INSTITUTE OF PHYSICS PUBLISHING NETWORK: COMPUTATION IN NEURAL SYSTEMS

Network: Comput. Neural Syst. **12** (2001) 47–74 www.iop.org/Journals/ne PII: S0954-898X(01)20331-5

Properties of synaptic transmission and the global stability of delayed activity states

Alexei A Koulakov

The Salk Institute, La Jolla, CA 92037, USA

Received 20 March 2000, in final form 29 November 2000

Abstract

The influence of synaptic channel properties on the stability of delayed activity maintained by recurrent neural networks is studied. The duration of excitatory post-synaptic current (EPSC) is shown to be essential for the global stability of the delayed response. The NMDA receptor channel is a much more reliable mediator of the reverberating activity than the AMPA receptor, due to a longer EPSC. This allows one to interpret the deterioration of the working memory observed in NMDA channel blockade experiments. The key mechanism leading to the decay of the delayed activity originates in the unreliability of synaptic transmission. The optimum fluctuation of the synaptic currents leading to the decay is identified. The decay time is calculated analytically and the result is confirmed computationally.

1. Introduction

Long-term memory is thought to be stored in biochemical modulation of the interneuron synaptic connections. As each neuron forms about 10^4 synapses, such a mechanism of memory is potentially very effective, allowing the storage of $>10^4$ bits per neuron. However, although synaptic modifications persist for a long time, it may take as long as several minutes to form them. Since the external environment operates on much shorter timescales, synaptic plasticity is virtually useless for short-term survival needs. Such needs are satisfied by working memory (WM), which is stored in the state of neuronal activity, rather than in modification of synaptic conductances (Miyashita 1988, Miyashita and Chang 1988, Sakai and Miyashita 1991, Funahashi *et al* 1989, Goldman-Rakic *et al* 1990, Fuster 1995).

WM is believed to be stored in recurrent neural circuits (Wilson and Cowan 1972, Amit and Tsodyks 1991a, b, Goldman-Rakic 1995). A recurrent neural network can have two stable states (attractors), characterized by low and high firing frequencies. External inputs produce transitions between these states. After transition the network maintains high or low firing frequency during the delay period. Such a bistable 'switch' can therefore store one bit of information. Similar solutions exist in associative memory networks (Gerstner and van Hemmen 1992a, b).

Over the last few years evidence has been accumulating that the NMDA receptor (NMDAR) plays an important role in the mechanism of WM. Studies in humans (Scherzer *et al*

0954-898X/01/010047+28$30.00

1998) and rats (Cotman *et al* 1987) show that this receptor is expressed at the highest density in the prefrontal cortex, the area involved in WM storage. Notably, the density of NDMAR expression in human prefrontal cortex is higher than in the hippocampus. The importance of NMDAR for WM is also evident from the impairment of the capacity to perform the delayed response task produced by the receptor blockade (Krystal *et al* 1994, Adler *et al* 1998, Pontecorvo *et al* 1991, Cole *et al* 1993, Aura and Riekkinen 1999). Similarly the injection of the NMDAR antagonists brings about weakening of the delayed activity demonstrated in electrophysiological studies (Javitt *et al* 1996, Dudkin *et al* 1997a). At the same time intracortical perfusion with the glutamate receptor agonists both improves the performance in the delayed response task and increases the activity of cells during the delay period (Dudkin *et al* 1997a, b). Such evidence is especially interesting since the administration of NMDAR antagonists (PCP or ketamine) reproduces many of the symptoms of schizophrenia including deficits in WM (Javitt and Zukin 1991, Krystal *et al* 1994, Goldman-Rakic 1994).

One of the distinguishing features of the NMDAR is the nonlinearity of its current–voltage characteristics. This feature was recently implicated as being crucial for WM (Lisman *et al* 1998). It was shown that by carefully balancing NMDA, AMPA, and GABA synaptic currents one can produce an N-shaped synaptic current–voltage characteristic, which, when complemented by the recurrent neural circuitry, results in the bistability. This paper implies therefore that a single neuron can be bistable and maintain a stable membrane state corresponding to high or low firing frequencies (Camperi and Wang 1998). Such bistability, therefore, is extremely fragile and can be easily destroyed by disturbing the balance between the NMDA, GABA or AMPA conductances. This may occur in the experiments with the NMDA antagonist.

Another possibility for the critical involvement of NMDAR in WM originates from its high affinity to glutamate, resulting in the long duration of the excitatory post-synaptic current (EPSC) ($\sim$100 ms (Clements *et al* 1992, Lester *et al* 1990)). This possibility was recently studied by Wang (1999)[1]. It was shown that WM cannot be mediated by the AMPA channel alone and slow synaptic dynamics is required to achieve a stable persistent state. Such dynamics can be provided by, for example, the NMDAR.

In this paper we accept the latter point of view, that is, that the duration of EPSC is critical to WM. Therefore, following Wang (1999), we disregard the phenomena associated with the nonlinearity of the synaptic current–voltage relation (as in Lisman *et al* (1998)). Instead of concentrating on delayed activity maintenance we study how this activity fails. More exactly we consider the processes of spontaneous decay of the delayed activity associated with the failures of WM. Such failures are frequently observed experimentally (Funahashi *et al* 1989). We evaluate the average WM storage time determined by such processes of spontaneous decay. We conclude that the WM storage time is significantly larger for synapses with longer EPSC. This finding is consistent with the apparent importance of NMDAR for WM and the conclusions of Wang (1999). In addition we find that the WM storage time decreases if the neuronal firing frequency is decreased. This allows us to interpret the deterioration of WM observed in the NMDAR blockade experiments.

