Global spontaneous activity and local structured (learned) delay activity in cortex

Daniel J. Amit
Racah Institute of Physics, Hebrew University, Jerusalem, and
INFN, Sezione di Roma, Istituto di Fisica
Universita di Roma, La Sapienza, Ple Aldo Moro, Roma
and
Nicolas Brunel
INFN, Sezione di Roma, Istituto di Fisica
Universita di Roma, La Sapienza, Ple Aldo Moro, Roma
January 28, 1995

Summary and conclusions

1. We investigate the conditions under which cortical activity alone makes spontaneous activity self-reproducing and stable against fluctuations of spike rates. Invoking simple assumptions about properties of integrate-and-fire neurons it is shown that the stochastic background activity, of 1-5 spikes/second, cannot be stabilized when all neurons are excitatory.

2. On the other hand, spontaneous activity becomes self-stabilizing in presence of local inhibition: given reasonable values of the parameters of the network spontaneous activity reproduces itself and small fluctuations in the rate are suppressed.

   a. If the integration time constants of excitatory and inhibitory neurons at the soma are equal, *local* excitatory and inhibitory inputs to a neuron must balance to provide *local* stablility.

   b. If inhibition integrates faster its synaptic inputs, spontaneous activity is stable even when local recurrent excitation predominates.

3. In a network sustaining spontaneous rates of 1-5 spikes/second, we study the effect of learning in a local module, expressed in synaptic modifications in specific populations of synapses. We find:

   a. Initially no stimulus specific delay activity manifests itself. Instead, there is a delay activity in which, locally, *all* neurons selective to any of the stimuli learned have rates which gradually increase with the amplitude of synaptic potentiation.

   b. When the average LTP increases beyond a critical value, specific local attractors appear abruptly against the background of the global uniform spontaneous attractor. This happens with either gradual or discrete stochastic LTP.

4. The above findings predict that in the process of learning unfamiliar stimuli, there is a stage in which all neurons selective to any of the learned stimuli enhance their spontaneous activity relative to the rest. Then, abruptly, selective delay activity appear. Both facts could be observed in single unit recordings in delayed match to sample experiments.

5. Beyond this critical learning strength the local module has two types of collective activity. It either participates in the global spontaneous activity, or it maintains a stimulus selective elevated activity distribution. The particular mode of behavior depends on the stimulus: if it is unfamiliar, the activity is spontaneous; if similar to a learned stimulus, the delay activity is selective. These new attractors (delay activities) reflect the synaptic structure developed during learning. In each of them a small population of neurons have elevated rates, 20-30 spikes/second, depending on the strength of LTP. The remaining neurons of the module have their activity at spontaneous rates.

1 Introduction

It has recently been pointed out[1, 2, 3] that the apparently low background spontaneous activity in cortex, about 1–5 spikes/sec (s^{-1}), is high and prominent when looked at from the vantage point of the post-synaptic neuron. The reason is that the connectivity in neo-cortex is very high (of the order of 20,000 synapses per neuron[4, 5]), so a neuron receives an enormous number of afferent spikes per second, before it engages in any, stimulus related, computational activity. In [1, 2] advanced analytical and numerical methods were employed to evaluate in detail, within cable theory [6], the effect of this high background input on neural physiological parameters, as well as on the single neuron's temporal performance as a detector of synchronization.

Ref. [3] emphasized the effects of this high background on the performance of the network as an attractor neural network. Observing that the additional afferents due to a reverberation provoked by a stimulus are small in comparison with the background, it was concluded that when operating as attractor networks the neurons can be described as linear elements *on the input side*. This has been an underpinning of essentially all neural network modeling (for a review see e.g. [7]).

The common conclusion to all three studies is that neurons operating in the cortex have their physiological parameters significantly renormalized by the flux of spontaneous activity. Another common assumption is that at this low activity rate, the spikes arriving on different synapses are uncorrelated. As a consequence the modification of the neurons becomes stationary, and networks in a living cortex are assemblies of effective elements, rather remote from the isolated neurons in a slice. The parameters of the renormalized neuron can, in principle, be computed from those of the bare one (see e.g. [1, 2]), or can be determined phenomenologically.

The prominent presence and the essential effect of the background neural activity, renews the question of the origin of this activity in the neural dynamics and of the mechanisms that keep it stable. Why would the cortex be stable at a rate of about 1–5 s^{-1}? One could invoke various deus ex machina. But an explanation in terms of the system's own dynamical rules and parameters would be more satisfying. In what follows we shall argue that such an explanation is available. What makes this program non-trivial is the fact that it must deal with many aspects concurrently:

- Since spontaneous activity is non specific and is global, it must be an attractor which is global on the cortical level.
- The existence of such an attractor implies
  - That the spontaneous rates at the observed levels can reproduce the same rates when considered as afferents in plausible cortical conditions;
  - That these rates are stable, i.e. that local fluctuations in rates, due either to variations in local spontaneous activity or to local selective activity do not disturb the global balance.

- That on top of this great cortical attractor, local variations in synaptic connectivity allow for selective (stimulus dependent) activity, either during the presence of a stimulus, or following it.
- That in the local, selective attractor neurons whose rates are not significantly enhanced maintain a spiking rate at levels similar to the global spontaneous activity.
- The selective rates in a local attractor may be close to the spontaneous rates [8, 9, 10, 11].

1.1 Assumptions

We shall make the following assumptions:

- Background activity is stochastic.
- In the absence of a stimulus, neurons in the cortex all see the same environment, apart from statistical fluctuations.
- The spontaneous activities of different neurons are uncorrelated.
- The marginal synaptic inputs to every cell, above the mean input due to spontaneous activity, is the linear sum of individual synaptic contributions. In fact, we shall make a stronger assumption, namely that the full input is linear. In Appendix D we discuss the implications of the possibility that the input be linear only in the marginal increments.
- The dynamics of both spontaneous and selective activity can be described by approximating the synaptic input by a Gaussian stationary process.
- That a reliable description is available in terms of spike rates and afferent currents, rather than actual spikes. This implies the supposition that neural activity does not synchronize in the range of biological parameters.

The first assumption is an expression of the empirical character of spontaneous activity, as well as a condition under which we can control the study of its stability. We conceive the origin of the stochasticity in the unknown and sporadic inputs that affect remote parts of the cortex.

The second expresses the position that the spontaneous activity is non-specific: Whatever variation among neurons exists, it is due entirely to statistical fluctuations in the distribution of synapses. It also expresses the view that spikes emitted in this dynamical phase are due to fluctuations in the afferent input arriving at a neuron. It is exactly these fluctuations which we shall estimate below, in order to determine the stability of this underlying activity state. A similar position was taken by Nelken[12], in a somewhat different context.

The third assumption is almost a consequence of the articulation of the second, combined with two physiological facts: first that the typical interspike interval of each neuron

is very long compared to neural integration and leak times, which are at most tens of milliseconds; and second that the cortex is weakly coupled, i.e. a large number (of the order of 100) of input spikes are required to provoke an output spike [13, 14, 15, 16, 17]. About one half of these spikes are originating locally, the rest coming from remote areas[18]. Significant correlations would produce either much higher rates, or very long finely tuned, but very fragile, cycles[19].

The fourth assumption implies that the mean spontaneous afferent is producing an effective neuron, with modified parameters, and small variations about this mean rate lead to linear effects[3].

The next assumption is justified if many low-rate spikes arrive on a large number of uncorrelated channels, (hence a sum of many Poisson processes) per neural integration time[20]. In the case of excitatory as well as inhibitory inputs, each of the two inputs can be considered as an uncorrelated Gaussian process, due to the large number of incoming spikes of either type. The sum of the two processes can therefore be approximated by a Gaussian process, whose mean is the difference of the two means and whose variance is the sum of both variances[26]. In this approach the fluctuating synaptic input is not replaced by its average, and in this sense it is not a naive 'mean-field' approach. Rather, it is replaced by a Gaussian stationary process whose two first moments are those of the synaptic input.

This competition between excitation and inhibition has been recently raised[21, 22], leading to rather radical conclusions concerning the spike dynamics of the network. Here the context is different mainly because the spike rates expected and produced are low (much lower than the neuron's integration time constant), even when selective activity occurs. This simplifies matters, leaving spikes uncorrelated. As a consequence one can reduce the description to a network of current-rate elements, as suggested in the last assumption. In this context we study the attractors and their stability against small fluctuations.

1.2 Results

In the present study we investigate the conditions which can make spontaneous activity be stable against local fluctuations of rates, in a network of leaky integrate-and-fire neurons. We find that when all neurons are excitatory it is impossible to stabilize realistic spontaneous rates. The stable states have all neurons either quasi-silent or firing at elevated rates. If inhibitory neurons are included, spontaneous rates of 1-5 spikes per second can be stabilized, for reasonable values of the parameters of the network. If the integration time constants of the membrane potentials (of the spike emitter) for excitatory and inhibitory neurons are equal, stability requires that the average excitatory and inhibitory synaptic inputs to every neuron of the network closely balance. If inhibition integrates its depolarization a couple of times faster, this condition can be relaxed and realistic spontaneous rates are stable even if the excitatory inputs dominate.

We find that the spontaneous activity is globally stable. So is also the state with all neurons silent in absence of external inputs. But a 'cortex' that is in an active state of spontaneous activity will not drift gradually into the silent state. A major coherent

fluctuation on a very large scale will be required for such a transition.

In that same network, which can maintain spontaneous rates of the order of 1-5s $^{-1}$, we observe the effects of learning in a local module, as expressed in the potentiation of the excitatory efficacies in some specific population of synapses. It is found that when long term potentiation (LTP) takes place one should observe two phases. At low potentiation levels there is a non-selective enhancement of the rates of all cells that have been selective to any of the stimuli. Then, when the average potentiated synaptic efficacy exceeds a critical value, the *local* network develops rather abruptly a variety of local attractors on the background of the stability of the uniform global one. The local attractors reflect the structure implied by the learning process. In such an attractor a small population of neurons, representative of a particular stimulus, are active at elevated rates (of the order of 20s^{-1}), depending on the potentiated synaptic value. The neurons of the local module which do not participate in the active representation of the particular stimulus (background neurons) remain at spontaneous rates. Even in presence of LTP and selective attractors with elevated rates, the local module has always the option of a stable uniform activity at rates close to the spontaneous ones. This state of activity is chosen by the module when it is stimulated by a stimulus uncorrelated with what has been learned in the past. See also e.g. ref. [23]. This is a direct prediction for single cell recording during learning in delay match to sample experiments.

To recapitulate the above findings: On top of the global, uniform, low-activity 'vacuum' local bubbles can be formed by learning. The number of modules in which this can take place is large, and in each module the number of different, selective, elevated rate attractors is also large. Each of these modules, if formed, retains the option of the spontaneous attractor, which it may use to indicate lack of recognition. This is the richer sense in which the term 'vacuum' is introduced here. It is the structureless global state on top of which local excitations (which are metastable) can express structured events.[1].

It should be emphasized that besides the structural features of the synthesis of single neuron and network effects this modeling effort provides another important bonus. It allows for the coexistence of low attractor rates with relatively high spontaneous rates, as is the case in the recordings during delayed match to sample experiments[8, 9, 10, 11]. The relatively high spontaneous activities are present both in the wide cortical area as well as in the neurons of the local module which do not participate in the foreground of a given retrieved memory.

The organization of the paper is as follows. In Section 2 we describe the model network and its elements. In Section 3 we study the spontaneous stable states in a purely excitatory network, first using a simple intuitive argument, then in a network of integrate-and-fire neurons. In Section 4 inhibitory units are added and it is shown how they stabilize a global attractor in which neurons have realistic spontaneous rates. Then in Section 5 we investigate the effect of learning on the stable states of the network.

[1]For the way in which this metaphor is employed to describe the world of physics, see e.g. [24].

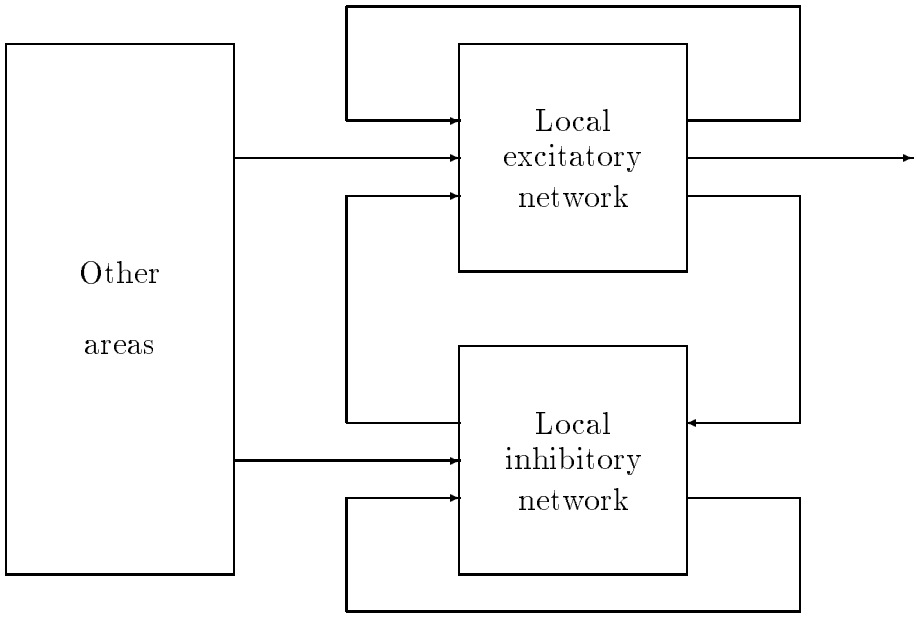

Figure 1: Connectivity scheme: Neurons in the local excitatory network (module) receive collateral connections from neurons in the same module and unstructured local inhibition. Both excitatory and inhibitory neurons in the module receive excitatory synaptic inputs from the wider cortical areas.

2 The model

2.1 The network

We consider a local module that is a large network, containing some 10^5 neurons (a cortical column) embedded in a vast sea of neurons (the entire cortex). We have in mind a local module of diameter of 1mm in a living cortex, as is observed in inferior temporal or prefrontal cortex[8, 11]. Two features distinguish the local module from the global network: the intensity of excitatory connectivity (high internally and low otherwise) and the range of the inhibitory neurons. Neurons receive three types of synaptic inputs: from recurrent (collateral) excitatory connections with other neurons in the same network; from inhibitory neurons in the module; from excitatory neurons in other unspecified areas. The latter are considered excitatory, since only pyramidal cells have long enough axons. These afferents represent unrelated activity in the rest of the cortex, and carries the stochastic component induced by sporadic inputs from remote parts of the cortex (see assumptions above). They are modeled as Poisson processes with a spontaneous spike rate which is kept fixed. Estimates as to the proportion of excitatory synaptic inputs on given neuron from local collaterals vary 50–80%. We will use the value 50%. A schematic representation of the connectivity is sketched in Fig. 1.

