj}(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))\boldsymbol{I}^{ext}}{\det((\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))} \tag{36}$$

was also zero at ξ. But then $\boldsymbol{I}^{ext}$ is an eigenvector of $\mathrm{adj}(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))$ with eigenvalue zero, which is impossible since $\det(\mathrm{adj}(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))) = 1$. Now consider any poles, ζ, of $\hat{\boldsymbol{\nu}}(s)$ that are not zeros $D(s)$. These must arise from the poles of $\hat{p}_{ij}(s)$, since otherwise $(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))$ is a finite invertible matrix so $\hat{\boldsymbol{\nu}}(s)$ is also finite. But a pole arising from the element $\hat{p}_{ij}(s)$ will appear to equal order in the determinant $D(s)$, and hence cancel, unless its cofactor in $(\boldsymbol{\theta} - \hat{\boldsymbol{P}}(s))$ is zero. This will not typically be the case; it requires a particular value of θ. Furthermore, even if it does occur, any poles arising from $\hat{p}_{ij}(s)$ must have negative real part since the function is a Laplace transform convergent for $Re(s) > 0$, so they will not affect the stability.

8.2 Stability of a purely excitatory network

We now prove the stability result for networks with purely excitatory synapses; that is, the steady state is stable if and only if $\theta > \Theta$. We know that the steady state is stable if and only if all the zeros of $D(s)$ have negative real part. First suppose that $\theta > \Theta$, and seeking a contradiction, hypothesise a zero, ξ, of $D(s)$ with non-negative real part. Then θ is an eigenvalue of $\hat{\boldsymbol{P}}(\xi)$. Consider the non-negative matrix $\hat{\boldsymbol{P}}^{abs}(\xi)$ formed by taking the absolute values of the elements of $\hat{\boldsymbol{P}}(\xi)$, and let its P-F eigenvalue be $\Theta^{abs}(\xi)$. Then a theorem of Frobenius tell us that $\theta \leq \Theta^{abs}(\xi)$ ([17], p. 25). Consider an element of the matrix $\hat{\boldsymbol{P}}^{abs}(\xi)$:

$$
\begin{aligned}
|\hat{p}_{ij}(\xi)| &= |\int_0^\infty dt e^{-\xi t} p_{ij}(t)| \\
&\leq \int_0^\infty dt |e^{-\xi t}||p_{ij}(t)| \\
&= \int_0^\infty dt e^{-Re(\xi) t} p_{ij}(t) \\
&\leq \int_0^\infty dt p_{ij}(t) \\
&= J_{ij}.
\end{aligned}
$$

We have used the assumption that $Re(\xi) > 0$. So $\hat{\boldsymbol{P}}^{abs}(s) \leq \hat{\boldsymbol{P}}^{abs}(0) = \boldsymbol{J}$ (elementwise) and so another theorem on non-negative matrices tells us then that $\Theta^{abs}(\xi) \leq \Theta$ ([17], p.2). Now we have $\theta \leq \Theta^{abs}(\xi) \leq \Theta < \theta$ which is a contradiction.

To show that there is always a zero of $D(s)$ with non-negative real part whenever $0 < \theta \leq \Theta$, we need only note that $\hat{\boldsymbol{P}}(s)$ is non-negative for positive real s, so there is always a P-F eigenvalue $\Theta(s)$ in this case. This decreases continuously and monotonically from Θ to zero along the positive real axis because $\hat{\boldsymbol{P}}(s) > \hat{\boldsymbol{P}}(s')$ if $s > s'$.