Our approach to the problem is based on finding the optimum fluctuation of the synaptic currents. Such a fluctuation leads most effectively to a decay in the delayed activity. Since the decay implies the termination of the recurrent current, such a fluctuation is internal with respect to the recurrent circuitry sources of noise. To address this issue we assume that the internal noise is provided by the probabilistic nature of neurotransmitter release in each synapse. This separates our model from that of Wang (1999). The optimum fluctuation produced by unreliable

[1] An earlier version of this paper, Koulakov (1999), is available.

synapses can be described by the formation of *critical nucleus* in the neuronal population. This implies that about 100 ms before the entire network shuts down a small subpopulation of the entire network ceases the persistent activity. Such a mechanism of decay of the delayed activity is more effective than that provided by external noise. The two systems of noise are shown below to be in an intricate relationship mediated by synchrony.

The decay time of the memory state evaluated below *increases* as the size of the network increases (see section 3). This is a natural result since in a large network the relative strength of noise is small according to the central limit theorem. Therefore, if one is given the minimum time during which the information has to be stored, there is a minimum network size needed to perform this task. We calculate therefore the minimum number of neurons necessary to store one bit of information in a recurrent network. This number weakly depends on the storage time and for the majority of cases is 5–15. This problem is similar to the calculation of storage capacity of a Hopfield network (see Hertz *et al* (1991) for an overview).

The analytical calculations presented below are confirmed by computer simulations. For individual neurons we use the modified leaky integrate-and-fire model due to Stevens and Zador (1998a), which is shown to correctly reproduce the timing of spikes *in vitro*. Such neurons have firing frequencies within the range 15–30 Hz for purely excitatory networks in the absence of any inhibitory input (section 2). This solves the high firing frequency problem (see e.g. Amit and Tsodyks (1991a)). The resolution of this problem is based on the unusual property of the Stevens–Zador (SZ) model, which, in contrast to a simple leaky integrator, has *two* timescales. These are the membrane time constant and the characteristic time during which the time constant changes. The latter, being much longer than the former, determines the minimum firing frequency in the recurrent network, making it consistent with the observed values (see section 2 for more detail). The SZ model is a special case of a spike response model (Gerstner 1995, Kistler *et al* 1997, Gerstner 2000).

2. The model and its approximate solution

In this section we first examine the properties of a single neuron, define our network model, and, finally solve the network approximately using the mean-field approximation.

The SZ model is an extension of the standard leaky integrator (see e.g. Tuckwell (1988)). It has been shown to accurately predict the spike timings in layer 2/3 cells of the rat sensory neocortex. The membrane potential V satisfies the leaky integrator equation with the time-varying resting potential E and integration time τ

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

Here t is the time elapsed since the last spike generated, and the input current $I(t)$ is measured in volts per second. When the membrane voltage reaches the threshold voltage θ the neuron emits a spike and the voltage is reset to V_{reset}.

This model is general enough to describe many types of neurons, differing only by the functions $E(t)$, $\tau(t)$, and the parameters θ and V_{reset}. For pyramidal cells in the cortical layer 2/3 of rats the functions can be fitted by

$$E(t) = E_0 - \Delta E[1 - \exp(-\alpha t/\tau_0)] \tag{2}$$

and

$$\tau(t) = \tau_0[1 - \exp(-\alpha t/\tau_0)]. \tag{3}$$

The parameters of the model for these cells have the following numerical values: $E_0 = -32$ mV, $\Delta E = 28$ mV, $\alpha = 0.3$, $\tau_0 = 10$ ms, $\theta = -22$ mV, and $V_{\text{reset}} = -32$ mV (Stevens

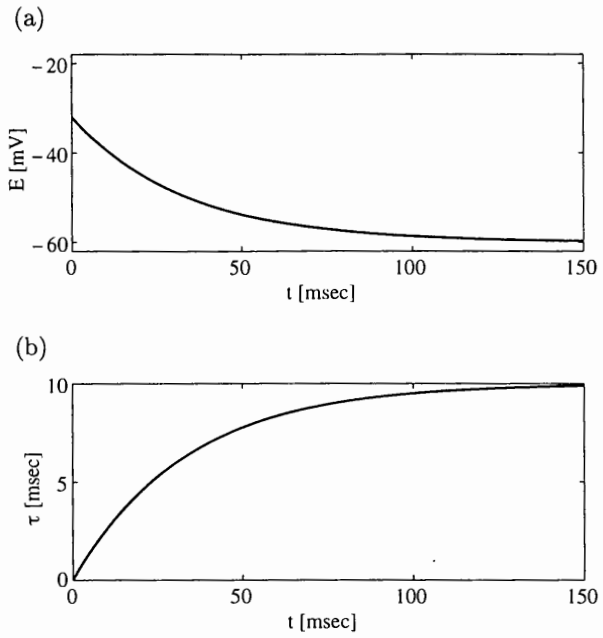

Figure 1. The somatic membrane resting potential (*a*) and the integration time (*b*) as functions of the time elapsed since the last spike.

and Zador 1998a). The resting potential and integration time with these parameters are shown in figure 1[2].

In the next step we calculate the transduction function $f = T(I)$, relating the average external current to the average firing frequency. This function is evaluated in appendix A and is shown in figure 2. The closed hand expression for this function cannot be obtained. Some approximate asymptotic expressions can, however, be found. For large frequencies ($f \gg \alpha/\tau_0 = 30$ Hz) it is approximately given by the linear function (dashed line in figure 2)

$$T(I) \approx g(I - I_0), \tag{4}$$

where

$$g = \frac{\alpha}{1+\alpha}\frac{1}{|E_0 - \theta|}, \tag{5}$$

and

$$I_0 = \frac{\Delta E}{\tau_0} + \frac{|E_0 - \theta|}{2\tau_0}\frac{1+\alpha}{1+2\alpha}. \tag{6}$$

For small frequencies ($f \ll \alpha/\tau_0$) we obtain

$$T(I) \approx \frac{\alpha}{\tau_0}\frac{1}{\ln[(|E_0 - \theta|)/(|E_0 - \theta| + \Delta E - \tau_0 I)]}. \tag{7}$$

The neuron therefore starts firing significantly when the current exceeds the critical value

$$I^* = \frac{|E_0 - \theta| + \Delta E}{\tau_0}. \tag{8}$$

[2] Since the time constant is zero at $t = 0$ to resolve the singularity an implicit Runge–Kutta scheme should be used in the numerical integration of (1).