To highlight the difficulties and their potential resolutions we first consider the system

in the absence of the inhibitory feedback. Then, in Sections 4 and 5 the context is enlarged and the network includes both excitatory and inhibitory units. The collateral connectivity in the local network has no geometric structure: a neuron has equal probability of synapsing on any other neuron in the local module. To remain within biologically acceptable territory, the number of inhibitory units in the module is taken to be 20% of the number of local excitatory neurons.

2.2 The integrate-and-fire neuron

Dynamics of the membrane depolarization. Our network is composed of leaky integrate-and-fire neurons[25, 26], converted into an interacting system of synaptic currents and output spike rates (see e.g. ref. [20] and discussion below). The assumptions underlying this description of the neurons will be continuously scrutinized as we go along. Each neuron in the local network is characterized by its depolarization at the soma $V(t)$, which obeys to the integrator equation

$$\tau \dot{V}(t) = -V(t) + I(t) \tag{1}$$

where $\tau \sim 10$ms is the integration time of the membrane depolarization at the spike emiting part of the soma and $I(t)$ is the synaptic current charging that part of the membrane. The resting potential has been set to zero.

When the depolarization reaches the threshold θ the neuron emits a spike, and the potential is reset to a hyperpolarizing value H, after an absolute refractory period $\tau_0 \sim$2ms. In the situation we consider, this absolute refractory period will not be relevant since the mean interspike interval will be very long. Yet it fixes the absolute scale of the rates, since it is the inverse of the maximal rate. We choose the time unit as the integration time of the membrane depolarization, $\tau = 1$, and so $\tau_0 = 0.2$. With this time unit the maximal rate is 5 and a rate of 1 amounts to 100s^{-1}. Post-spike hyperpolarization will be set to zero in numerical studies. We have checked that at low rates its actual value is immaterial.

Synaptic inputs to the cells. A neuron receives some $C \sim 20,000$ synaptic contacts from other cells. A fraction $x \sim 0.5$ of these synapses come from the local module, while the remaining synapses come from outside. In the unstructured situation, the potential variability in synaptic efficacies on a neuron is schematized in the following way: the synaptic efficacies J_i $(i = 1, \ldots, C)$ on the dendrites of a given neuron are assumed to be randomly distributed, according to a distribution $P(J)$, with mean J and standard deviation $J\Delta$. This distribution can account for randomness in the synaptic transmission, different electrotonic distances of the synaptic sites on the dendrites, etc.

Dynamics of the afferent currents. The effective, afferent current $I(t)$, due to the temporal variation of the synaptic conductances provoked by afferent spikes and charging the spike-emitting integrator in Eq. 1, obeys the equation:

$$\tau' \dot{I}(t) = -I(t) + \sum_{i=1}^{C} J_i \sum_{\alpha} \delta(t_i^{\alpha} - t) \tag{2}$$

The sum over i is over the synaptic sites on the dendrites, while the sum over j is over all

spikes arriving at a given site, and t_i^α is the time of arrival of spike α at synapse i. τ' is the time constant of the conductance changes at the synaptic sites. See e.g. [28, 20].

If $\tau' \ll 1$, i.e. the time constant of the currents is much shorter than that of the potentials, one can replace I in Eq. (1) by the source term on right hand side of Eq. (2), i.e.

$$\dot{V} = -V + \sum_{i=1}^{C} J_i \sum_{\alpha} \delta(t_i^\alpha - t). \tag{3}$$

In this case, on a time scale of the order of 1 (τ), the depolarization $V(t)$ is a sum of unit contributions each with amplitude J_i.

Dynamics of a single EPSP. Let us assume that at $t = 0$ the membrane potential is at rest, $V = 0$, and a spike arrives on a synapse i with efficacy J_i. In the absence of other spikes, the potential, for $t > 0$, will be given by

$$V(t) = J_i \exp(-t)$$

This is the unitary EPSP in the model, when the time constant of the currents is very short. We shall choose the mean of J_i (denoted above J) as the unit of synaptic efficacy. J is then also the average EPSP. We therefore choose J as the unit of the depolarization and threshold.

Estimates of J range from 0.05 to 0.5mV [13, 14, 15, 16, 17]. We pick an average of 0.2mV and with a threshold of 20mV some 100 simultaneous synaptic events will reach threshold. The variability in amplitude of unitary EPSPs from different synapses is taken of the same order as their average amplitude, thus $\Delta \sim 1$ (see also e.g. ref. [14]).

Independent Poisson activation of the input channels. Suppose that each synapse is activated by an identical, independent Poisson process, for collateral as well as external afferents, and spikes arrive at each channel with rate ν. Suppose further that ν is low, but that in a unit time interval the number of arriving spikes is high due to the large number of input channels C, then the source of the depolarization will have a Gaussian distribution, and the output rate of the neuron is given by the solution of the first passage problem[25, 26], namely

$$\nu_{out} = \frac{1}{\langle T \rangle} = \left(\tau_0 + \int_{\frac{H-\mu(\nu)}{\sigma(\nu)}}^{\frac{\theta-\mu(\nu)}{\sigma(\nu)}} du \psi(u) \right)^{-1}, \tag{4}$$

where the function ψ is defined as

$$\psi(u) = \sqrt{\pi} \exp(u^2)[1 + \mathrm{erf}(u)] \tag{5}$$

where $\mu(\nu)$ and $\sigma(\nu)$ are the rate-dependent mean and RMS of the Gaussian distribution of EPSP's depolarizing the cell.

A few words concerning the assumptions made above: That the average activity is the same inside and outside the module is reasonable for pure spontaneous activity, which we assume to have no structure and variations are attributed to fluctuations. The rates arriving at each synapse on a given neuron are low, yet the number of synapses is so high

(see below) that the assumption of a high number of spikes arriving in an integration time of the depolarization is reasonable. The most tenuous assumption concerns probably the ratio of the two time constants. If the time constant of conductance changes is of the order or larger than that of the depolarization (see e.g. ref. [29, 30, 31]), the replacement leading to Eq. 3 is not valid and the conditions leading to the current-rate transduction function, Eq. 4, are not satisfied. We return to this question in Appendix C.

3 Excitatory unstructured network

3.1 Reproducibility vs stability in simplified context

To obtain a preliminary idea concerning the possible stable, self-reproducing rate distributions we start by making the following approximate treatment of the spike emission process described by Eq. 2: In every integration time interval $\tau(=1)$ we assume that there is no decay. In each of these intervals the depolarization V is equal to the sum of unitary inputs J_i arriving in that interval.[2] This sum would be a Gaussian variable. At the end of the interval either a spike is emitted and the depolarization becomes zero. Or no spike is emitted and still the depolarization is reset to zero, this time as an expression of the exponential decay of the integrator. The probability for the emission of a spike in such an interval is taken equal to the probability that the depolarization arrives at threshold. Both are quite crude approximations, eliminated below. But they do provide reasonable guides.

Within the above framework, if the afferent in the unit interval has mean μ and RMS σ and the threshold is θ, the probability that a spike be emitted in an interval $\tau(=1)$ is:

$$\Pr(\nu_{in}) = \int_{\alpha(\nu_{in})}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \tag{6}$$

which depends on the rate ν_{in} of the afferent spikes via the rate dependence of μ and σ entering

$$\alpha(\nu) = \frac{\theta - \mu(\nu_{in})}{\sigma(\nu_{in})}. \tag{7}$$

The resulting output spike rate is therefore given by

$$\nu_{out} = \Pr(\nu_{in}) \tag{8}$$

spikes per unit time, which is the integration time of the potential. In other words, the output rates are given in units of 100s^{-1}. If the resulting output spike rate is to be very low, we must have $\mu < \theta$. Otherwise the probability is greater than 0.5 per τ and if $\tau \sim$10ms the rate will be greater than 50s^{-1}.

In the unit interval the mean μ and standard deviation (SD) σ of the afferent source are:

[2] We take the neuron to be fully linear on the input side. See discussion in Appendix D.

$$\mu = C\nu, \quad \sigma = \sqrt{C\nu(1+\Delta^2)}, \tag{9}$$

where C is the number of synapses per neuron and the excitatory synapses have a distribution of efficacies with mean 1 and SD Δ, as in Section 2.2. For example, in cortical conditions, taking τ=10ms, C=20,000, Δ=0, i.e. uniform synapses and ν=0.02, i.e. 2s^{-1}, then

$$\mu = 400 \text{ and } \sigma = 20.$$

Thus the mean number of spikes arriving at a neuron per integration time is typically a few hundred, with fluctuations of a few tens.

Self-reproducibility

As was mentioned above, when the activity is purely spontaneous it is considered uniform globally. In that case the input to every neuron is equal on the average, irrespectively of what proportion of the synapses are from the local module and what from outside. Since it is the rate itself which determines both μ and σ, the reproducibility of the spontaneous rate distribution requires that the rate which determines the parameters be equal to the rate at which spikes are produced given those parameters, i.e. we must have

$$\nu = \Pr(\nu) = \int_{\alpha(\nu)}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \tag{10}$$

where J=1, and $\Delta = 0$ in Eqs. (9). Given the number of synapses C per neuron and the threshold θ, one can compute from Eq. 10 the value of the spontaneous spike rate which reproduces itself on average.

The function $\Pr(\nu)$ is represented in Fig. 2 for $C = 20,000$ and two values of θ. The solutions of Eq. (10) are the intersections of any of the curves with the straight line of slope unity in the figure. The intersections away from zero are at self-consistent average rates which are in the range of realistic spontaneous rates – 1–5s^{-1}.

The above formulation applies only when the spike rates are much lower than 1 (100s^{-1}). Then the emission probability per τ is much smaller than 1, and the probability for the emission of multiple spikes during an interval is negligible.

But are these rate distributions stable against local fluctuations?

Local stability

If a small fluctuation in the local mean rate displaces the rates in the system with no restoring force, this state is of little interest. Hence, we must look for solutions of Eq. (10) which are stable, i.e. which suppress fluctuations and restore the mean value of the spontaneous rates. For example, for the rate distribution to be stable against a fluctuation in the local mean rate, we must have that the change in the output rate is smaller than the change in the input rate that caused it. The external synaptic input is assumed fixed, with each synapse coming from outside the local module activated at a rate ν_{ext}. The logic is that the dynamics of the external rates is outside our description, and the global constraint is that all spontaneous rates be equal. Local synapses are again activated at mean rate ν. They constitute a fraction x of the synapses on any neuron.

The mean μ and standard deviation σ of the afferent input in a unit interval become:

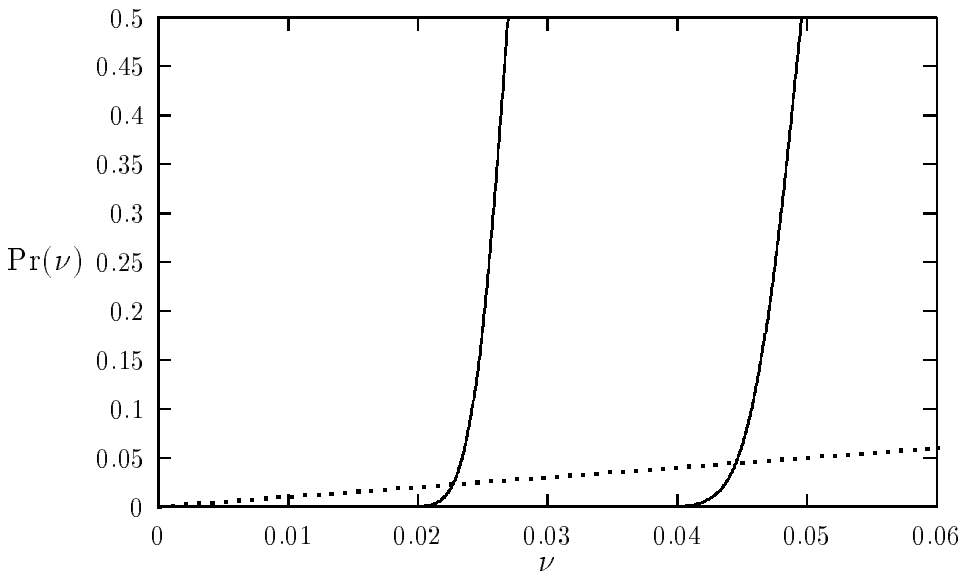

Figure 2: Numerical solution for self-reproducing rates. Full curves: function $\Pr(\nu)$ for $\theta =$ (from left to right) 500 and 1000. In both cases the self-reproducing rate is the intersection with the straight line of slope unity (dotted line) for low positive rates. In both cases it is highly unstable.

$$\mu = C[x\nu + (1-x)\nu_{ext}] \tag{11}$$

and

$$\sigma = \sqrt{C(1+\Delta^2)[x\nu + (1-x)\nu_{ext}]}. \tag{12}$$

The condition for self-reproducibility of rates is, Eq. 10,

$$\nu = \Pr(\nu) = \int_{\alpha(\nu)}^{\infty} \frac{dt}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \tag{13}$$

where $\alpha(\nu)$ is given by Eq. (7) with μ and σ given by Eqs. (11, 12). The condition for stability against fluctuations in the local rates is

$$\frac{d\Pr}{d\nu}(\nu_\star) < 1 \tag{14}$$

where $\nu_\star$ is the self-reproducing rate, the solution of Eq. (13). If this condition holds, then for $\nu > \nu_\star$ the probability of generating a spike is lower than the input rate and ν decreases. For $\nu < \nu_\star$ the probability of generating a spike is higher than the input rate, and ν will increase.

We set $x = 0.5$ [18], i.e. afferents come in equal proportion from the local module and from the outside. ν_{ext}, the external spike rate, is set at a fixed plausible value of spontaneous rate. Then the threshold is varied so that $\nu_\star = \nu_{ext}$ be a solution of Eq. 13. This would be a globally consistent self-reproducing rate. The stability of a given self-reproducing rate is investigated by varying $\Pr(\nu)$ with respect to ν, *at fixed* θ *and* ν_{ext}, to check the stability condition, Eq. (14).