8.3 Stability of a network with a dominant eigenvalue for $\boldsymbol{J}$ and identical time courses for $p_{ij}(t)$

In the case where there are inhibitory synapses, with the assumptions that $\boldsymbol{J}_f$ has a dominant eigenvalue Θ_f and all synapses have the same time course, $p_{ij}(t) = J_{ij} g(t)$, we stated that steady state for this partition is stable if and only if $\theta > \Theta_f$. We shall prove this now. Following the proof of the excitatory case, we let $\theta > \Theta_f$, and suppose that there is a zero, ξ, of $D_f(s)$ with non-negative real part. Note that if

$\lambda_1, \ldots, \lambda_N$ are the eigenvalues of $\boldsymbol{J}_f$ then we may factorise $D_f(s) = \Pi_i(\theta - \lambda_i \hat{g}(s))$. Therefore $\theta = \lambda_i \hat{g}(\xi)$ for some i. Noting that θ is a real positive number we have

$$\begin{aligned} \theta &= |\lambda_i||\hat{g}(\xi)| \\ &\leq \Theta_f \end{aligned}$$

where we have used the facts that Θ_f is a dominant eigenvalue and $|\hat{g}(\xi)| \leq 1$ for $Re(\xi) \geq 0$. This is a contradition, so all the zeros of $D_f(s)$ have negative real part for $\theta > \Theta_f$. If on the other hand $0 < \theta \leq \Theta_f$, then there is always a solution to $\theta - \Theta_f \hat{g}(s) = 0$ with positive real part, since $\hat{g}(s)$ is a continuous positive real function on the positive real axis with $\hat{g}(0) = 1$ and $\hat{g}(\infty) = 0$.

We also stated that this stability condition continues to hold for any synaptic response function, $p_{ij}(t)$, peturbed from the original by shifting it to the left or right by a time less than $\epsilon > 0$. We shall also prove this. Let the Laplace transform of the matrix of peturbed response functions be $\hat{\boldsymbol{P}}_f(s)$. It has elements $p_{ij}(s) = \exp(-\epsilon_{ij}s)\hat{g}(s)J_{ij}$, where ϵ_{ij} is the time shift of magnitude less that ϵ. Suppose that $\theta > \Theta_f$ and let ξ be a zero of $D_f(s) = \det(\theta - \hat{\boldsymbol{P}}_f(s))$ with non-negative real part. We still have that $\hat{\boldsymbol{P}}_f(0) = \boldsymbol{J}_f$, which has dominant eigenvalue Θ_f. Assuming it is not repeated, it may be continued analytically as an eigenvalue of $\hat{\boldsymbol{P}}_f(\xi)$ in some disc $B(\delta)$ of radius δ centred on the origin:

$$\Theta_f(\xi) = \Theta_f + \boldsymbol{y}^T \hat{\boldsymbol{P}}'_f(0)\boldsymbol{x}\xi + \mathrm{O}(\xi^2), \tag{37}$$

where $\boldsymbol{y}^T$ and $\boldsymbol{x}$ are left and right eigenvectors of $\boldsymbol{J}_f$ respectively, such that $\boldsymbol{y}^T\boldsymbol{x} = 1$, and $\hat{\boldsymbol{P}}'_f(0)$ is the derivative of $\hat{\boldsymbol{P}}_f(s)$ with respect to s at 0 ([13], pp. 383-405). Since θ is an eigenvalue of $\hat{\boldsymbol{P}}_f(\xi)$ we first consider the possiblity that it described by the equation (37) for some $\xi \in B(\delta)$. Since $\hat{\boldsymbol{P}}'_f(0) = \hat{g}'(0)\boldsymbol{J}_f + \boldsymbol{\gamma}$, where $\gamma_{ij} = \epsilon_{ij}J_{ij}$, we have

$$\theta = \Theta_f + \Theta_f(\hat{g}'(0) + \mathrm{O}(\epsilon))\xi + \mathrm{O}(\xi^2). \tag{38}$$

Now $\hat{g}'(0) < 0$, so we can choose an ϵ such that $(\hat{g}'(0) + \mathrm{O}(\epsilon)) < 0$. Since θ is real so must be ξ, and so our assumption that $Re(\xi) \geq 0$, becomes $\xi \geq 0$, which implies that $\theta \leq \Theta_f$, which is a contradiction.