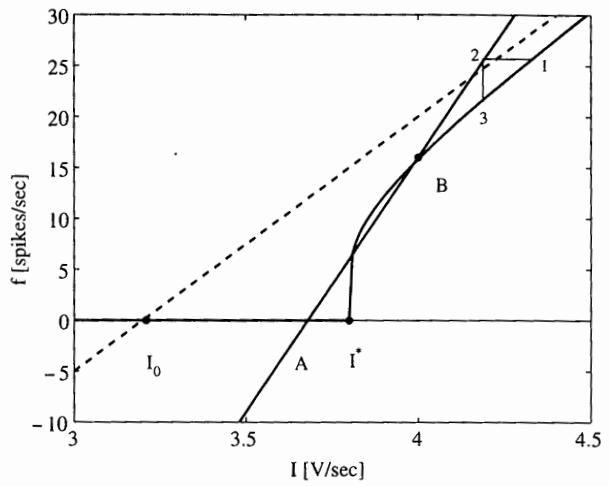

Figure 2. Neuronal transduction function relating the firing frequency and the input current for one neuron (solid line $I_0 A I^* B$). The straight line AB describes the network feedback. The dashed line represents the asymptote of the transduction function for large frequencies. It is given by equation (4). The expressions for I_0 and I^* are provided by equations (6) and (8). The points A and B are the low- and high-frequency attractors, respectively.

In the following, we ignore spontaneous activity by setting all firing frequencies below 5 Hz to zero. We also disregard the effects of the refractory period since they are irrelevant at frequencies of 10–50 Hz.

We consider a network consisting of N SZ neurons establishing all-to-all connections. After each neuron emits a spike an EPSC is generated in the input currents of all cells with the probability κ. This is intended to simulate the finite probability of neurotransmitter release in each synapse. The total input current of the nth neuron is therefore

$$I_n(t) = \sum_{k=1}^{N} \sum_{t_{s_k}} b_{n,s_k} j(t - t_{s_k}) + I_n^{\text{ext}}(t), \tag{9}$$

where k labels the neurons making synapses on the nth neuron (all neurons in the network), t_{s_k} is the time of spike number s_k emitted by cell k, and I_n^{ext} is the external current. The binary variable b is equal to 1 with probability κ (in our computer simulations always equal to 0.3 (Dobrunz and Stevens 1997)). It is the presence of this variable that distinguishes our approach from Wang (1999). We choose the EPSC represented by $j(t)$ to be (see Amit and Brunel (1997))

$$j = j_0 \exp(-t/\tau_{\text{EPSC}}) H(t), \tag{10}$$

where $H(t) = 1$, if $t > 0$, and $H(t) = 0$ otherwise. As evident from (10) τ_{EPSC} is the duration of EPSC and is therefore the central variable in our consideration.

Equation (9) implies many approximations about the synaptic input current. First, we assume that the synaptic channels are not saturated. For the NMDAR the decrease in the amplitude of EPSC in pair-pulse stimulation is $p \approx 0.8$ (Mainen *et al* 1999). In the case of a spike train bombarding the synapse the reduction in amplitude of a single EPSC can be estimated as $\delta j_0/j_0 \approx (1 - p) \exp[-1/(\kappa f \tau_{\text{EPSC}})]$. For a realistic firing rate $f \approx 20$ Hz $\delta j_0/j_0 \sim 6\%$. This slightly reduces the network feedback (see below), making the system slightly more stable against noise. This does not change our results, however, since they do

not depend on the amplitude of a single EPSC. Second, we ignore the nonlinear interactions in the dendrite, in view of the existing controversy in the experimental literature (see discussion in Kogo and Ariel (1999)). Third, the synaptic current given by equation (10) is not voltage dependent. This approximation is good for an AMPA receptor and for a NMDAR it is good as long as we consider the high-frequency attractor, in which the membrane voltage oscillates in a relatively narrow range. Finally, we disregard the inhibitory interneurons in our model. The role of the interneurons in the decay of delayed activity warrants further study but some experimental studies have suggested that the inhibitory current is small compared to the excitatory current (Stevens and Zador 1998b). Our model is quite approximate but we do feel that it captures the main features of the phenomenon. The other features can be accounted for using a rescaling of the parameters involved.