It turns out that at the desired rates 0.01-0.05 (1-5s^{-1}) the condition (14) never holds. For Eq. (13) to have a solution at these rates one must have

$$\alpha(\nu_\star) \sim 2 \tag{15}$$

i.e. the mean synaptic input at about two standard deviations below the threshold. For high values of α one can approximate $\Pr(\nu)$, Eq. (13), by [34]

$$\Pr(\nu) \sim \frac{1}{\sqrt{2\pi}\alpha} \exp(-\frac{\alpha^2}{2})$$

which yields

$$\nu_\star \sim \frac{1}{\sqrt{2\pi}\alpha} \exp(-\frac{\alpha^2}{2})$$

The left hand side of Eq. (14) can therefore be approximated by

$$\alpha x \sqrt{C\nu} \left(1 + \frac{(1+\Delta^2)\alpha}{2\sqrt{C\nu}}\right)$$

For the parameters chosen, $C\nu$ is of order 100, and thus

$$\frac{d\Pr}{d\nu}(\nu_\star) \sim \alpha x \sqrt{C\nu} \gg 1$$

Thus there is no way of having a stable self-reproducing rate of this order. Figure 2 shows a graphical representation of the instability of the self-reproducing state. Any perturbation of the rates will drive the network either to a state where all neurons are essentially quiescent, or to a state of high activity, in which the assumptions of the analysis do not hold anymore. This is related to the fact that there are only excitatory elements in the network. In the quiescent state the activity is essentially zero. The network lacks regulatory units that would control local perturbations in the rates. Note also that the threshold required to produce the low self-reproducing rates is very high. This is due to the fact that we have only excitation. Including inhibition, the same self-reproducing rates are obtained with a lower, more plausible, threshold. See e.g. Section 4.

Global stability

The requirement of global stability is more stringent, since upon a global uniform variation of the rates the variation of the excitatory input to a cell is larger than in the local case. As the purely excitatory network is unstable to local fluctuations, it will be so a fortiori for the global ones. It would therefore make sense to return to the question of global stability only after the conditions for local stability will have been clarified.

The simplified argument presented above serves to define the problem and to give an indication of the difficulties involved in finding a solution for the existence and the stability of self-reproducing spontaneous spike rate distributions. As far as networks composed exclusively of excitatory neurons are concerned, a more careful treatment does not modify the conclusions. This is shown in Section 3.2. To have a stable attractor with low rates, it is necessary to introduce inhibitory neurons. This is done in Section 4.

It was assumed above and will be assumed below that at low rates the process of spike emission of each neuron is a Poisson process. Effects of the absolute refractory period, or of the resetting of the membrane depolarization following the emission of a spike are negligible, if the interspike interval is long compared with the integration time. Hence we shall consider only solutions with rates much smaller than 1. In such situations the variability of the interspike intervals is naturally unity, avoiding the complications raised in refs. [32, 33].

3.2 Leaky integrate-and-fire neurons

Turning to a somewhat more realistic description of the neuron we take Eq. (3) for $V(t)$ as a description of the neuron's depolarization. In a neuron whose active membrane is charged by a current of Gaussian distribution, uncorrelated at different times, as considered in Section 2.2, the mean interspike interval, $\langle T \rangle$, is the mean time of first passage across the threshold. See e.g. refs. [26, 25, 20]. It is given, as a function of the mean μ and standard deviation σ of the input; of the neural threshold θ; and of the post-spike hyper-polarization H, as

$$\langle T \rangle = \tau_0 + \int_{\frac{H-\mu}{\sigma}}^{\frac{\theta-\mu}{\sigma}} \psi(u) du \tag{16}$$

where the function ψ is defined in Eq. (5). The mean spike rate is its inverse, i.e

$$\nu_{out} = \frac{1}{\langle T \rangle} = \left(\tau_0 + \int_{\frac{H-\mu(\nu)}{\sigma(\nu)}}^{\frac{\theta-\mu(\nu)}{\sigma(\nu)}} du \psi(u) \right)^{-1} \equiv \phi(\nu). \tag{17}$$

If the activation rate of a synapse is $\nu = \nu_{in}$, the output rate ν_{out} is obtained by substituting ν_{in} for ν on the right hand side of Eq. 17. ϕ depends on ν_{in} via μ and σ of the synaptic input. For a uniform activation rate ν and synaptic efficacies drawn from the same distribution $P(J)$ with mean 1 and standard deviation Δ

$$\mu(\nu) = C\nu \quad \text{and} \quad \sigma(\nu) = \sqrt{C(1+\Delta^2)\nu}.$$

If x is the fraction of synapses from local collaterals and the local rate is ν while the rate of neurons outside the column is ν_{ext}, the mean and standard deviation of the synaptic input are [26]

$$\mu = C[x\nu + (1-x)\nu_{ext}] \quad \text{and} \quad \sigma = \sqrt{C(1+\Delta^2)[x\nu + (1-x)\nu_{ext}]}$$

A uniform distribution of local rates will be reproduced if

$$\nu_\star = \phi(\nu_\star), \tag{18}$$

and $\nu_{star} = \nu_{ext}$, as explained in Section 3.1 above.

Eq. (18) has a solution with $\nu_\star$ of several spikes per second when the threshold for spike emission is about two standard deviations above the mean of the input, i.e.

$$\theta \sim \mu + 2\sigma, \tag{19}$$

much like in Section 3.1.

For the parameters chosen in Section 2.2, the neural threshold would be at about $440J$ above the resting membrane potential while the stationary synaptic input keeps the neuron at a depolarization of $400J$. About 440 synaptic events should arrive in an interval of the integration time constant of the neuron if the cell is to fire. The average single synaptic contribution is very small compared to the threshold.

If θ=20mV above the resting potential [27], then the average synaptic contribution is $J \sim 0.05$mV. Not an implausible value for the mean EPSP contribution[13, 14, 15, 16, 17]. As in Section 3.1, at such low rates spike emission will be a Poisson process to a good approximation.

Next we consider the stability of this solution. A small perturbation in the rates $\nu \rightarrow \nu_\star + \delta\nu$ in the *local* module at ν_{ext} and θ fixed, will provoke a corresponding perturbation in the mean and variance of the synaptic input to any neuron in the local network. The perturbations in the mean and standard deviation of the potentials $\delta\mu_V$ and $\delta\sigma_V$ can then be obtained from Eq. (3). Details of the stability analysis are given in Appendix A. It is found that stability is obtained when both eigenvalues of the stability matrix, Eq. (47) are negative. In this case any perturbation in the rates will be suppressed and the self-reproducing solution of Eq. (18) will be stable.

To obtain numerical estimates we take $\Delta = 1$, $H = 0$, $C = 20,000$, $x = 0.5$. The fixed external spike rate, ν_{ext}, is taken to be equal to the self-reproducing rate inside the local module, $\nu_\star$, which in turn is the solution of Eq. (18). What we actually do is to prescribe the spontaneous rate and use Eq. (18) to determine the threshold producing it. We find that for $\nu_\star$ in the range from 1 to 5 $s^{-1}$the solution of Eq. (18) is unstable. Thus, as the simple argument of section 3.1 suggested, there is no attractor with realistic spontaneous rates in a network of excitatory units only. The network lacks regulatory units and neurons are either all active at very high rates or all essentially silent. We will see in the next section how the inclusion of inhibitory cells changes the picture and allows for stability at realistic spontaneous rates.

4 The role of unstructured local inhibition

In the cortex about 20% of the neurons are inhibitory interneurons with axons that branch only locally. Both excitatory and inhibitory neurons will be assumed to have similar physiological characteristics, differing only by the values of numerical parameters. It will be further assumed that both types of neurons receive only excitatory inputs from outside the local network. There are two main effects of the inhibitory inputs on the Gaussian process arriving at the soma of a neuron. First, it reduces the mean depolarizing current. Second, it increases the standard deviation of the net depolarizing current. These effects will make it possible to relax the stability condition.

4.1 Model anatomy and physiology

Both types of neurons are leaky integrate-and-fire neurons, as described in Section 2.2. Each type is characterized by a threshold θ_α, a post-spike hyperpolarization H_α, an integration time constant τ_α, with $\alpha = E, I$ indicating whether the neuron is excitatory or inhibitory, respectively. Each neuron receives a Gaussian white noise current of mean μ_α and standard deviation σ_α. Each neuron receives C_α synaptic contacts, which are divided in $C_{\alpha E}$ excitatory synapses and $C_{\alpha I}$ inhibitory ones. The synapses in the network are now of four types, depending on all the possible types of pre- and post-synaptic neurons. For each synaptic type the efficacies are drawn randomly from the distribution $P_{\alpha\beta}(J)$ (α and β denote the post- and pre-synaptic type of neuron, respectively). It has mean $J_{\alpha\beta}$ and SD $J_{\alpha\beta}\Delta_{\alpha\beta}$. A fraction x_α of the excitatory connections come from outside the network and are activated at a rate ν_{ext}. The local excitatory synapses are activated at a rate ν_E, while the inhibitory ones are activated at a rate ν_I.

The number of parameters is large. To reduce this number we make the following choices, which we believe to be in a plausible region:

- both types of neurons receive an equal number of inputs $C = 22,000$, divided into $C_{\alpha E} = 20,000$ excitatory and $C_{\alpha I} = 2,000$ inhibitory ones. The latter are assumed to originate all in the local module;
- both types of neurons receive an equal fraction of excitatory inputs from inside the local module. We choose $x = 0.5$, i.e. 10,000 local excitatory inputs;
- the relative standard deviations of the synaptic efficiacies are equal for all four synaptic types, i.e. $\Delta_{\alpha\beta} = \Delta = 1$ for all α and β (see e.g. ref. [14]).
- both excitatory and inhibitory post-spike hyperpolarizations are set to zero, i.e. $H_\alpha = 0$ for all α.

The set of parameters is further reduced by taking the mean excitatory synaptic efficacy as the unit of the potentials in each population of neurons, $J_{EE} = J_{IE} = 1$. The time unit is the excitatory integration constant. The remaining parameters are

- The average strength of the inhibitory synapses, J_{EI} and J_{II};
- The integration time constant of the inhibitory neuron τ_I;
- The rates ν_E and ν_I.

The neural thresholds θ_E and θ_I are determined by our requirement that preselected values for the rates ν_E and ν_I be self-reproducing given all the other parameters, and that the resulting $\nu_E = \nu_{ext}$. A moment's reflection on this point may make the reading of the rest of this section easier, especially as one approaches Fig. 3.

The input to each type of neuron is again Gaussian. If the depolarization is taken again to be fully linear in the synaptic events, then the mean and variance of the input to an excitatory neuron are:

$$\mu_E = C_{EE}[x\nu_E + (1-x)\nu_{ext}] - C_{EI}J_{EI}\nu_I \tag{20}$$

and

$$\sigma_E^2 = \lambda\left(C_{EE}[x\nu_E + (1-x)\nu_{ext}] + C_{EI}J_{EI}^2\nu_I\right), \tag{21}$$

where $\lambda = 1 + \Delta^2$. Though the above expression are largely self-explanatory, we discuss them in some detail in Appendix B. For inhibitory units we have

$$\mu_I = C_{IE}\tau_I[x\nu_E + (1-x)\nu_{ext}] - C_{II}J_{II}\tau_I\nu_I \tag{22}$$

and

$$\sigma_I^2 = \lambda\tau_I\left(C_{IE}[x\nu_E + (1-x)\nu_{ext}] + C_{II}J_{II}^2\nu_I\right). \tag{23}$$

We may now proceed to study the stable rates of the coupled system of excitatory and inhibitory units, first in the framework of the Gaussian argument of Section 3.1, and then in a network of leaky integrate-and-fire neurons.

4.2 Qualitative considerations

To simplify the qualitative argument we set in Eqs. (20–23) τ_I=1 (i.e. equal inhibitory and excitatory time constants); $J_{EI} = J_{II} = g$ (equal average inhibitory coupling to both types of neurons); $\theta_I = \theta_E$ (equal thresholds); $\Delta = 0$ (uniform synapses). In this case both types of neurons have equal rates ν and receive a synaptic input of the same mean and variance, i.e.

$$\mu = \mu_{ext} + (xC_{EE} - gC_{EI})\nu, \quad \sigma^2 = \sigma_{ext}^2 + (xC_{EE} + g^2C_{EI})\nu$$

where μ_{ext} and σ_{ext}^2 are the mean and variance of the external synaptic inputs.

As in Section 3.1 we approximate the probability of emission of a spike in the integration interval τ (=1), by

$$\Pr(\nu) = \int_{\alpha(\nu)}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right), \tag{24}$$

where α, as a function of θ, μ and σ, is given by Eq. (7). When $\Pr(\nu)$ is low, Eq. (24) is also the output spike rate per τ. Since with the present choice of parameters τ, ν, μ and σ are equal for both types of neurons, their self reproducing rates $\nu_\star$ are the solution of the same equation, i.e.

$$\nu_\star = \Pr(\nu_\star) \tag{25}$$

where $\nu_\star = \nu_{ext}$, as above, and the condition for the stability of a solution is again

$$\frac{d\Pr}{d\nu}(\nu_\star) < 1. \tag{26}$$

Local stability analysis

The stability analysis follows the discussion of Section 3.1. We choose again ν_{ext} in the range 0.01-0.05, and the self-reproducing rate $\nu_\star = \nu_{ext}$. First the self-reproducing rates are arrived at by varying the thresholds $\theta_{E,I}$ to enforce condition (25). Then the variation of Pr with the local rate ν is computed at fixed ν_{ext} and θ. Eq. (26) reduces to the condition

$$\frac{d\,\mathrm{Pr}}{d\nu}(\nu_\star) \sim \alpha\frac{\mu_l}{\sigma} + \alpha^2\frac{\sigma_l^2}{2\sigma^2} < 1, \tag{27}$$

where μ_l and σ_l are, respectively, the mean and the variance of the synaptic input coming from the local collaterals, i.e.

$$\mu_l = (xC_{EE} - gC_{EI})\nu \quad \text{and} \quad \sigma_l^2 = \lambda(xC_{EE} + g^2C_{EI})\nu,$$
$$\sigma^2 = \sigma_l^2 + \sigma_{ext}^2$$

The first term on the right hand side of Eq. (27) is the relative perturbation of the output rate due to the perturbation of the mean of the synaptic input, while the second term corresponds to the relative perturbation due to the variation of the variance of the synaptic input. When $xC_{EE} \gg gC_{EI}$, i.e. there is a predominance of excitatory over inhibitory inputs, the first term dominates. Since there is a large number of excitatory incoming spikes, $\mu_l \gg \sigma$ and condition (26) is violated. On the other hand, if excitatory and inhibitory inputs roughly balance, i.e. $xC_{EE} \sim gC_{EI}$ (With our parameters: $xC_{EE} = 10,000$ and $C_{EI} = 2,000$, implying $g \sim 5$, i.e. inhibitory synapses about 5 times more efficient than excitatory ones), the standard deviation of the synaptic input becomes of the same order or larger than the mean local input μ_l, which makes it possible for the stability condition to be enforced.