Now consider the possibility that θ is some other eigenvalue of $\hat{\boldsymbol{P}}_f(\xi)$, either with $\xi \notin B(\delta)$ or continued from $\lambda_i \neq \Theta_f$. Note that

$$|\hat{p}_{ij}(s) - \hat{g}(s)J_{ij}| = |\int_0^\infty dt\, \mathrm{e}^{-st}(p_{ij}(t) - g(t)J_{ij})|$$

$$
\begin{aligned}
&\leq \int_0^\infty dt\,|p_{ij}(t) - g(t)J_{ij}| \\
&= |J_{ij}| \int_0^\infty dt\,|g(t-\epsilon_{ij}) - g(t)| \\
&= |J_{ij}\epsilon_{ij}| \int_0^\infty dt\,|g'(t) + \mathrm{O}(\epsilon_{ij})| \\
&= 2|J_{ij}\epsilon_{ij}|(\sum_{\text{local max.}} g(t) - \sum_{\text{local min.}} g(t)) + \mathrm{O}(\epsilon_{ij}^2) \\
&\leq \mathrm{O}(\epsilon)
\end{aligned}
$$

We have assumed that $g(t)$ is at least piecewise differentiable and has finitely many local extrema. From the general theory of peturbed matrices we can guarentee that $|\theta - \hat{g}(\xi)\lambda_i| = \mathrm{O}(\epsilon)$ ([13], p. 388). Since $\sup_{\lambda_i \neq \Theta_f, Re(\xi)>0}(\Theta_f - \hat{g}(\xi)\lambda_i) > 0$ and $\sup_{\xi \notin B(\delta), Re(\xi)>0}(\Theta_f - \hat{g}(\xi)\lambda_i) > 0$, we can always find an ϵ such that $\theta < \Theta_f$, which is again a contradiction. So there can not be any zeros ξ of $D_f(s)$ with non-negative real part when $\theta > \Theta_f$.

Conversly, we can argue along similiar lines as we have just done that the eigenvalue $\Theta_f(s)$, perturbed from $\Theta_f\hat{g}(s)$, is never repeated for any s. Then it can be shown that it is analytic every where in the half plane $Re(s) > 0$ and that the equation (37) and its extensions apply on this plane (this is based on the theory in [13] and the theory of analytic continuation as detailed in [15], for example). This implies that $\Theta_f(s)$ is real and continuous on the positive real axis, such that $\Theta_f(0) = \Theta_f$ and $\Theta_f(\infty) = 0$. So there is always a zero of $D_f(s)$ with non-negative real part for $0 < \theta \leq \Theta_f$.

8.4 Other conditions for stability

We prove that a condition sufficient for stability of the steady state for any matrix $\boldsymbol{J}_f$ is that $\theta > \Theta_f^{abs}$, where Θ_f^{abs} is the P-F eigenvalue of the matrix, $\boldsymbol{J}_f^{abs}$, obtained by reversing the sign of any inhibitory synapses of $\boldsymbol{J}_f$. We further prove that a set of extreme delays can be chosen to make the condition necessary.

That the condition is sufficient follows directly from the sufficiency proof of the excitatory case if we replace Θ by Θ_f^{abs} and $\boldsymbol{J}$ by $\boldsymbol{J}_f^{abs}$ throughout. To choose a set of delays such that the condition is necessary, let the delays, τ, to inhibitory synapses be very long and the delays to the remaining synapses be zero. The introduction of the delay τ at an inhibitory synapse, (i,j), means that the Laplace transform of the response function becomes $\hat{p}_{ij}^{delay}(s) = \exp(-\tau s)\hat{p}_{ij}(s)$. Setting $s = \pi i/\tau$ we get

$$\begin{aligned} \hat{p}_{ij}^{delay}(s) &= -\hat{p}_{ij}(\pi i/\tau) \\ &\to -J_{ij} \quad \text{as } \tau \to \infty. \end{aligned}$$

So the limit of large inhibitory delays has the effect of transforming $\boldsymbol{J}_f$ to $\boldsymbol{J}_f^{abs}$ and the result follows.