If the number of neurons in the network is large its dynamics is well described by the mean-field approximation. Although called an approximation it correctly reproduces collective behaviour of the network. The deviations from the mean-field picture are discussed in the next section. In the mean-field approximation the average firing frequency $\bar{f}$ and the average input current $\bar{I}$ describe the network completely. Thus the network dynamics can be approximated by one 'effective' neuron receiving the average input current and firing at the average frequency. The behaviour of system in this approximation, reduced to the behaviour of one effective neuron, can be understood from figure 2. The stationary points of the system are given by the points of intersection between the neuronal transduction curve and the line describing the feedback. There are three such points in figure 2. Two of them are denoted by letters A and B. The third stationary point is between A and B. Indeed, if the system is put exactly at one of these points the recurrent feedback and the neuronal transduction curve will reproduce the state of the system indefinitely, whereby the system will remain at the initial point forever. However, only two of these three stationary points are stable; these are points A and B, which represent the low- and high-frequency attractors, respectively. The stability of state B is demonstrated in figure 2 graphically; assuming that the system received a brief external perturbation, which increased its input current to the value corresponding to point 1 in figure 2, the numbers 123 then demonstrate the trajectory of the system after such a perturbation. The value of the input current corresponding to point 1 produces the firing rate in the hypothetical neuron associated with a mean-field approximation, which is described by the transduction function. This firing rate induces a new feedback value, described by the line AB and given by point 2. This feedback current induces a new value of the firing rate given by point 3. When moving from point 1 to 3 the system approaches point B, which is therefore an attractor. The stability of point A can be shown in a similar way. The point between A and B is unstable due to similar considerations and represents the edge of the two attraction basins of points A and B.

We now address the dynamics of the system in the mean-field approximation on a more quantitative level. Assume that the hypothetical neuron associated with the approximation emits spikes at a frequency of $f(t)$ (point 1 in figure 2). Due to the network feedback this results in the input current equal to the average of equation (9), displaced in time by the average duration of an EPSC, i.e.

$$I(t + \tau_{\text{EPSC}}) \approx \bar{I} = N\kappa j_0 \tau_{\text{EPSC}} \bar{f}(t) + \bar{I}^{\text{ext}}. \tag{11}$$

This corresponds to the transition between points 1 and 2 in figure 2. Finally the transition between the input current and the firing frequency is accomplished by the transduction function $T(I)$ (points 2 and 3). The delay due to this transition is of the order of a somatic membrane time constant ($\sim$10 ms) and is negligible compared to $\tau_{\text{EPSC}} \sim 100$ ms. We, therefore, obtain the equation for the firing frequency $\bar{f}(t + \tau_{\text{EPSC}}) = T(\bar{I}(\bar{f}(t)))$ (Wilson and Cowan 1972),

which has the following Taylor expansion:

$$\tau_{\text{EPSC}}\dot{\bar{f}} = T(\bar{I}(\bar{f})) - \bar{f}. \tag{12}$$

To obtain the steady state solutions we set the time derivatives in (12) to zero

$$T(I) = I^{-1}(I). \tag{13}$$

Here $I^{-1}(I)$ is the function which is inverse to (11). It is shown in figure 2 by a straight solid line. This equation has three solutions, two of which are stable and are shown by points A and B in the figure.

Point B represents the high-frequency attractor. The frequencies obtained in the SZ model neurons are not too high. They are of the order α/τ_0 (=0.3/10 ms in our case), i.e. in the range 15–30 Hz (see e.g. figure 3(b)). This coincides with the range of frequencies observed in the delayed activity experiments (Miyashita 1988, Miyashita and Chang 1988, Sakai and Miyashita 1991, Funahashi *et al* 1989, Goldman-Rakic *et al* 1990, Fuster 1995). The reason for the relatively low firing frequency is as follows. For the leaky integrator model the characteristic firing rates in the recurrent network are of the order $1/\tau$ (1/10 ms), i.e. they are in the range of 50–100 Hz. For the SZ neuron τ is even smaller (see figure 1(b)), and therefore it seems that the firing rates should be larger than for a leaky integrator. This, however, is not true, since for the SZ neuron, in contrast to the leaky integrator, there is a second timescale. It is the characteristic time of variation of the time constant $\tau_0/\alpha \sim$ 30–50 ms (see figure 1(b)). Since the time constant itself is very short it becomes irrelevant for spike generation purposes at low firing rates and the second timescale determines the characteristic frequency. It is, therefore, in the range 15–30 Hz.

Finally we would like to discuss the stability of the attractor B. It can be locally stable, i.e. small perturbations cannot produce a transition from state B to state A. The condition for this follows from the linearizing equation (12) near the equilibrium

$$\nu = \frac{\mathrm{d}T}{\mathrm{d}\bar{I}}\frac{\mathrm{d}\bar{I}}{\mathrm{d}\bar{f}} < 1. \tag{14}$$

Very rarely, however, a large noise fluctuation can occur that kicks the system out of the basin of attraction of state B. It is, therefore, *never* globally stable, which is the topic of the next section.

It should be noted that the dynamics of spiking neural networks is described by equation (12) only approximately. One can justify the firing rate approach implied by (12) if the averaging time τ_{EPSC} includes a few spikes, i.e. $\tau_{\text{EPSC}} f \gg 1$. Since $\tau_{\text{EPSC}} f \approx 2$ in our numerical experiments this equation applies only marginally. The other case when equation (12) is valid is for the case of an asynchronous state. The stability of this state is analysed in detail in Abbot and van Vreeswijk (1993), Gerstner and van Hemmen (1993) and Gerstner (1995, 2000). In the following we first assume that the asynchronous state is stable, and then consider the effects of synchronization on the global stability of the high-frequency state.

3. Spontaneous decay of persistent activity

Computer simulations show that our network can successfully generate delayed activity responses. An example is shown in figure 3(a) where a short pulse of an external current (dashed line) produces a transition to the high-frequency state. This state is well described by the mean-field treatment given in the previous section.