The critical value of g at which the self-consistent solution of Eq. (24) becomes stable is actually slightly larger than 5 in a wide range of rates. It varies from $g \sim 5.5$ for $\nu = 0.01$ (1 spike per second) to $g \sim 5.1$ for $5\mathrm{s}^{-1}$. This means that the inhibitory inputs have to be slightly larger than the excitatory ones for the self-consistent rates to be stable. The reason is that when $g = 5$, i.e. excitation exactly balances inhibition, the variation in the SD of the inputs acts to destabilize the solution of Eq. (24). Note that the total synaptic input remains depolarizing even for $g > 5$, due to the excitatory input from outside the local module.

To visualize the effect of inhibition on the stability of the solution of Eq. (25) we present in Fig. 3 the graphical solution for two values of the rate, (ν=0.023, 0.045), each for three values of the relative inhibitory synaptic efficacy (g=3, 5, 5.5). To obtain the same self-reproducing rate for different values of g the threshold is adjusted. Thus in the figure, there are two rates at each of which the straight line, of the left hand side of Eq. (25), (bold-dotted) intersects three curves, the right hand side of (25), marked a (g=3), b (5.0) and c (5.5). As g increases the slope of $\mathrm{Pr}(\nu)$ at the intersection decreases. At $g = 3$ it is essentially as for $g = 0$ (Fig. 2). At $g = 5$ both slopes are still larger than one, because of the variation of the variance of the local synaptic input. At $g = 5.5$ both self-reproducing rates are stable.

For low g the low-rate stable self-reproducing solution to Eq. (25) is essentially at $\nu = 0$. It is kept away from zero by the fixed excitatory input from outside the module.

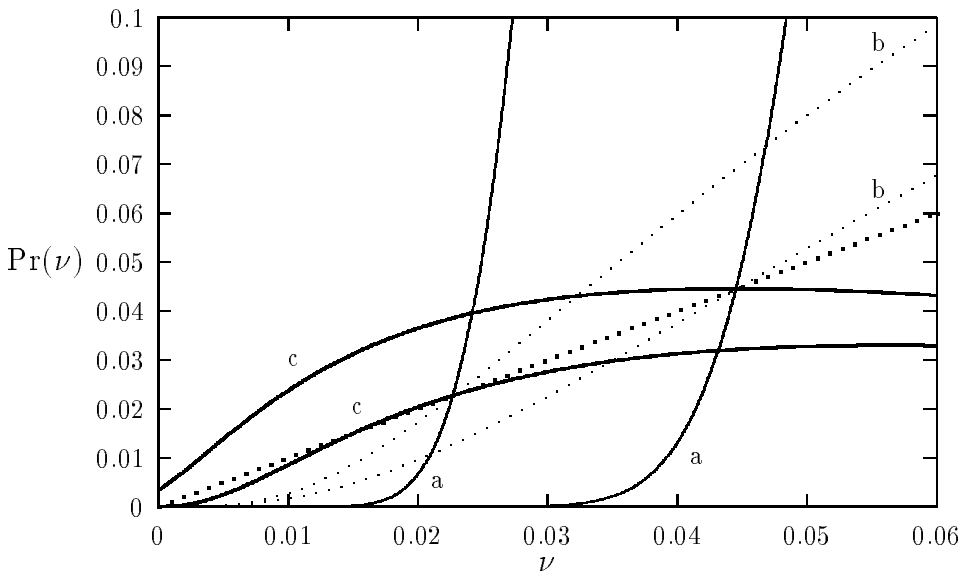

Figure 3: Graphical representation of the stability condition for two self-reproducing spontaneous spike rates $\nu_\star$=0.023, 0.045. Each $\nu_\star$, the intersection of the (bold-dotted) straight line, is obtained for 3 values of g by adjusting the threshold. (a) g=3 (b) g=5 (c) g=5.5. Only for g=5.5 (bold) are the solutions stable.

Note that the external afferents alone leave the neurons at extremely low rates. When g crosses its critical value the state in which all neurons have a spontaneous rate of several spikes per second, depending on the neural threshold, becomes stable and small variations around this rate are suppressed.

There is an intermediate regime in which the network has two stable self-reproducing rates, the quiescent state with ν essentially at 0 and a state with spontaneous rates of several spikes per second. The lower curve c in Fig. 3 represents such a situation. It corresponds to $\nu_\star$=2.3s^{-1}and g=5.5. Large enough fluctuations may shift the network from one state to the other. When the inhibition increases further, the quiescent state becomes unstable.

What is actually going on? To preserve the same self-reproducing spontaneous rate with increased inhibition the threshold must be lowered or, equivalently, the mean unitary EPSP must increase. In the present analysis it is assumed that the fixed excitatory rate from outside the module is equal to the internal self-reproducing rate $\nu_\star$ solving Eq. (25). This has the consequence that the external currents are higher relative to the threshold, destabilizing the quiescent state. In this case the state with spontaneous rates of several spikes per second remains the only stable state of the network. The quiescent state becomes unstable when the external inputs are such that the curve $\nu \to \Pr(\nu)$ has a unique solution $\nu_\star$. This is the case for the upper curve c of Fig. 3. In fact, both curves c in Fig. 3 correspond to the same value of the gain, $g = 5.5$. The reason why in one case the quiescent state is stable (lower curve c) while in the other it is not (uper curve c) is related to the relative magnitude of the thresholds which is lower in the upper curve, to produce the higher rate. This has the effect of increasing the external afferent current,

which in turn destabilizes the quiescent state as explained above. The self-reproducing rate $\nu_\star$=0.023 is stable for $g > 5.45$, but the quiescent state is stable for $g < 5.7$. Thus both states are stable between these two values of g. On the other hand, the self-reproducing rate $\nu_\star = 0.045$ becomes stable for $g > 5.18$, and coexists with the quiescent state only until $g = 5.35$.

4.3 Integrate-and-fire neurons: stable spontaneous rates

We turn now to the corresponding network of integrate-and-fire neurons, in which we can investigate the influence of different characteristics of excitatory and inhibitory neurons on the stability of the spontaneous state. It turns out that as long as the time constants of both types of neurons are equal, the simple considerations of section 4.2 capture the qualitative features. However, we shall see that when τ_I is allowed to become much smaller than τ_E, the synaptic constraints for a stable solution are significantly relaxed. Then we investigate the stability of the spontaneous activity to a variation of the external afferents. We show, for a particular set of parameters, that the spontaneous activity state is stable for a wide range of values of ν_{ext}.

In a network of integrate-and-fire neurons the self-reproducing rates for both populations are given by the solutions of the coupled equations

$$\nu_E = \phi_E(\mu_E, \sigma_E) \tag{28}$$

$$\nu_I = \phi_I(\mu_I, \sigma_I) \tag{29}$$

in which the means μ_α and SDs σ_α are functions of the rates of the excitatory and the inhibitory neurons, ν_E and ν_I, via Eqs. (20–23). The right hand sides are, as in Section 3.2,

$$\phi_\alpha(\mu_\alpha, \sigma_\alpha) = \left(\tau_0 + \tau_\alpha \int_{\frac{H_\alpha - \mu_\alpha}{\sigma_\alpha}}^{\frac{\theta_\alpha - \mu_\alpha}{\sigma_\alpha}} du \psi(u) \right)^{-1} \tag{30}$$

with ψ given in Eq. (5).

The stability analysis consists of considering the effect of a general *uniform* displacement of the two rates: $(\nu_E, \nu_I) \to (\nu_E + \delta\nu_E, \nu_I + \delta\nu_I)$ in the local module. The elements of the analysis have been discussed in Section 3. The details are given in Appendix B. In the following we discuss a few simple examples in the space of parameters for which the self-reproducing solution of Eqs. (28,29) is stable against small perturbations.

4.3.1 Equal time constants and inhibitory couplings

We take the integration time constant equal for excitatory and inhibitory neurons, τ_I=τ_E=1, and equal inhibitory efficacies to both types of neurons, $J_{EI} = J_{II} = g$. Solving the equations numerically, we find that stability of the spontaneous rates can be achieved if excitatory and inhibitory *local* inputs are approximately balanced. In fact, the mean local synaptic input should be slightly dominated by the inhibition, much as in Section 4.2. Hence, the smaller number of inhibitory inputs (20%) must be compensated either

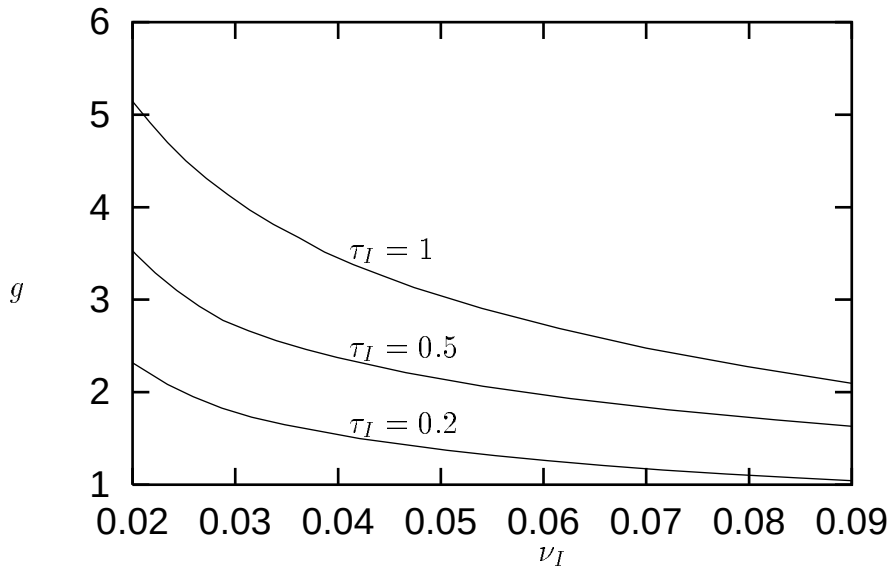

Figure 4: Stability region of excitatory spontaneous activity of $2s^{-1}$and J_{EI}=J_{II}=g, in the space of relative average inhibitory efficacy g and inhibitory neurons' rates ν_I, for several values of the inhibitory integration time (indicated on the figure). The stable region for each τ_I is above the corresponding curve. As τ_I decreases the stability region expands.

by stronger inhibitory synapses ($J_{EI} > J_{EE}$) or by relatively higher inhibitory rates. It is not a very stringent condition. In hippocampus inhibitory couplings are indeed more efficient than the excitatory ones [35], and there is evidence that inhibitory rates are actually higher[36, 37].

In Fig. 4 we show regions of stability for several values of τ_I in the g-ν_I plane, for ν_E= $2s^{-1}$and fixed external rates $\nu_{ext} = \nu_E$. Consider first the curve corresponding to $\tau_I = 1$. The spontaneous activity is stable above the curve, and is unstable below it. Higher inhibitory rates allow lower inhibitory efficacies and vice versa. We see that the critical inhibitory coupling g goes down from 5.15 for $\nu_I = \nu_E$, to about 2 when the inhibitory spontaneous rate is much higher ($\nu_I = 9$ spikes per second).

In Fig. 5 we show the region of stability of the spontaneous activity upon varying the thresholds θ_E and θ_I, for relative inhibitory efficacies: g= 2, 3 and 5. For each value of g, the spontaneous activity state is stable **below** the corresponding curve. Taking 20mV for both excitatory and inhibitory thresholds, θ_E=300 corresponds to an average efficacy J_{EE}=0.067mV, while θ_I=50 corresponds to J_{IE}=0.4mV. These values are well within the experimentally reported range[13, 14, 15, 16, 17]. Fig. 4 indicates that for the corresponding values of g, the inhibitory threshold has to be smaller than the excitatory one (or conversely the average excitatory synapses have to be stronger on inhibitory cells), which in turn means that inhibitory cells have a higher rate. A decrease in g requires a compensating decrease of θ_I/θ_E, which implies higher inhibitory spike rates.

For example, in a network with $g = 3$ (inhibitory synapses on both excitatory and inhibitory neurons are three times more efficient than an excitatory synapse on the same

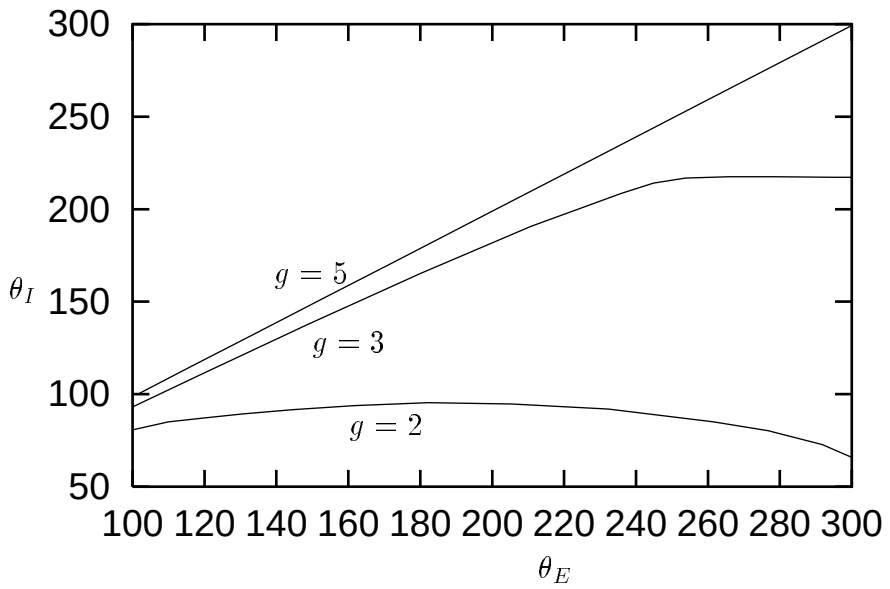

Figure 5: Stability region of spontaneous activity, in the space of thresholds θ_E-θ_I for three values of the relative inhibitory efficacy g (indicated in the figure). For each value of g the low rate state is stable below the corresponding curve.

neuron), $\theta_E \sim 180$, $\theta_I \sim 160$ there is a spontaneous activity attractor with ν_E=2s^{-1}and ν_I=5.4s^{-1}. In this case the mean afferent on any excitatory neuron in its integration time constant is $\mu \sim 80$ and its variance is $\sigma \sim 50$. The mean synaptic input is still largely positive because of the external excitatory afferents. Taking realistic values for the thresholds: $\theta_E = \theta_I = 20$mV above the resting potential, we have that the mean excitatory synaptic contribution is of order 0.1mV while an inhibitory spike hyperpolarizes, on average, by some 0.3mV.