Finally we show that the condition is necessary and independent of the choice of synaptic response functions precisely when $\boldsymbol{J}_f = \boldsymbol{D}\boldsymbol{J}_f^{abs}\boldsymbol{D}$. If $\boldsymbol{J}_f$ is of this form then it is clearly necessary and independent of the choice of synaptic response function, since $\boldsymbol{J}_f$ and $\boldsymbol{J}_f^{abs}$ share the same eigenvalues. Conversely, if the condition is always necessary there exists a ξ with non-negative real part such that $\hat{P}_f(\xi)$ has an eigenvalue Θ_f^{abs}. Then a theorem of [17] tell us that $\hat{P}_f(\xi) = \boldsymbol{D}\boldsymbol{J}_f^{abs}\boldsymbol{D}^{-1}$ where $\boldsymbol{D}$ is the a matrix with diagonal entries of absolute value one, and zeros elsewhere. Since $|\hat{p}_{ij}(\xi)| < J_{ij}$ for $\xi \neq 0$, we conclude that $\xi = 0$ and $\boldsymbol{J}_f = \boldsymbol{D}\boldsymbol{J}_f^{abs}\boldsymbol{D}$ where $\boldsymbol{D}$ now has diagonal entries that are ± 1.

9 Appendix C

This appendix is concerned with the stability of a graded response network.

Consider the system given by equation (26) for a particular partition (f, b) of the neurons into foreground and background, that is, $I_i > 0$ if neuron i is in the foreground and $I_i \leq 0$ if neuron i is in the background. Then according to equation (27) the system of differential equations becomes linear

$$\frac{d\boldsymbol{I}}{dt} = -(\boldsymbol{\theta} - \boldsymbol{J}_{(f,b)})\boldsymbol{I} + \boldsymbol{I}^{ext} \tag{39}$$

where $\boldsymbol{J}_{(f,b)}$ is $\boldsymbol{J}$ with the background columns replaced by zeros. Assuming that $\boldsymbol{J}_{(f,b)}$ is diagonalisable, then

$$\boldsymbol{D}_{(f,b)} = \boldsymbol{P}^{-1}\boldsymbol{J}_{(f,b)}\boldsymbol{P} \tag{40}$$

where $\boldsymbol{D}_{(f,b)}$ is the diagonal matrix of eigenvalues of $\boldsymbol{J}_{(f,b)}$ and $\boldsymbol{P}$ is the column-wise matrix of normalised eigenvectors. Note that every eigenvalue of $\boldsymbol{J}_f$ is an eigenvalue of $\boldsymbol{J}_{(f,b)}$ and the remaining eigenvalues of $\boldsymbol{J}_{(f,b)}$ are all zero. This leads to the following solution to equation (39):

$$\begin{aligned}\boldsymbol{I}(t) &= \boldsymbol{P}\exp\{-(\boldsymbol{\theta}-\boldsymbol{D}_{(f,b)})\boldsymbol{t}\}\boldsymbol{P}^{-1}\boldsymbol{I}(0) \\ &\quad +[\mathbf{1}-\boldsymbol{P}\exp\{-(\boldsymbol{\theta}-\boldsymbol{D}_{(f,b)})\boldsymbol{t}\}\boldsymbol{P}^{-1}][\boldsymbol{\theta}-\boldsymbol{J}_{(f,b)}]^{-1}\boldsymbol{I}^{ext}, \quad (41)\end{aligned}$$

provided that for all times t' prior to and including t, $I_i(t') > 0$ if neuron i is in the foreground and $I_i(t') \leq 0$ if neuron i is in the background. $\boldsymbol{t}$ is the $N \times N$ diagonal matrix with elements t and $\mathbf{1}$ is the $N \times N$ identity matrix. This solution has a unique stationary solution, $(\boldsymbol{\theta}-\boldsymbol{J}_{(f,b)})^{-1}\boldsymbol{I}^{ext}$ if and only if the foreground and background components are greater than and less than or equal to zero, respectively. It is easily seen that this gives same set of conditions for steady states as given by equations (28) and (29). According to equation (41), this solution is stable if and only if $\theta - \lambda_k > 0$ for all eigenvalues λ_k of $\boldsymbol{J}_{(f,b)}$. The eigenvalues of $\boldsymbol{J}_{(f,b)}$ and $\boldsymbol{J}_f$ are the same or zero, and so if $\boldsymbol{J}_f$ has a dominant eigenvalue this is equivalent to the stability criterion given in the text, namely $\theta > \Theta_f$.