This is true, however, only for networks containing a large number of neurons N. If the size of the network is smaller than some critical number N^*, the following phenomenon is observed

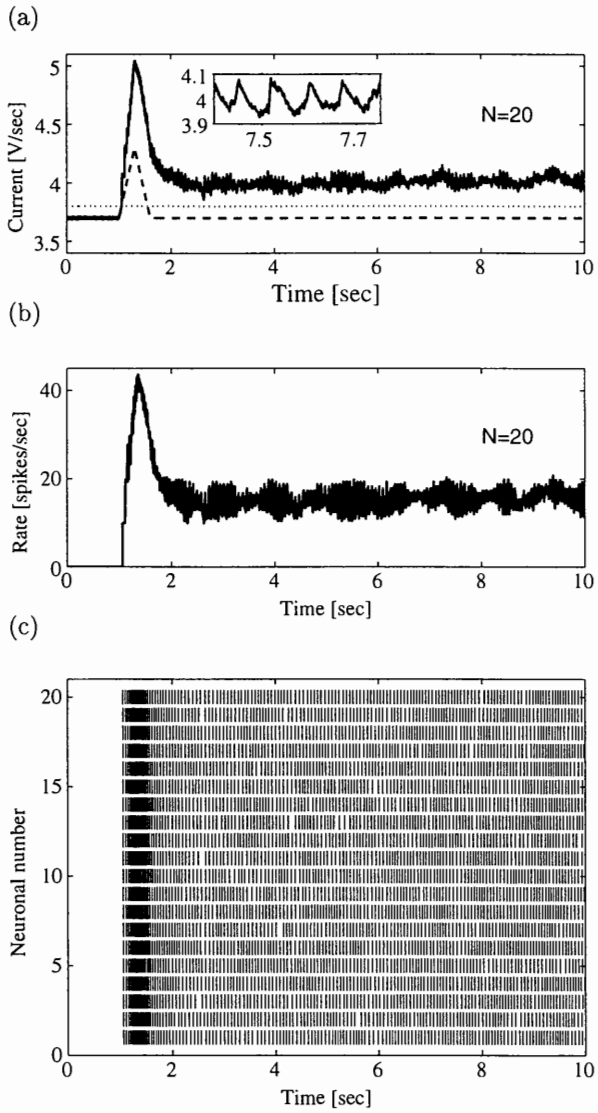

Figure 3. In a large networks ($N = 20$) the delayed activity can persist virtually forever. (*a*) Total input current averaged over all the neurons in the network (solid curve). A short pulse of the external current (dashed line) brings about the transition to the high-frequency state, in which the total current is larger than the external current by the value of feedback. The edge of the attraction basin is displayed by the dotted line. The inset shows the traces of neuronal synchronization: the average current experiences oscillations at the average firing frequency. (*b*) The firing rate averaged over the network as a function of time. (*c*) The spiking rastergram for each neuron in the network. All plots are shown for the same run of the simulation program.

(see figure 4(*a*)). The fluctuations of current reach the edge of the attraction basin (dotted line) and the network spontaneously shuts down, jumping from the high-frequency state to the low-frequency state. The quantitative treatment of these events is the subject of this section.

Similar decay processes have been observed by Funahashi *et al* (1989) in the prefrontal cortex. One example of such an error trial is shown in figure 5.

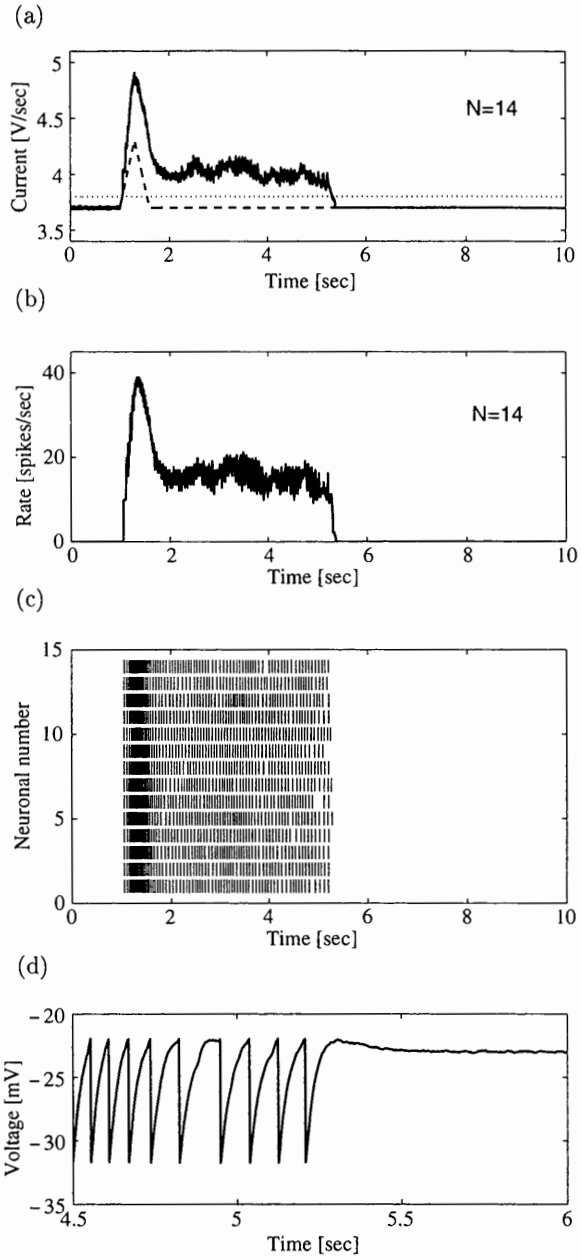

Figure 4. Same as in figure 3 but for the network containing $N = 14$ neurons. In the small network the high-frequency state can decay and the delayed activity terminates. This results in the average current crossing the edge of the attraction basin (the edge is shown by the dotted line in (*a*)) or in a sudden decrease of the firing rate, as in (*b*). (*d*) Also shown is the membrane voltage of neuron no 1 (very bottom of (*c*)) as a function of time for the moment of decay of the persistent activity.