4.3.2 Fast inhibition

Next we explore the effects of faster inhibitory neurons, $\tau_I < 1$. If inhibition is faster, any perturbation in the rates in the excitatory network is more easily controlled, and the network should be more stable. We find that decreasing the integration time constant of the inhibition relaxes the condition on the relative inhibitory couplings and on the inhibitory spike rates. One cannot, however, decrease this time constant arbitrarily since the number of spikes received by an inhibitory unit in an interval equal to τ_I needs to be large enough for the Gaussian approximation to the synaptic input to be valid. In other words, we need

$$\min(C_{IE}\nu_E\tau_I, C_{II}\nu_I\tau_I) \gg 1$$

i.e. the afferent number of spikes of either type to an inhibitory neuron per integration time must be large. For example if ν_I=ν_E=5s^{-1}the requirement is $\tau_I \gg 0.01$, in which a single spike will arrive. A more severe constraint comes from the fact that we have considered the

time constants for the conductance changes to be smaller than the integration constants for the potentials. This means that we cannot reasonably take values of less than 2ms for the inhibitory time constant.

Fig. 4 presents the regions of stability, in the space of g and ν_I, for a spontaneous activity distribution with excitatory rates ν_E=2s^{-1}and equal inhibitory efficacy on both excitatory and inhibitory neurons (J_{EI}=J_{II}=g), for three different values of τ_I=1, 0.5 and 0.2 (10, 5 and 2 ms), indicated in the figure. The activity distribution is stable above the corresponding curve. Decreasing the integration time constant has the effect of increasing the stability of the rate distribution. Therefore, at parity of excitatory rates the spontaneous attractor will be stable for lower values of the average inhibitory efficacy or inhibitory rate. For example, if ν_I=ν_E=2s^{-1}, the critical inhibitory strength g is 3.53 for τ_I=0.5 and 2.24 for τ_I=0.2. In particular, the spontaneous rates will be stabilized even with large dominance of excitatory inputs over the inhibitory ones. This is due to the fact that the inhibition adjusts its rates faster in case of a perturbation. In the limit $\tau_I \to 0$ any positive value of the inhibitory couplings and of the inhibitory rates would stabilize the spontaneous rates, since inhibition would adjust instantaneously to any perturbation in the excitatory network.

We have also considered the case $J_{EI} \neq J_{II}$: different inhibitory efficacy on excitatory and inhibitory neurons. For example, we take an inhibitory neuron with τ_I=0.2 (2ms). In a network with $J_{EI} = 1.25$, $J_{II} = 1$, $\theta_E \sim 560$, $\theta_I \sim 140$, we have a spontaneous attractor with ν_E of 3s^{-1}and ν_I of 4.2s^{-1}. The mean of the currents afferent to an excitatory neuron is very close to the threshold, $\mu_E \sim 480$, with $\sigma_E \sim 40$. Taking the thresholds to be 20mV above the resting potential, we have $J_{EE} \sim 0.03$mV, $J_{EI} \sim 0.05$mV, $J_{IE} = J_{II} \sim 0.1$mV.

Note that very different values of the excitatory threshold have appeared along the way. This is because the value of the threshold is very sensitive to the ratio of inhibitory to excitatory unitary PSP. Without inhibition, (Section 3.1) the thresholds were 500–1000 synaptic inputs. As long as the inhibitory synaptic inputs remain weak compared to the excitatory ones, as in the case described above in which stability relied on fast rather than strong inhibition, the thresholds remain of the order of 500. When inhibitory synaptic inputs get stronger, as in Section 4.3.1, the thresholds decrease and become of the order of 200 synaptic inputs, which in turn yields average EPSPs of 0.1mV, apparently more acceptable. These are the thresholds obtained for average excitatory spontaneous rates of order 2–3s^{-1}. If the average rates vary significantly from those values, so will the thresholds, decreasing with the level of the spontaneous activity. Another point to keep in mind is that the analysis has been carried out for neurons that sum linearly their synaptic inputs. The values of the thresholds might be significantly modified in the case of nonlinearities. See Appendix D for a discussion of a simple non-linear case. But, even within this simplified context one easily finds oneself in biologically plausible territory.

In the above discussion we have neglected the potential influence of transmission delays. If delays are small compared to the integration time constants, one may expect such delays to have little effect on the stability of the attractors. This is no longer the case if delays become of order of one of the integration time constants. To check the effect of delays we have simulated the dynamics of the mean and variance of the depolarization of each neuron, according to Eqs. (48,49), in a network of 1,000 excitatory and 200 inhibitory

units. Random delays were introduced at each synaptic site whose mean ranges from 1–2ms. The instantaneous rates were in turn determined at each integration step from the transduction function of Eq. (4). (The simulation is somewhat more general than the current-rate dynamics of ref. [20] in that we also take into account the dynamics of the variance of the depolarization of each neuron, while in ref. [20] only the dynamics of the contribution from the selective activity in the network was considered.)

The inhibitory time constant τ_I was taken to be 5 and 2ms. We find that when τ_I=5ms, the network is robust to the delays, even if the parameters are close to the onset of instability. On the other hand, if inhibition is faster (τ_I=2ms) and correspondingly weaker, the network is sensitive to the delays. The spontaneous attractor is destabilized if the parameters of the network are close to the instability line in Fig. 4. The rates of both excitatory and inhibitory populations begin to oscillate until all activity in the network dies down. With the parameters used for the simulation the all-silent network is stable as in the case of lower curve c in Fig. 3. The introduction of delays causes the network to switch from the state with mean spontaneous activity of a few spikes per second into the silent state.

Qualitatively the picture is as follows. When τ_I decreases, the region of parameters in which spontaneous self-reproducing rates are stable *in absence of delays* grows. However, in this region the spontaneous self-reproducing attractor coexists with the quiescent attractor in which all neurons are silent. Delays have the effect of amplifying small fluctuations, which drive the network into the more stable, quiescent state. However, if inhibitory synaptic efficacies and/or inhibitory firing rates are sufficiently raised, stability is restored even in presence of delays. These simulations will be described in detail in a separate report.

To conclude, fast inhibition stabilizes spontaneous self-reproducing rates with weaker relative inhibitory synaptic efficacies. But, this is done at the price of a sensitivity to random delays. A rather strong inhibition seems therefore to be required to obtain robust spontaneous activity distribution.

4.3.3 Stability to a variation in the external afferents

We have shown that in a biologically plausible region of parameters the state of spontaneous activity is stable to fluctuations in the local rates, at fixed external afferents ν_{ext}. We turn next to the question of the robustness of this stability to a variation of the external currents. We start from the stable state in which $\nu_E = \nu_{ext}$, and fix all parameters. Then the external rate ν_{ext} is varied. The change in the external rate has the effect of adjusting the equilibrium rates in the network (ν_E, ν_I).

Fig. 6 presents the variation of the excitatory and inhibitory stable spontaneous rates upon variations of the external rate ν_{ext}, for fixed parameters. The low rate solution is stable in the entire range ν_{ext} between 0 and 5s^{-1}. The inhibition adjusts itself to the variation of the external rates, leaving the network in a spontaneous activity state of almost constant excitatory rates. When the external rate is zero, the network has another stable solution, the quiescent state $\nu_E = \nu_I = 0$. However a major fluctuation would be needed to reach this state.

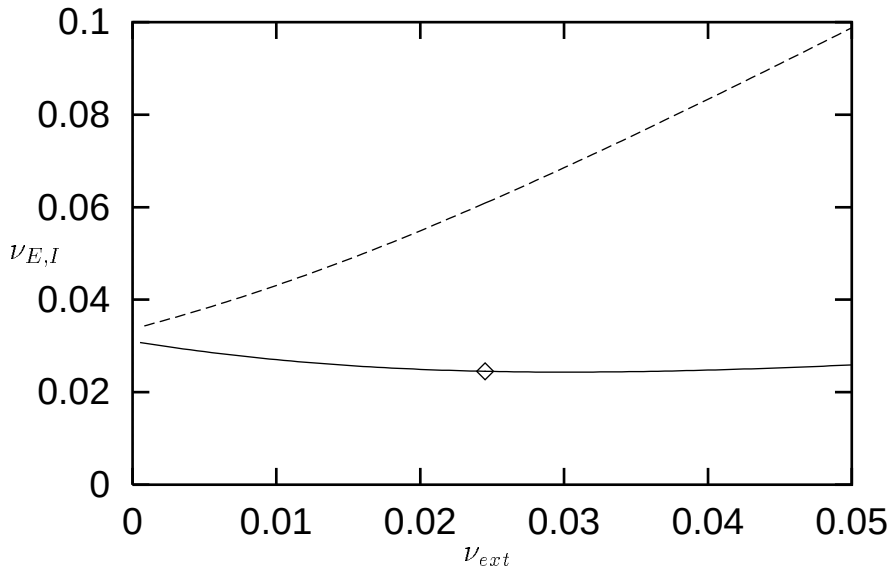

Figure 6: Robustness of spontaneous activity against variation of external afferents. Stable rates ν_E (full curve) and ν_I (dashed curve) as a function of the external rate ν_{ext}. $\diamond$: $\nu_E = \nu_{ext}$. Fixed network parameters $\theta_E = 136$, $\theta_I = 85$, $J_{EI} = J_{II} = 4$, $\tau_I = 0.4$.

Global stability.

Global stability can be considered in two different ways. One may consider a sudden fluctuation in the spontaneous rates of all cells in the cortex. To have stability against such an eventuality significantly larger inhibition would be required. If one allows for such global fluctuations then the system may also tend to settle in the global quiescent state which is stable.

But that is a very unlikely situation. The more relevant eventuality would be the spreading of a local fluctuation in rates. However, as we have shown here our system is stable against such spreading. If we consider a particular module, and the rates fluctuate in another with excitatory projections into the first, the local inhibition in the first module is sufficient to suppress the variation in the excitatory input arriving from outside.

5 Learning and selective activity

So far we have considered a completely unstructured network, and the *distribution* of afferent synaptic efficacies on every neuron was assumed identical. The conditions on the parameters leading to stable spontaneous activity are not very stringent and the region of parameters includes biologically plausible intervals. The next question is whether a network characterized by the same set of parameters defining neuronal physiology and anatomy, which provided a stable spontaneous activity distribution (even at relatively high rates), can provide stable local delay activity distributions by the mere modification of synapses between excitatory neurons.

We imagine the local, structured distribution of synaptic efficacies to have been generated by stimuli arriving at the module and being learned. Here we will not expand on the issue of the learning dynamics of attractors (see e.g. [38, 23]). Instead, each stimulus either activates a given cell in the local module significantly above the spontaneous level or leaves it at spontaneous rates (i.e. binary stimuli). The excitatory synapses in the local network have Hebbian plasticity: a pair of neurons simultaneously activated by a given stimulus will enhance the efficacy of the connecting synapse while a pair of neurons with anti-correlated activities reduce the efficacy of the synapse connecting them. Inhibitory synapses remain unstructured. This is admittedly a very schematic representation of learning, but it suffices for the discussion of the issue raised here which concerns the possible stability of local, selective activity attractors on the background of the global spontaneous activity 'vacuum'.

In order to obtain an intuition into the behavior of a network in which stimuli choose their selective neurons randomly and independently — which we discuss briefly in Section 6 below — we revert to a simpler, approximate situation in which stimuli are subject to the constraint that a neuron is excited by at most one stimulus. This case is amenable to detailed analysis (see also e.g. ref. [39]) and provides a reasonable approximation for the unconstrained case, provided the coding level is low and the number of learned stimuli is not too high.

Suppose that p stimuli had been learned. Each stimulus activates an equal fraction f (coding level) of the excitatory neurons in the local network. With non-overlapping stimuli, $fp < 1$. We take f=0.01, and p=50. For the same values of p and f, if stimuli were composed of neurons selected randomly and independently, about 9% of the neurons in the module would be activated by more than one stimulus. We have chosen f=0.01 for the coding level as it is close to the value deduced from experiments in infero-temporal cortex of monkeys[39]. See also below.

The analysis of the case of non-overlapping stimuli is simplified by the fact that the excitatory network can be divided into two sub-populations: one of pfN neurons excited by one of the p learned stimuli (selective neurons); and the remaining $(1-pf)N$ non-selective neurons, i.e. those not activated by any stimulus (See e.g. [39]).

The excitatory synaptic input to each population is:

Selective neurons. With the parameters of Section 4, each of these neurons receives about fxC_{EE}=100 synaptic contacts from other neurons in the sub-population of neurons activated by the same stimulus. These synaptic efficacies are characterized by a new probability distribution $P_+(J)$ with mean $J_+ > J$ where J is the average, excitatory, unstructured synaptic efficacy sustaining the spontaneous activity (J=1 as in Sections 3 and 4). The enhanced mean of this distribution expresses the potentiation (LTP) due to coactivation of pre and post synaptic neurons by the same stimulus. The structured local synaptic efficacies have a standard deviation $J_+\Delta$. A selective neuron has $(1-f)xC_{EE}$ synaptic contacts with other excitatory neurons which are not excited by the same stimulus. These neurons are either neurons selective for a different stimulus or non-selective neurons. The corresponding synaptic contacts are characterized by a distribution $P_-(J)$ with mean $J_- < J$, and standard deviation $J_-\Delta$, expressing depression (LTD) in these synapses, due to anti-correlation of the activities of pre- and post-synaptic neurons

when the stimulus is present.

Non selective neurons. These neurons receive $pfxC_{EE} = 5000$ contacts from selective neurons. These contacts are drawn from the depressed distribution $P_{-}(J)$. They also receive $(1-pf)xC_{EE} = 5000$ contacts from other non-selective neurons. These synapses have not been affected by learning and therefore are drawn from the original distribution $P(J)$ of Sect. 3.

The distributions $P_{\pm}$ characterize the learning of the set of stimuli in the network. We suppose that the mean excitatory synaptic efficacy in the entire module is not changed by the learning process. So on average synaptic depressions compensate for potentiations.[3] This is expressed as

$$pf^2 J_+ + pf[2 - f(p+1)]J_- + (1-pf)^2 = 1$$

which is a statement that in the local module the mean potentiation plus the mean depression plus the mean unmodified efficacies gives the pre-learning mean. J_- can therefore be expressed as a function of J_+, i.e.

$$J_- = \frac{2 - f(p + J_+)}{2 - f(p+1)} \tag{31}$$

When the coding level f is very low, we have to first order in f

$$J_- = 1 - \frac{J_+ - 1}{2} f + O(f^2)$$

Usually, the performance of an attractor network, in terms of the number of items it can recall, increases as the coding level decreases[41, 42, 38]. However, in a network with a fixed, finite connectivity one cannot lower f too much, since a selective neuron of a given stimulus will eventually get too few inputs from the other selective neurons in the same stimulus. For a recurrent connectivity of 10,000 and $f \sim 0.01$, every neuron receives about 100 synaptic contacts from other selective neurons, which is a number that may suffice to keep activity going. This is an attractive coincidence since, as was mentioned above, both numbers are biologically plausible.