References

[1] 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: Comput. Neural Syst.* **2** 259-273

[2] Amit D J Tsodyks M V 1991 Quantitative study of attractor neural networks retrieving at low spike rates: II. Low-rate retrieval in symmetric networks *Network: Comput. Neural Syst.* **2** 275-294

[3] Budelli R, Torres J, Catsigeras E and Enricht H. 1991 Two neuron networks: I Integrate and fire pacemaker models. *Biol. Cybern.* **66** 95-110

[4] Burkitt A N 1994 Attractor neural networks with excitatory neurons and fast inhibitory interneurons at low spike rates *Network: Comput. Neural Syst.* **5** 437-448

[5] Burkitt A N 1996 Retrieval properties of attractor neural networks that obey Dale's law using a self-consistent signal-to-noise analysis. *Network: Comput. Neural Syst.* **7** 517-532

[6] Cox D R and Miller H D 1965 *The Theory of Stochastic Processes* (London: Chapman and Hall)

[7] Gomez L and Budelli R 1996 Two neuron networks: II Leaky integrator pacemaker models. *Biol. Cybern.* **74** 131-137

[8] Gerstner W and Van Hemmen J. L 1992 Universality in neural networks: the importance of the 'mean firing rate' *Biol. Cybern.* **67** 195-205

[9] Hansel D, Mato G, Meunier C 1995 Synchrony in excitatory neural networks *Neural Comp.* **7** 307-337

[10] Herz A V M and Hopfield J J 1995 Earthquake Cycles and Neural Reverberations: Collective Oscillations in Systems with Pulse-Coupled Threshold Elements *Phys. Rev. Lett.* **75** 6 1222-1225

[11] Hopfield J J 1984 Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Nat. Acad. Sci. USA***81**, 3088-3092.

[12] Knight B W 1972 Dynamics of Encoding in a Population Neurons *J. Gen. Physiol.* **59** 734-766

[13] Lancaster P and Tismenetsky, 1985 *The Theory of Matrices: Second edition with applications* (Academic Press)

[14] Mirollo R E and Strogatz S H 1990 Synchronisation of pulse-coupled biological oscillators *SIAM J.Appl. Math.* **50** 1645-1662.

[15] Nevanlinna R and Paatero V 1969 *Introduction to Complex Analysis* (Addison-wesley Publishing Company)

[16] Segev I 1992 Single neurone models: oversimple, complex and reduced *Trends in the Neurosciences* **15** 414-421

[17] Seneta E 1973 *Non-negative Matrices* (London: Allen and Unwin)

[18] Sofky W and Koch C 1995 Single-cell models *The Handbook of Brain Theory and Neural Networks* ed M A Arbib (Cambridge, MA: MIT Press) pp 879-884

[19] Stimmen M, Treves A and Rolls E T 1996 Pattern retrieval in threshold-linear associative nets *Network: Comput. Neural Syst.* **7** 109-122.

[20] Treves A 1990 Graded-response neurons and information encodings in autoassociative memories *Phys. Rev. A* **42** 2418-2430

[21] Treves A 1993 Mean-field analysis of neuronal spike dynamics *Network: Comput. Neural Sys.* **4** 259-284

[22] Tuckwell H C 1988 *Introduction to Theoretical Neurobiology, vols 1 and 2* (Cambridge: Cambridge University Press)

[23] Van Vreeswijk C, Abbott L F and Ermentrout G B 1994 When inhibition not excitation synchronises neural firing *J. Comp. Neurosci.* **1** 313-321

[24] Wilson M A and Bower J M 1989 The simulation of large-scale neural networks *Methods in Neuronal Modeling - From Synapses to Networks* ed C Koch and I Segev (Cambridge MA: MIT Press) pp 291-333

[25] Zygmund A 1968 *Trigonometric Series: Volume I* (Cambridge: Cambridge University Press) pp 68