Although the decay is abrupt, the moment at which it occurs is *not* reproducible from experiment to experiment. It is of interest, therefore, to study the distribution of the time intervals between the initiation of the delayed activity and the moment of its decay. Our

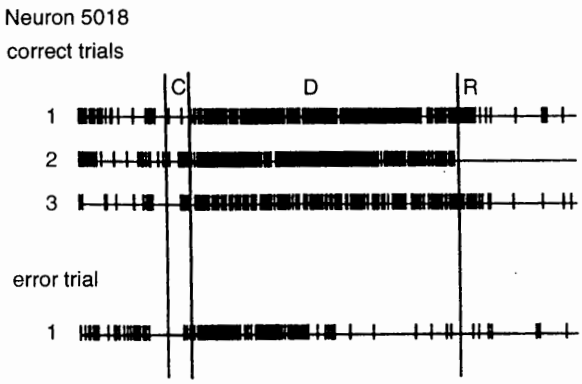

Figure 5. Sudden decay of the delayed activity in a neuron in the prefrontal cortex during an error trial (bottom figure) observed by Funahashi *et al* (1989). Compare to the rastergram for neuron no 6 in figure 4(*c*).

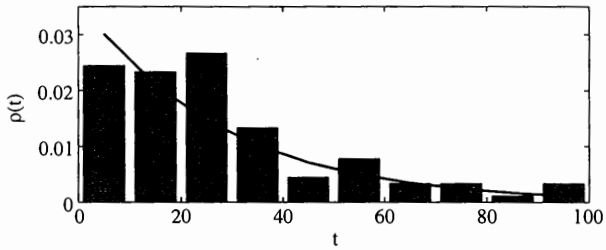

Figure 6. The probability density of decay at time t obtained numerically (bars) and the exponential distribution with $\bar{t} = 28$ s (solid curve). The latter is given by equation (B.8). The numerical results include 100 trials with $\bar{f} \approx 16$ Hz, $N = 14$, $\tau_{\mathrm{EPSC}} = 80$ ms.

simulations and the arguments given in appendix B show that the decay can be considered to be a Poisson process. The decay times have, therefore, an exponential distribution (see figure 6). The reason why the decay is a Poisson process originates in the large difference between the EPSC duration (~100 ms) and the average delayed activity decay time (~10 s). Since the fluctuations in the network last no longer than the EPSC time, the attempts of the system to decay during the long delayed activity time (~10 s), become independent of each other. This implies that the probability of decay during each of the τ_{EPSC} periods is independent of time, which leads to the Poissonian statistics for the decay times.

Since failure to maintain the persistent activity entails the loss of the memory and incorrect performance in the delayed response task, at least in the monkey prefrontal cortex (Funahashi *et al* 1989), one can use the Poisson distribution to interpret some psychophysical data. In some experiments on rats performing the binary delayed matching to position tasks the deterioration of WM is observed as a function of delay time (Cole *et al* 1993). The deterioration was characterized by a 'forgetting' curve, with matching accuracy decreasing from about 100% at zero delay to approximately 70% at 30 s delay. Performance approaches the regime of random guessing with 50% of correct responses in this binary task (figure 7). Our prediction for the shape of the 'forgetting' curve follows from the Poisson distribution of delayed activity decay times:

$$\text{Correct} = (1 + \gamma \exp(-t/\bar{t})) \times 50\%. \tag{15}$$

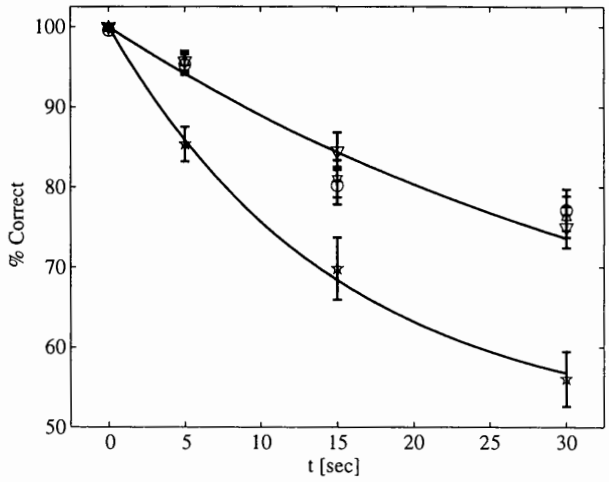

Figure 7. The performance of rats in a delayed matching to position task (from Cole *et al* (1993)). Different markers show various degrees of sedation with the competitive NMDA antagonist CPP: circles 0 mg kg^{-1}; triangles 1 mg kg^{-1}; hexagons 3 mg kg^{-1}; pentagons 10 mg kg^{-1}. The solid curves represent the fits by equation (15). The upper and lower curves correspond to the average decay times of 40 and 15 s, respectively.

Here $0 \leqslant \gamma \leqslant 1$ is the probability that the delayed activity is initiated after presentation of a stimulus. This prediction is used to fit the experimental data in figure 7.

This figure also shows the effect of the competitive NMDA antagonist CPP. Application of the antagonist reduces the average memory retention time $\bar{t}$ from about 40 to 15 s. The presence of non-competitive antagonists impairs performance even at zero delay (Cole *et al* 1993, Pontecorvo *et al* 1991). Non-competitive antagonists have, therefore, an effect on non-WM components of animal behaviour (parameter γ in (15) is less than one). When the delayed component of the 'forgetting' curve is extracted it can be well fitted by an expression (15). This also shows that the average WM storage time $\bar{t}$ decreases with application of an NMDA antagonist.

Our calculation in appendix B shows that the average memory storage time is given by

$$\bar{t} \sim \tau_{\text{EPSC}} \exp[\kappa f \tau_{\text{EPSC}} N^2 (\Delta I/I)^3]. \tag{16}$$

Here ΔI is the distance to the edge of the attraction basin from the stable state and I is the average feedback current. This result holds if $f\tau_{\text{EPSC}} \gg 1$. Because $\bar{t}$ is of the order of tens of seconds and τ_{EPSC} is approximately 100 ms the exponential in (16) is realistically of the order 10^2–10^3.