We may now investigate the effect of learning, represented by the distributions $P_{\pm}(J)$, on the stable rate distributions supported by the local module. We consider two situations:

1. No stimulus-related activity in the local network; corresponding to stimulation by an unfamiliar stimulus.

2. Stimulus-related persistent activity in the local network; corresponding to stimulation by a familiar stimulus.

In both cases we consider the situation following the removal of the stimulus, much like in the delay period of delayed match-to-sample experiments [8, 9, 10, 11].

What makes the problem simple and amenable to detailed solution is the combination of the facts that (a) each neuron's behavior is fully determined by the mean and the variance

[3] See in this context e.g. [40, 38].

of the synaptic input to the depolarization at the cell body and (b) that when there is no sharing of selective neurons between different stimuli, both quantities are uniform in each class of neurons and easily computable. Thus we first compute the mean and the variance of the distribution of the synaptic inputs arriving to a neuron of every type in the local network. Then we solve the equations for self-reproducing rate distributions, i.e. the rates in the different populations of neurons in the network. In both cases neurons receive an external, excitatory synaptic input (spontaneous activity from other cortical areas) of mean and variance:

$$\mu_{\alpha,ext} = (1-x)C_{\alpha E}\tau_\alpha \nu_{ext}; \;\; \sigma^2_{\alpha,ext} = \lambda \mu_{\alpha,ext}$$

where $\alpha = E, I$ indicates if the receiving neuron is excitatory or inhibitory, and $\lambda = 1+\Delta^2$ measures the synaptic variability. They also receive an inhibitory input from the local inhibitory neurons whose mean and variance are

$$\mu_{\alpha I} = C_{\alpha I} J_{\alpha I} \tau_\alpha \nu_I; \;\; \sigma^2_{\alpha I} = \lambda J_{\alpha I} \mu_{\alpha I}.$$

Next we consider in detail the synaptic contribution from the local excitatory neurons, which depends on whether or not there is stimulus-related activity in the local network.

1. **No stimulus-related activity in the local network**: All selective neurons have initially the same rates. Selective neurons receive currents with mean μ_+ and variance σ_+. Non selective neurons have afferent currents with mean μ_0 and variance σ_0. Their rates are respectively ν_+ and ν_0. The mean *recurrent excitatory* synaptic input to a selective neuron, is the sum of three contributions: $xC_{EE}fJ_+\nu_+$, from other neurons selective for the same stimulus; $xC_{EE}(p-1)fJ_-\nu_+$, from neurons selective for other stimuli; and $xC_{EE}(1-pf)J_-\nu_0$, from non-selective neurons. After adding the external excitatory afferent, $\mu_{E,ext}$ and the local inhibition, μ_{EI}, the total average synaptic input becomes

$$\mu_+ = xC_{EE}\left[(fJ_+ + (p-1)fJ_-)\nu_+ + (1-pf)J_-\nu_0\right] + \mu_{E,ext} - \mu_{EI} \tag{32}$$

and its variance is

$$\sigma^2_+ = \lambda x C_{EE}\left[(fJ^2_+ + (p-1)fJ^2_-)\nu_+ + (1-pf)J^2_-\nu_0\right] + \sigma^2_{E,ext} + \sigma^2_{EI} \tag{33}$$

A non-selective neuron receives $xC_{EE}[pfJ_-\nu_+]$ recurrent excitation from selective cells; $xC_{EE}(1-pf)\nu_0$ from other non-selective cells. Its total mean input is:

$$\mu_0 = xC_{EE}\left[pfJ_-\nu_+ + (1-pf)\nu_0\right] + \mu_{E,ext} - \mu_{EI} \tag{34}$$

and

$$\sigma^2_0 = \lambda x C_{EE}\left[pfJ^2_-\nu_+ + (1-pf)\nu_0\right] + \sigma^2_{E,ext} + \sigma^2_{EI}. \tag{35}$$

The inhibitory units receive a mean synaptic input of

$$\mu_I = xC_{IE}\tau_I\nu_E + \mu_{I,ext} - \mu_{II} \tag{36}$$

with variance

$$\sigma_I^2 = \lambda \tau_I x C_{IE} \nu_E + \sigma_{I,ext}^2 + \sigma_{II}^2, \tag{37}$$

where ν_E is the average rate of an excitatory neuron in the local network, i.e.

$$\nu_E = pf\nu_+ + (1-pf)\nu_0.$$

The stable rates are given by the system of three coupled equations

$$\nu_+ = \phi_E(\mu_+, \sigma_+) \tag{38}$$

$$\nu_0 = \phi_E(\mu_0, \sigma_0) \tag{39}$$

$$\nu_I = \phi_I(\mu_I, \sigma_I) \tag{40}$$

where $\phi_{E,I}$ are the transduction functions, given by Eq. (30), for the excitatory and the inhibitory neurons, respectively. The means $\mu_{+,0,I}$ and the variances $\sigma_{+,0,I}^2$ depend in turn on the rates $\nu_{+,0,I}$ via Eqs. (32-37). This makes the above equations a closed set of equations for the rates in the different populations.

Equations (38–40) are studied numerically for the three stable, self-reproducing rates $\nu_{+,0,I}$. The parameters are the same as for the last numerical example of Section 4, i.e. $\tau_I = 0.2\tau_E$, $J_{EI} = 1.5 J_{EE}$, $J_{IE} = J_{II}$, $\theta_E \sim 560 J_{EE}$, $\theta_I \sim 140 J_{EI}$. In the absence of learning (J_+=1), there is stable spontaneous activity of ν_E=3s^{-1} and ν_I=4.2s^{-1}.

Next the solutions of these equations are studied as a function of J_+, the enhancement due to learning of the mean synaptic efficacy. The observed stable rates for the three populations of neurons are plotted as light curves in the lower figure in Fig. 7. Note that as J_+ increases, the stable rate of the selective neurons increases, from 3s^{-1}at J_+=1 (no learning) to 5.5s^{-1} at J_+=5 (the average LTP synaptic value is five times the background one). Correspondingly, there is a decrease of the rates of the non selective neurons, which become as low as ν_0=1s^{-1} when J_+=5. The inhibitory neurons have a milder variation of their rates, and stay between 4 and 5 spikes per second.

We conclude that the spontaneous attractor is robust against the synaptic variations induced by learning, though it acquires some structure on the local level. The spontaneous rates carry some indication of the learning history. It is a precursor of the impending selective attractor. Neurons that had been strongly activated by external stimuli during learning will have enhanced rates, while those that have not been activated at all will have their rates depressed relative to the bare spontaneous activity. Note, however that the enhanced *spontaneous* activity is non selective. In other words, neurons partaking in any of the stimuli will have somewhat elevated spontaneous rates simultaneously.

2. Stimulus-related activity in the local network: After the removal of a given stimulus, neurons selective to the particular stimulus have higher rates than neurons selective of other stimuli. In that case there may be four types of rates: of neurons selective of the afferent stimulus, ν_{sel}; of neurons selective for other stimuli, ν_+; of non-selective excitatory neurons, ν_0; of inhibitory neurons, ν_I.

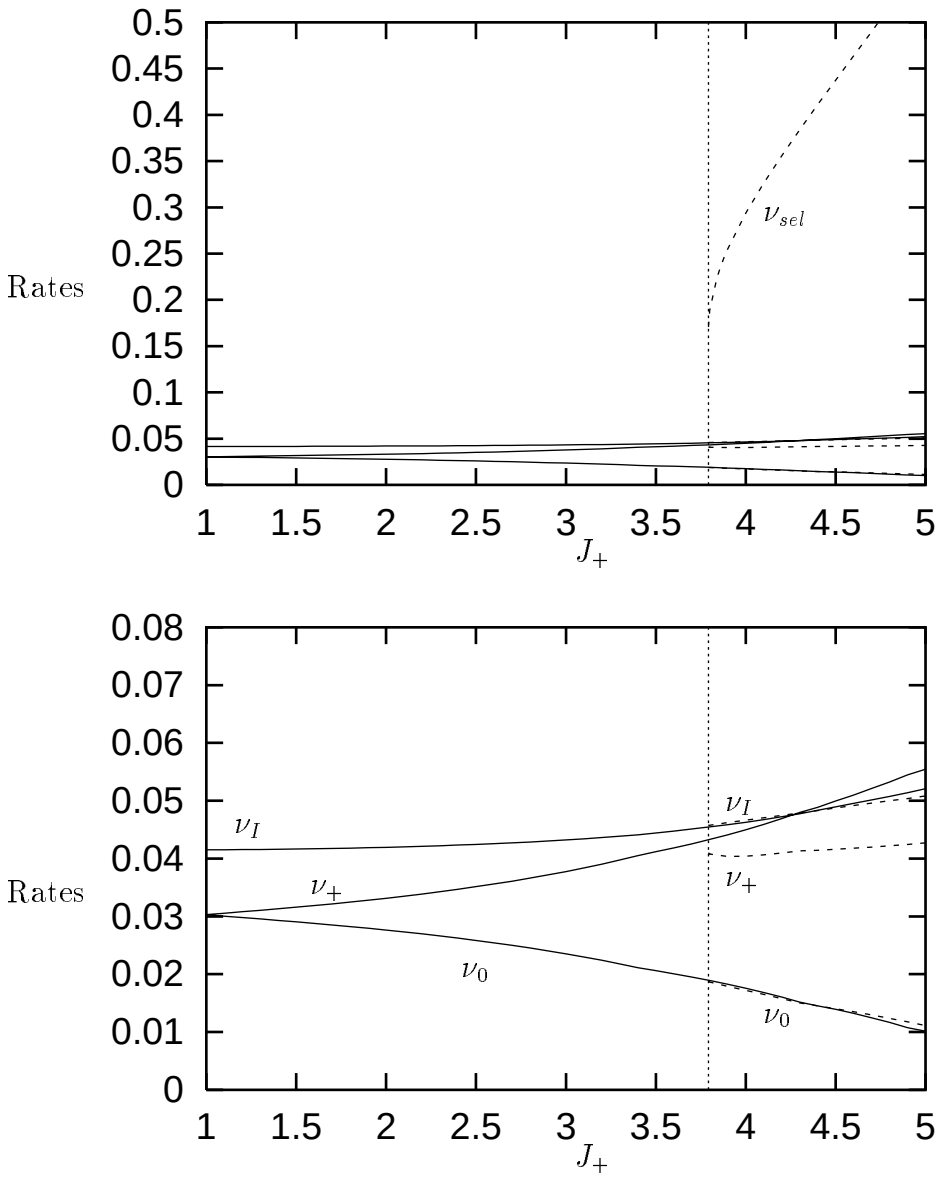

Figure 7: Stable, self-reproducing rates for selective, non-selective and inhibitory neurons vs J_+ the relative mean synaptic potentiation due to learning. To left of dotted vertical line a single solution (full curves) – spontaneous activity. To the right, an abrupt development of a second stable solution (dashed curves). The group of neurons selective for the current stimulus has a higher, selective rate, marked ν_{sel}. The bottom figure is a zooming in of the lower part of the top figure. The spontaneous activity, non-selective rate distribution, continued from the left, coexists with the selective one to the right.

The mean synaptic input to a neuron selective to the present stimulus, in an interval corresponding to the integration time of the membrane potential, is:

$$\mu_{sel} = xC_{EE}\left[fJ_+\nu_{sel} + (p-1)fJ_-\nu_+ + (1-pf)J_-\nu_0\right] + \mu_{E,ext} - \mu_{EI}$$

where the different terms correspond to the synaptic contributions coming from, respectively, the neurons selective to the same stimulus; the other selective neurons; the non-selective excitatory neurons in the module; the input from spontaneous activity external to the module; and the inhibitory input. For the corresponding variance we have

$$\sigma^2_{sel} = \lambda xC_{EE}\left[fJ^2_+\nu_{sel} + (p-1)fJ^2_-\nu_+ + (1-pf)J^2_-\nu_0\right] + \sigma^2_{E,ext} + \sigma^2_{EI}$$

Neurons selective of other stimuli have a mean input:

$$\mu_+ = xC_{EE}\left[fJ_-\nu_{sel} + f[J_+ + (p-2)J_-]\nu_+ + (1-pf)J_-\nu_0\right] + \mu_{E,ext} - \mu_{EI}$$

and a variance

$$\sigma^2_+ = \lambda xC_{EE}\left[fJ^2_-\nu_{sel} + f[J^2_+ + (p-2)J^2_-]\nu_+ + (1-pf)J^2_-\nu_0\right] + \sigma^2_{E,ext} + \sigma^2_{EI}$$

The mean synaptic input to a non-selective neuron is:

$$\mu_0 = xC_{EE}\left[fJ_-[\nu_{sel} + (p-1)\nu_+] + (1-pf)\nu_0\right] + \mu_{E,ext} - \mu_{EI}$$

with variance:

$$\sigma^2_0 = \lambda xC_{EE}\left[fJ^2_-[\nu_{sel} + (p-1)\nu_+] + (1-pf)\nu_0\right] + \sigma^2_{E,ext} + \sigma^2_{EI}.$$

Finally, an inhibitory neuron has its mean input and variance given again by Eqs. (36,37), but with ν_E replaced by

$$\nu_E = f[\nu_{sel} + (p-1)\nu_+] + (1-pf)\nu_0.$$

The stable, self-reproducing rates are determined by the stable solutions of the four coupled equations

$$\nu_{sel} = \phi_E(\mu_{sel}, \sigma_{sel}) \tag{41}$$

$$\nu_+ = \phi_E(\mu_+, \sigma_+) \tag{42}$$

$$\nu_0 = \phi_E(\mu_0, \sigma_0) \tag{43}$$

$$\nu_I = \phi_I(\mu_I, \sigma_I) \tag{44}$$

When J_+ is close to 1, the only solution are the spontaneous rates $(\nu^\star_+, \nu^\star_0, \nu^\star_I)$, which are the solutions of Eqs. (38–40) encountered above, i.e.

$$\nu_{sel} = \nu_+ = \nu_+^\star, \;\; \nu_0 = \nu_0^\star, \;\; \nu_I = \nu_I^\star.$$

When J_+ crosses a critical value ($\sim$3.78), Eqs. (41-44) develop *abruptly* an *additional* solution over and above the spontaneous rate distribution. The full scenario is displayed in Fig. 7, where all the stable frequency solutions are plotted as a function of the relative potentiation due to learning J_+. To the left of the dotted vertical line, the critical value of J_+, there is a single stable solution for each of the three neuron types: selective, non-selective excitatory, inhibitory. There is no selectivity, in the delay activity, to any particular stimulus. The only effect of learning is to enhance gradually the spontaneous activity of *all* selective neurons. These spontaneous rate distributions vary mildly with increasing J_+. The picture is expanded in the bottom figure.