There are, therefore, two ways in which a synaptic receptor blockade can affect $\bar{t}$. First, the attenuation of the EPSC decreases the average firing frequency f. Second, it moves the system closer to the edge of the attraction basin, reducing ΔI. Both factors increase the effect of noise on the system and decrease the average memory storage time. This is manifested by equation (16).

Another consequence of the formula is the importance of the NMDAR for WM storage. It is based on the large affinity of the receptor to glutamate leading to long EPSCs ($\tau_{\text{NMDA}} \approx$ 100 ms), compared for example to the AMPA receptor ($\tau_{\text{AMPA}} \approx 15$ ms). Equation (16) implies that if the AMPA receptor is used in the bistable neural net and all other parameters (κ, f, N, and $\Delta I/I$) are kept the same, the memory storage time is equal to $\tau_{\text{AMPA}} \approx 15$ ms. Thus it is not surprising that the NMDAR is chosen by evolution as a mediator of WM and the highest

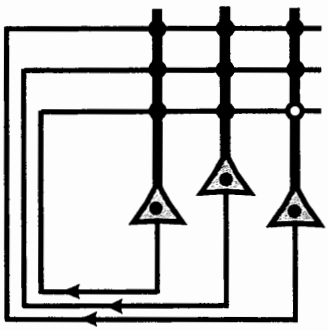

Figure 8. The network used in the text to illustrate the optimum fluctuation leading to a delayed activity decay. The synapses are shown by circles. The synapses releasing the neurotransmitter are shown by full circles. The failing synapse is shown by the open circle. The dendrites and axons are represented by thick and thin lines, respectively.

density of the receptor is observed in places involved in WM storage, i.e. in the prefrontal cortex (for example in frontal and anterior cingulate cortices (a rostral part of the medial prefrontal cortex) in rats: Cotman *et al* (1987); in human prefrontal cortex: Scherzer *et al* (1998)).

We now would like to illustrate what processes lead to the decay time given by (16). Consider a simple network consisting of three neurons (figure 8). Assume that the ratio $\Delta I/I$ for this network is equal to $1/3$. This implies that the neurons have to lose only $1/3$ of their recurrent input current due to a fluctuation to stop firing. This can be accomplished by various means. Our investigation shows that the most effective fluctuation is as follows. Due to the probabilistic nature of synaptic transmission some of the synapses release glutamate when the spike arrives onto the presynaptic terminal (full circles in figure 8) whereas some fail to do so (open circle). It is easy to see that if the same synapse fails to release a neurotransmitter in response to any spike arriving during the time interval τ_{EPSC} the reverberations of the current terminate. Indeed, the failing synapse (open circle) deprives the neuron of $1/3$ of its current. This is just enough to put the input current into the neuron below the threshold and the neuron, therefore, stops firing. This deprives the entire network, consisting of three neurons, of $1/3$ of its feedback current and, therefore, the delayed activity in this network is terminated.

Since in this example only one out of nine synapses fail, the network current $\bar{I}$ (the mean-field current) experiences a fluctuation equal to $\bar{I}/9$. In the the most reasonable alternative decay mechanism the mean-field current would reach the edge of the attraction basin, i.e. would be reduced by $\bar{I}/3$. Thus the mechanism proposed above involves a much *smaller* fluctuation in the current than the mean-field one. Such a fluctuation, therefore, can be realized more frequently and the corresponding decay is faster. Quantitatively the difference between the above (non-mean-field) and the mean-field decay rates is manifested in reducing the exponent of the factor containing $\Delta I/I < 1$ in equation (16) from 3 to 2 (see appendix C for the mean-field estimate). This brings about an *increase* in the average storage time. The mean-field mechanism is therefore less destructive than the proposed one and is disregarded in this paper.

The single neuron which initiates the decay in the persistent activity (the rightmost in figure 8) constitutes the critical nucleus mentioned in the introduction. In general the critical nucleus involves $n = N\Delta I/I$ neurons.

Since an increase in the number of neurons in the network N dramatically influences the memory storage time according to equation (16), another characteristic of the reliability of the WM circuit is the minimum number of neurons N^* necessary to store one bit of information for a $\bar{t}$. We determine this quantity as follows. We run the network simulation many times with the same values of N and τ_{EPSC}. We then determine the average decay time $\bar{t}$. Having done this we decrease or increase N depending on whether $\bar{t}$ is larger or smaller than a given value (20 s in all our simulations). This process converges to the number of neurons necessary to sustain the delayed activity N^* for the given value of τ_{EPSC}. The process is then repeated for different

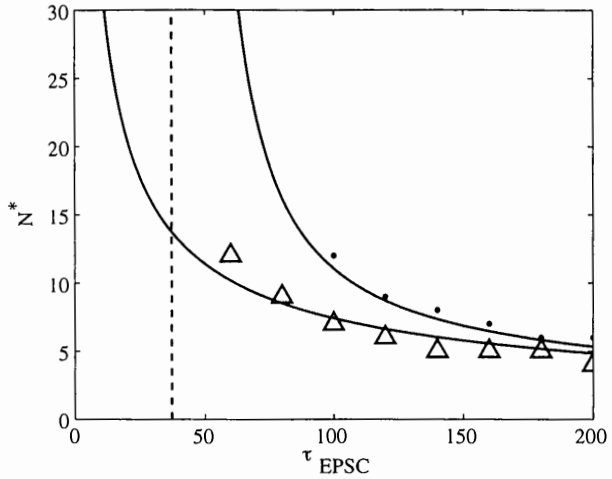

Figure 9. The minimum number of neurons N^* needed to maintain delayed activity with an average decay time of 20 s for $\bar{f} \approx 16$ Hz, $\Delta I/I = 2/3$. The dots show computational results for the case of no noise in the external inputs. The triangles correspond to the case of 10% white noise added to the external current. The solid curves represent the result of the analytical calculations described in the text. The dashed line is the vertical asymptote of N^* in the case of no external noise. No delayed activity can exist to the left of this line, i.e. for the synaptic receptor time constant < 37 ms.