Upon crossing the dotted vertical line there is a bifurcation: a new branch appears. A class of neurons develops an alternative stable rate. Those are the neurons selective to the particular afferent stimulus. It is represented by the upper bold curve marked ν_{sel}. These are stable, self-reproducing selective rates in a delay period. They vary in a rather sensitive way with J_+, from 17s^{-1} at $J_+ \sim 3.78$ to above 50s^{-1} when J_+=5. The neurons in the other classes have low rates. There are three such classes, as explained above: neurons selective of other stimuli; non selective excitatory neurons and inhibitory neurons. Their rates are shown in the figure. To appreciate the fine structure the lower figure should be consulted. It expands the low rate part of the full figure.

There are six curves in this part of the figure. Three (dashed) corresponding to the rates of the three classes in the presence of selective activity, and three (full) corresponding to the continuation of the spontaneous activity curves from the low J_+ region. The main effect is that the rates of neurons selective to other stimuli (ν_+) are significantly reduced. Non-selective and inhibitory neurons are hardly effected by the presence or absence of the selective structured activity. The reason is that selective activity affects only midly the overall activity in the excitatory network, due to the fact that the coding level is low, and only 1% of the excitatory neurons have significantly higher activity, combined with the fact that the selective activity is not much higher than the spontaneous rates. For $J_+ \sim 4$, the selective rates are between 20 and 30 spikes per second, and thus only 4–6 times the spontaneous rates.

At the risk of repetition, we recall that the two solutions are both self-reproducing and both stable. Which of the two is actually manifested in the local module depends on how the module was stimulated. If stimulated by an afferent similar to one of the memories, the high selective branch is maintained after the stimulus is removed. If stimulated by an afferent uncorrelated with any of the previously learned stimuli, the network relaxes to the spontaneous activity state.

This scenario takes place either if single synapses are potentiated gradually or if synapses have only discrete stable states but on each presentation only a fraction of the synapses that should be potentiated are actually potentiated. For example, if each synapse has only two stable states, J and J_P, and a stochastic learning dynamics [38, 23], then the resulting distribution of synapses has

$$J_+ = c_P J_P + (1 - c_P) J,$$

where c_P is the fraction of actually potentiated synapses out of the total which Hebbian learning would have potentiated had it been deterministic. Thus, if the potentiated efficacy is J_P=5J, and the critical value for the appearance of the selective attractors is $J_+ = 3.78$, we need a fraction c_P=0.70 of potentiated synapses to obtain the corresponding attractors.

The above observation could be either a prediction for cortical regions such as IT or pre-frontal cortex or as a test of the learning dynamics: If LTP takes place gradually, then during the process of training there should appear a phase with enhanced delay activity distribution identical for all stimuli in delay match-to-sample experiments. But this enhancement of the *spontaneous* activity during learning depends on the presumed learning process. In the present picture any selective neuron sees an increase of the *average* excitatory synaptic efficacy on its dendrites upon learning. This is related to the requirement that the synaptic strength averaged over the whole network remain constant during learning, and to the resulting fact that efficacies of synapses on non-selective neurons decrease. If during learning the average synaptic efficacy remained constant on the dendrites of every individual neuron, the *spontaneous* activity of any neuron would not be modified during learning. At the other extreme, if learning decreases the mean average synaptic strength, because in average depression is stronger than potentiation, the *spontaneous* activity of selective neurons would decrease with learning. In any case the observation of the evolution of the spontaneous activity during learning would give rather interesting indications on the learning dynamics.

It is rewarding to note that the range of values of both selective and spontaneous rates includes rates observed in experiments in IT cortex[8, 9, 10] and in prefrontal cortex[11].

6 Discussion

The present study integrates the structureless spontaneous neural activity into the framework of attractor neural networks. It makes spontaneous activity into another stable collective property of a system of interacting neurons. The particular stable state of spontaneous activity differs from usual attractors previously considered in neural dynamics first by its global nature, second by its extra stability and third by the fact that it can sustain local structured metastable states, the learned selective delay activity distributions. The combination of the above properties has led to the label 'vacuum' for this state.

Even if the present description does not deflect all criticisms, some of which we will ourselves suggest below, we consider it a significant step forward in the long road to the construction of a neuro-biologically plausible neural network dynamics, directly confrontable with experiment and at the same time expressing cognitively related tokens – delay activities in cognitively performing animals. It demonstrates that no external agent is required to maintain this pervasive, low rate activity and that selective delay activity can coexist with the spontaneous activity even if the rates of the two are not very widely separated. It connects levels of spontaneous activity to learning and identifies the appearance of selective delay activity as a bifurcation caused by the quantitative level of learning. To our knowledge this ensemble of properties in a single framework is here exhibited for the first time.

All the above requires an interplay between excitation and inhibition. But that interplay is rather robust: It is unstructured; it has a wide scope in the space of neural and network parameters; and the functional domain of parameters includes biologically plausible values.

Yet a number of issues remain open. We start with the seemingly lighter ones. The study has focused on the stationary states of the dynamics and on their stability. We have also simulated the current-rate dynamics for a network of 2000 neurons. This network was considered as the local module, the global environment was introduced by hand. The findings confirm the reported results. The main difference between the results of the simulated dynamics and the above analysis is that in the simulation, the rates in populations of functionally similar neurons are not uniform. Rather, in each population there is a wide distribution of rates around a mean rate that is close to the rate determined by the analysis. These wide distributions are, of course, more realistic.

For lack of space these simulations will be reported whenever we overcome the next hurdle which is the verification that all is well when our simple neurons actually emit spikes (not rates).

Deeper are the issues of the full linearity on the input side of our neurons and of the assumed relation of the time constants. Concerning the first issue our position is as follows: Had we had our choice the scenario would be that when a couple of hundred spikes arrive independently each 10ms and are scattered randomly on the dendrites, the non-linear cable effects are negligible. In fact, it has been argued[21, 46] that the simple leaky integrate-and-fire neuron and a complex reconstructed layer V pyramidal cell have similar temporal dynamics.

A second line of defence is that even in the presence of non-linearities it may be the case that one can approximate the distribution of depolarization at the soma by a Gaussian, whose first moments can be calculated as a single valued function of the afferent rates. The calculation then can be redone to obtain the conditions for self-reproduction and stability, though the resulting conditions may be different. Such is the case if the underlying synaptic dynamics includes reversal potentials, but no geometrical structure. We find that in that case the system is more stable, or that the acceptable range of parameters expands. See e.g. Appendix D. Beyond the stabilization of the spontaneous activity, the depolarization due to the learned structured stimuli can be considered marginal and hence linear, in the spirit of ref. [20].

For what concerns the issue of the time constants associated with the synaptic conductance changes, we argue in Appendix C that the introduction of finite time constants associated with the input currents does not have necessarily severe effects on the stability of the network.

Finally, as mentioned in the Introduction, all the above analysis holds in the absence of synchronization of neuronal spike activities. The threat of synchronization cannot be addressed within the terms of the present analysis. It is being tested by simulations. While the possibility of synchronization is real, see e.g. refs. [48, 19], preliminary results indicate that in a network of sufficiently many neurons operating at low rates, with a randomly diluted connectivity matrix and stochastic external afferents, no synchronization is observed. We have considered a network in which all neurons in a given population are

equivalent. Fluctuations of the connectivity from neuron to neuron produces a variability of rates between the neurons of the local module. Our preliminary simulations indicate that this variability of rates does not compromise the stability of the low rate state. The analysis of such a situation is currently under way.

Acknowledgments

We have benefitted from discussions with Simone Sergi and Stefano Fusi. We are grateful to Carl van Vreeswijk and Marc Potters for a critical reading of the manuscript, and particularly to Misha Tsodyks for his many detailed comments. This work was supported in part by a grant from the Consiglio Nazionale delle Ricerche.

Appendix

A. Stability in networks of excitatory neurons

The dynamic behavior of our neurons (Eq. 17) is determined by the mean and the variance of the depolarization. In fact[26], one can obtain the mean interspike interval (the inverse rate) in terms of the mean and variance of the 'free' depolarization, in the absence of a threshold. In a steady state situation the mean and the variance are related to each other and both are directly deducible from the afferent rates. When the depolarization evolves according to Eq. 3, and each synapse is activated by an independent Poisson process with rate ν or ν_{ext}, its steady-state mean and variance, in absence of a threshold, are (e.g. ref. [26], Eqs. 9.148, 9.149)

$$\mu[V] = JC[x\nu + (1-x)\nu_{ext}] \quad \text{and} \quad \sigma[V]^2 = \frac{J^2\lambda C}{2}[x\nu + (1-x)\nu_{ext}]$$

From now on we set $J = 1$.

If the steady state situation is perturbed, as would be the case in testing stability, the expectation values are no longer simply related, there is a adjustment time for both due to the finite integration time of the neuron. They obey the evolution equations ([26], Eqs. 9.146, 9.147)

$$\partial_t \mu[V] = -\mu[V] + C[x\nu + (1-x)\nu_{ext}] \tag{45}$$

and

$$\partial_t \left(\sigma[V]^2\right) = -2\sigma[V]^2 + \lambda C[x\nu + (1-x)\nu_{ext}] \tag{46}$$

which can be rewritten

$$\partial_t \sigma[V] = -\sigma[V] + \frac{\lambda C[x\nu + (1-x)\nu_{ext}]}{2\sigma[V]}.$$

A perturbation in the rates of neurons in the local module $\nu \to \nu + \delta\nu$ induces a variation in $\mu[V]$ $(\delta\mu)$ and in $\sigma[V]$ $(\delta\sigma)$, which determine the output spike rates via the transduction function ϕ. Making the linear variations in Eqs. (45, 46) one finds

$$\partial_t \delta\mu = -\delta\mu + xC\delta\nu \quad \text{and} \quad \partial_t \delta\sigma = -2\delta\sigma + \frac{\lambda}{\sigma[V]} xC\delta\nu.$$

The variation in the output rate $\delta\nu$ is a function of ($\delta\mu$, $\delta\sigma$) given by

$$\delta\nu = \frac{\partial\phi}{\partial\mu}\delta\mu + \frac{\partial\phi}{\partial\sigma}\delta\sigma$$

where the partial derivatives are taken at the self-reproducing rate $\nu_\star$ given by Eq. (18), and ν_{ext} is kept fixed. Stability is assured if the eigenvalues of the matrix

$$\begin{pmatrix} -1 + xC\dfrac{\partial\phi}{\partial\mu} & xC\dfrac{\partial\phi}{\partial\sigma} \\ \\ xC\dfrac{\partial\phi}{\partial\mu}\dfrac{\lambda}{\sigma} & -2 + xC\dfrac{\partial\phi}{\partial\sigma}\dfrac{\lambda}{\sigma} \end{pmatrix} \tag{47}$$

are all negative. If $\partial\phi/\partial\mu > 0$ the condition reduces to

$$xC\left(\frac{\partial\phi}{\partial\mu} + \frac{\partial\phi}{\partial\sigma}\frac{\lambda}{2\sigma}\right) < 1$$

This is the analog of Eq. (14) for integrate-and-fire units.

B. Stability in presence of inhibition

If excitatory and inhibitory inputs are independent the mean of the depolarization is simply the sum of the depolarizing and hyperpolarizing terms. The variance is the sum of the depolarizing and hyperpolarizing input variances. This leads to Eqs. (20-23). The evolution of the moments is determined from evolution equations

$$\partial_t \mu[V_E] = -\mu[V_E] + C_{EE}[x\nu_E + (1-x)\nu_{ext}] - C_{EI}J_{EI}\nu_I \tag{48}$$

and

$$\partial_t \left(\sigma[V_E]^2\right) = -2\sigma[V_E]^2 + \lambda\left(C_{EE}[x\nu_E + (1-x)\nu_{ext}] + C_{EI}J_{EI}^2\nu_I\right) \tag{49}$$

and we have similar equations for the inhibitory neurons. These are minor extensions of Eqs. 9.146 and 9.147 of ref. [26]. The effect of the perturbation in the rates $(\nu_\alpha) \to (\nu_\alpha + \delta\nu_\alpha)$ on the means and the variances is determined from these equations as in Appendix A and we obtain

$$\partial_t \delta\mu[V_E] = -\delta\mu[V_E] + xC_{EE}\delta\nu_E - C_{EI}J_{EI}\delta\nu_I$$

$$\partial_t \delta\sigma[V_E] = -2\delta\sigma[V_E] + \frac{\lambda}{\sigma[V_E]}\left(xC_{EE}\delta\nu_E + C_{EI}J_{EI}^2\delta\nu_I\right),$$

for the excitatory neurons, and for the inhibitory cells

$$\partial_t \delta\mu[V_I] = -\frac{1}{\tau_I}\delta\mu[V_I] + xC_{IE}\delta\nu_E - C_{II}J_{II}\delta\nu_I$$

$$\partial_t \delta\sigma[V_I] = -\frac{2}{\tau_I}\delta\sigma[V_I] + \frac{\lambda}{\sigma[V_I]}\left(xC_{IE}\delta\nu_E + C_{II}J_{II}^2\delta\nu_I\right).$$

The shifts in the rates are given via the corresponding ϕ_α of Eq. (30), i.e.

$$\delta\nu_\alpha = \frac{\partial\phi_\alpha}{\partial\mu_\alpha}\delta\mu_\alpha + \frac{\partial\phi_\alpha}{\partial\sigma_\alpha}\delta\sigma_\alpha,$$

taken at the self-reproducing rates satisfying Eqs. (28, 29). These solutions will be stable if all the eigenvalues of the following matrix are negative

$$\begin{pmatrix} -1+xX_E & xY_E & -J_{EI}X_I & -J_{EI}Y_I \\ xX_E\frac{\lambda}{\sigma_E} & -2+xY_E\frac{\lambda}{\sigma_E} & J_{EI}^2X_I\frac{\lambda}{\sigma_E} & J_{EI}^2Y_I\frac{\lambda}{\sigma_E} \\ xX_E & xY_E & -\frac{1}{\tau_I}-J_{II}X_I & -J_{II}Y_I \\ xX_E\frac{\lambda}{\sigma_I} & xY_E\frac{\lambda}{\sigma_I} & J_{II}^2X_I\frac{\lambda}{\sigma_I} & -\frac{2}{\tau_I}+J_{II}^2Y_I\frac{\lambda}{\sigma_I} \end{pmatrix} \tag{50}$$

in which

$$X_\alpha = C_\alpha\frac{\partial\phi_\alpha}{\partial\mu}; \qquad Y_\alpha = C_\alpha\frac{\partial\phi_\alpha}{\partial\sigma}.$$

In the simplified case of $\tau_I = 1$ (i.e. equal inhibitory and excitatory time constants); $J_{EI} = J_{II} = g$ (equal average inhibitory coupling to both types of neurons) we have $\sigma_E = \sigma_I$.

The matrix (50) has two eigenvalues, -1 and -2 and the other two are negative if

$$xX_E - gX_I + (xY_E + g^2Y_I - xg(g+1)[X_EY_I - X_IY_E])\frac{\lambda}{2\sigma} < 1 \tag{51}$$

which is the analog of Eq. (26) for the integrate-and-fire neurons. If excitatory and inhibitory thresholds are equal, $\theta_E = \theta_I$, excitatory and inhibitory transduction functions become equivalent and Eq. (51) becomes

$$(xC_E - gC_I)\phi'_\mu + (xC_E + g^2C_I)\phi'_\sigma\frac{\lambda}{2\sigma} < 1.$$

In other cases one has to find numerically the eigenvalues of (50).

C. Finite conductance time constants

In this appendix we present some preliminary arguments that tend to indicate that the effect of finite conductance time constants will not effect drastically the stability analysis.

Since the major contribution to variation of the rates comes from the variation of the mean depolarization, we neglect the variations in the variance of the depolarization when the rates are perturbed. We consider first a purely excitatory network. There are coupled perturbations of $\mu[I]$, the mean of the afferent currents to the soma, and $\mu[V]$, the mean of the depolarization, in response to a perturbation in the rates $\nu_\star \to \nu_\star + \delta\nu$. The perturbations $\delta\mu[I]$ and $\delta\mu[V]$ follow their respective evolution equations:

$$\tau' \frac{d\delta\mu[I]}{dt} = -\delta\mu(I) + xC\delta\nu$$

$$\frac{d\delta\mu[V]}{dt} = -\delta\mu[V] + \delta\mu[I].$$

Neglecting the fluctuation of the variance, the response of the rate, $\delta\nu$, is given by

$$\delta\nu = \phi'_\mu \delta\mu[V] \tag{52}$$

where $\phi'_\mu = \frac{\partial\phi}{\partial\mu}$ taken at the self-reproducing rates, Eq. (18), at fixed ν_{ext}. Note that in Eq. (52) the transduction function is no longer given by Eq. (4), because the source to the depolarization is now correlated in time by the finite time constant τ'.

The evolution of the perturbation is governed by the eigenvalues

$$\lambda_\pm = -\frac{1+\omega}{2} \pm \frac{1}{2}\sqrt{(1+\omega)^2 - 4\omega(1 - xC\phi'_\mu)}$$

where $\omega = \frac{1}{\tau'}$. The condition for both eigenvalues to be negative is $xC\phi'_\mu < 1$, and the dependence of the stability condition on τ' is contained in the derivative of the transduction function at the self-reproducing rate. When $\tau' = 0$ we have seen in Section 3 that $xC\phi'_\mu$ is so much greater than 1 that we tend to conclude that a finite τ' is not likely to lower ϕ'_μ enough to satisfy the stability condition. In fact we expect that the suppression of the fluctuations by the preceding filter will tend to increase the slope[49]. The all-excitatory network in the self-reproducing rates is likely to be unstable for any τ'.

Next we include inhibition and repeat the analysis. For simplicity we discuss the case of

- equal integration time constants for the depolarizations of excitatory and inhibitory units;
- equal relative efficacies of inhibitory synapses on excitatory and inhibitory units $J_{EI} = J_{II} = g$;
- equal integration time constants for the currents afferent onto excitatory and inhibitory neurons.

In this case the eigenvalues of the linear stability matrix are given by

$$\lambda_{\pm} = -\frac{1+\omega}{2} \pm \frac{1}{2}\sqrt{(1+\omega)^2 - 4\omega(1-X)}$$

where $X = xC_E\phi'_{E,\mu} - C_I g\phi'_{I,\mu}$, and $\phi'_{\alpha,\mu}$ are the respective derivatives of the excitatory and inhibitory transduction function at the self-reproducing rates, which again depend on the synaptic time constant τ'. The stability condition is $X < 1$. The condition can be expressed as a bound on g for the spontaneous rates to be stable, which, for $xC_E\phi'_{E,\mu} \gg 1$, reads

$$g > \frac{xC_E\phi'_{E,\mu}}{C_I\phi'_{I,\mu}}$$

If the synaptic time constant is equal for synapses on excitatory and inhibitory neurons, which is the case we assume here, τ' will have a similar effect on both transduction functions $\phi_{E,I}$. As a consequence the critical relative inhibitory strength will not change significantly when τ' is finite.

In conclusion, the introduction of finite decay times for conductance changes may not have a strong effect on the stability of the self-reproducing rates, provided they are equal in excitatory and inhibitory synapses and that the transduction function does not change too abruptly. However, the finite integration time at the synapses acts as a low pass filter, generating temporal correlations in the source of the depolarization. This may change significantly the output frequency as a function of the input rates and the above analysis may no longer hold. Some work investigating the dependence of the transduction function on τ'/τ has been started[49]. But it is still too early to draw conclusions from it.

D. Nonlinear dynamics and reversal potentials

We present here a tentative argument indicating that stability may be improved by non-linearities on the input side. We consider a neuron without a geometric structure and the nonlinearities are due to reversal potentials in the in the synaptic conductivities. The membrane depolarization evolves according to

$$\dot{V} = -V + (V_E - V)\sum_i g_i^E \sum_j \delta(t - t_{ij}^E) + (V_I - V)\sum_i g_i^I \sum_j \delta(t - t_{ij}^I)$$

where $V_{E,I}$ are, respectively, the excitatory and inhibitory reversal potentials. The index i refers to the synaptic site, at which the conductance change caused by a spike arriving at synaptic site i is $g_i^{E,I}$ when a spike arrives, and t_{ij} is the arrival time of the jth spike. V_I=0 is the case of shunting inhibition.

In absence of a threshold the steady-state mean of the depolarization is given by ([26], 9.190)

$$\mu[V] = \frac{V_E g_E C_E \nu_E + V_I g_I C_I \nu_I}{1 + g_E C_E \nu_E + g_I C_I \nu_I} \tag{53}$$

and the variance is

$$\sigma[V]^2 = \lambda \frac{V_E^2 g_E^2 C_E \nu_E + V_I^2 g_I^2 C_I \nu_I}{2(1 + g_E C_E \nu_E + g_I C_I \nu_I)} \tag{54}$$

where λ is related to the variance of the synaptic distributions by $\lambda = 1 + \Delta^2$. See also e.g. ref. [50].

The means and variances will be less sensitive to variations in the input rates than in the linear case. These variations will typically be reduced by the shunting factor: $1 + g_E C_E \nu_E + g_I C_I \nu_I$. The conclusion is that, at parity of parameters, the network with nonlinear elements will have more stable self-reproducing rates than in the linear case.

Of course in this case the neuronal transduction function will differ from Eq. (4). A complete treatment of this situation would require the computation of the neural output frequency in a diffusion approximation, much like the analysis for the neuron that sums linearly its inputs[26]. The mean inter-spike interval is the solution $f(x)$ of the differential equation

$$\frac{1}{2}\left[\gamma_E (V_E - x)^2 + \gamma_I (V_I - x)^2\right] f''(x) + (\alpha - \beta x) f'(x) = -1$$

satisfying the boundary conditions $f(\theta) = 0$ and $f(x) = 0$ for $x \to -\infty$, evaluated at the post-spike hyper-polarization $x = H$. The various parameters in the equation are defined by:

$$\alpha = V_E g_E C_E \nu_E + V_I g_I C_I \nu_I$$

$$\beta = 1 + g_E C_E \nu_E + g_I C_I \nu_I$$

and

$$\gamma_{E,I} = \lambda g_{E,I}^2 C_{E,I} \nu_{E,I}.$$

Given the transduction function it is possible to repeat the analysis to obtain the self-reproducing rates, and their stability. We have not pursued this program, which is a manageable numerical task. At this stage we leave the argument at the intuitive level discussed at the beginning of this appendix.

References

[1] Bernander O, Douglas RJ, Martin KAC and Koch C 1991, Synaptic background activity determines spatio-temporal integration in single pyramidal cells *Proc. Natl Acad Sci USA* **88** 1569

[2] Rapp M, Yarom Y and Segev I 1992, The impact of parallel background activity on the cable properties of cerebellar Purkinje cells *Neural Computation* **4** 518

[3] Amit DJ and Tsodyks MV 1992 Effective neurons and attractor neural networks in cortical environment, *Network* **3** 121

[4] Larkman A 1991 *J. Comp. Neurol.* **306** 332.

[5] Martin KAC 1988 *Q. J. Exp. Physiol.* **73** 637

[6] Rall W 1977 Core conductor theory and cable properties of neurons, In Kandel ER ed. *Handbook of Physiology* (Amer. Physiol. Soc., Bethesda) Vol. I.

[7] Amit DJ 1989 *Modeling Brain Function* (Cambridge University Press, NY)

[8] Miyashita Y and Chang HS 1988 Neuronal correlate of pictorial short-term memory in the primate temporal cortex, *Nature*, **331** 68

[9] Miyashita Y 1988 Neuronal correlate of visual associative long-term memory in the primate temporal cortex, *Nature* **335** 817

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

[11] Wilson FAW, Scalaidhe SPO and Goldman-Rakic PS 1993 Dissociation of Object and Spatial Processing Domains in Primate Prefrontal Cortex, *Science*, **260** 1955

[12] Nelken I 1988 Analysis of the activity of single neurons in stochastic settings, *Biol. Cybern.*, **59**.

[13] Komatsu Y, Nakajima S, Toyama K, Fetz E 1988 Intracortical connectivity revealed by spike-triggered averaging in slice preparations of cat visual cortex, *Brain Res.* **442** 359

[14] Mason A, Nicoll A, Stratford K 1991 Synaptic transmission between individual pyramidal neurons of the rat visual cortex *in vitro*, *J. Neurosci.*, **11** 72

[15] Sayer R, Friedlander M, Redman S 1990 The time course and amplitude of EPSPs evoked at synapses between pairs of CA3/CA1 neurons in the hippocampal slice, *J. Neurosci.* **10** 826

[16] McNaughton B, Barnes C, Andersen P 1981 Synaptic efficacy and EPSP summation in granule cells of rat fascia denta studied *in vitro J. Neurophysiol.* **46** 952

[17] Thompson A, Girdlestone D, West D 1988 Voltage dependent currents prolong single-axon postsynaptic potentials in layer III pyramidal neurons in rat neocortical slices *J. Neurophysiol.* **60** 1896

[18] Braitenberg V and Schuz A 1991 *Anatomy of the cortex*, (Springer-Verlag, Berlin).

[19] VanVreeswijk C and Abbott LF 1992 Self-sustained firing in populations of integrate and fire neurons, *SIAM J. Appl. Math.* **53** 253

[20] Amit DJ and Tsodyks MV 1991 Quantitative study of attractor neural network retrieving at low spike rates I: Substrate – spikes, rates and neuronal gain *Network* **2** 259

[21] Softky WR and Koch C 1993 The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs, *J. Neurosci.* **13** 334

[22] Tsodyks M and Sejnowski T 1994 Rapid state switching in balanced cortical network models, Salk Institute preprint

[23] Amit DJ and Brunel N 1994 Learning internal representations in an attractor neural network with analogue neurons, *Network* to be published

[24] See e.g. Schweber SS 1961 *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson and Co., Evanston Ill.), ch. 17.

[25] Ricciardi LM 1977 *Diffusion processes and Related topics on biology* (Springer-Verlag, Berlin)

[26] Tuckwell HC 1988 *Introduction to theoretical Neurobiology* (Cambridge: Cambridge University Press)

[27] Douglas R, Martin K 1991 A functional microcircuit for cat visual cortex, *J. Physiol. (Lond)* **440** 735

[28] Frolov AA and Medvedev AV 1986 Substantiation of the 'point approximation' for describing the total activity of the brain with the use of a simulation model *Biofizika 31* 304 (Engl. transl. 1986 *Biophysics* **31** 332)

[29] Hestrin S, Sah P, Nicoll R 1990 Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices *Neuron* **5** 247

[30] Bekkers JM and Stevens CF 1989 NMDA and non-NMDA receptors are co-localized at individual excitatory synapses in cultured rat hippocampus *Nature* **341** 230

[31] Zador A, Koch C and Brown TH 1990 Biophysical model of a Hebbian synapse *Proc. Natl. Acad. Sci. USA* **87** 6718

[32] Usher M, Stemmler M, Koch C and Olami Z 1994 Network amplification of local fluctuations causes high spike rate variability, fractal firing patterns and oscillatory local field potentials *Neural Computation* **6** 795

[33] Softky WR 1993 Sub-millisecond coincidence detection in active dentritic trees, *Neuroscience* **58** 13

[34] Abramowitz M and Stegun IA 1970 *Tables of Mathematical Functions* (Dover Publications, NY).

[35] Traub RD and Miles R 1991 *Neuronal Networks of the hippocampus* (Cambridge University Press, Cambridge)

[36] Simons DJ 1978 Response properties of vibrissa units in rat SI somatosensory neocortex *J. Neurophysiol.* **41** 798

[37] Connors BW and Gutnick MJ 1990 Intrinsic firing patterns of diverse neocortical neurons *TINS* **13** 99

[38] Amit DJ and Fusi S 1994 Dynamic learning in neural networks with material synapses, *Neural Computation* **6** 957

[39] Brunel N 1994 Dynamics of an attractor neural network converting temporal into spatial correlations, *Network*, to be published.

[40] Stanton PK and Sejnowsky TJ 1989 Associative long-term depression in the hippocampus induced by Hebbian covariance, *Nature* **339** 215

[41] Willshaw D, Buneman O P and Longuet-Higgins H 1969 Non-holographic associative memory, *Nature*, **222** 960

[42] Tsodyks MV and Feigel'man MV 1988, The enhanced storage capacity in neural networks with low activity level, *Europhys. Lett.*, **46** 101

[43] Brindley GS 1969 Nerve net models of plausible size that perform many simple learning tasks, *Proc. Roy. Soc. Lond. B* **174** 173

[44] Marr D 1970 A theory for cerebral neocortex, *Proc. Roy. Soc. Lond. B. 176* 161

[45] Amit DJ and Brunel N 1993 Adequate input for learning in attractor neural networks, *Network* **4** 177

[46] Bernander O, Koch C and Usher M 1994 The effect of synchronized inputs at the single neuron level, *Neural Computation*, **6**, 622

[47] Amit DJ Brunel N and Tsodyks MV 1994 Correlations of cortical Hebbian reverberations: experiment *vs* theory, *J. Neurosci.* **14**

[48] Tsodyks M, Mitkov I and Sompolinsky H 1993 Pattern of synchrony in inhomogeneous networks of oscillators with pulse interactions *Phys. Rev. Lett.* **71** 1280

[49] Sergi S 1994, thesis University of Firenze, unpublished.

[50] Abbott L 1991 Realistic synaptic inputs for model neural networks *Network* **2** 245