values of the synaptic time-constant. However, the network feedback is always renormalized so that the firing frequency stays the same, close to the reasonable value $\bar{f} \approx 16$ Hz. This corresponds to the attractor state shown in figure 2. The resulting dependence of N^* versus τ_{EPSC} is shown in figure 9.

There are two sets of computational results in figure 9. The higher values of N^* are obtained for the network with no noise in the external inputs (dots in figure 9). The only source of noise in such a network is therefore the unreliability of synaptic connections. It appears for this case that *no* delayed activity can exist for τ_{EPSC} below 37 ms. The lower values of N^* are obtained for the case of white noise added to the external input (triangles). The values for the noise component of the current of each neuron are uniformly distributed between $-A$ and A, where the amplitude of noise A is equal to 10% of the total input current and its correlation time is 1 ms. The limiting value of the synaptic time constant for which the delayed activity is not possible is much smaller for this case ($\approx$5 ms).

This result may seem counter-intuitive. Having added the external noise we increased the viability of the high-frequency state, reducing the effects of the internal synaptic noise. This, however, is not so surprising if one takes the neuronal synchrony into account. The synchrony of neuronal firing results in oscillations in the average input current (see inset in figure 2(a)). Such oscillations periodically bring the system closer to the edge of the attraction basin, creating additional opportunities for decay. Thus a synchronous network is less stable than an asynchronous one. External noise attenuates neuronal synchrony, smearing the oscillations of the average input current. Thus the system with noise in the external inputs should have a larger decay time and a smaller critical number of neurons N^*. This idea is discussed quantitatively in this section. This is similar to the stabilization of the mean-field solutions in the networks of pulse-coupled oscillators (Abbot and Van Vreeswijk 1993).

Figure 10 shows the results of similar calculations for the attractor state with a higher average firing frequency ($\bar{f} = 28$ Hz). This network shows a higher level of reliability and

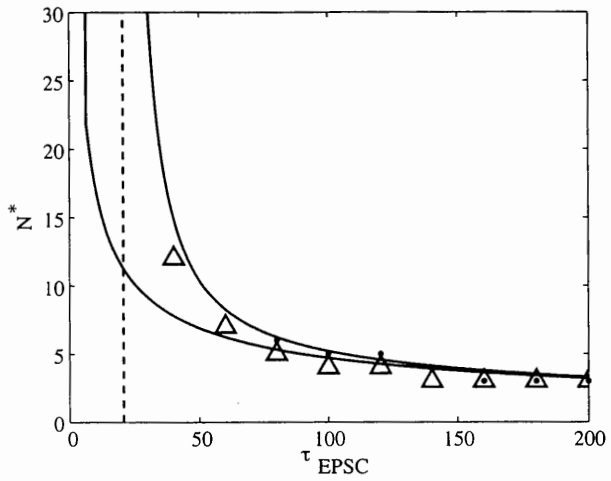

Figure 10. The minimum number of neurons N^* needed to maintain delayed activity with an average decay time of 20 s for the attractor with an average firing frequency of 28 Hz ($\Delta I/I = 5/6$). Notations are the same as in figure 9. The cut-off value of the synaptic time constant is 27 ms. The numerical results with external noise in this case are slightly above those predicted due to insufficient suppression of the synchrony.

smaller values of N^* for both synchronous and asynchronous (10% noise added) regimes. This is consistent with equation (16). The results of the analytical calculations (solid curves) show good agreement with the computer modelling (markers). The rest of the section is dedicated to the discussion of different aspects of the analytical calculations and their results. A more complete treatment of the problem can be found in appendix B.

To evaluate the minimum number of neurons in the closed hand form we solve equation (16) for N

$$N^* \approx N_1^* \equiv \frac{1}{\sqrt{\kappa f \tau_{\text{EPSC}}}} \left(\frac{I}{\Delta I}\right)^{3/2} \sqrt{\ln\left(\frac{\bar{t}}{\tau_{\text{EPSC}}}\right)}. \tag{17}$$

This equation is valid for $f\tau_{\text{EPSC}} \gg 1$. In the opposite case $f\tau_{\text{EPSC}} \lesssim 1$ it has to be amended. To obtain the correct expression in the latter case the following considerations should be taken into account.

3.1. Fluctuations of the average mean-field current

It is obvious that when the mean-field attractor becomes less and less stable in the local sense, i.e. the dimensionless feedback coefficient $\nu < 1$ (equation (14)) approaches unity, the global stability should also suffer. Indeed, if the system is weakly locally stable the fluctuations in the average current are large. This should facilitate global instability. The facilitation can be accounted for by noticing that the transition from the high-frequency state B to the low frequency one A is most probable when the average current is low. Hence to obtain the most realistic transition probability and the correct values for $\bar{t}$ and N^* one has to decrease the values of the average current and distance to the edge of the attraction basin by

$$\begin{aligned} I &\to I - \delta I, \\ \Delta I &\to \Delta I - \delta I, \end{aligned} \tag{18}$$

where the standard deviation of the average current δI is calculated in appendix B

$$\delta I = \frac{I}{N\sqrt{2(1-\nu)\kappa f \tau_{\text{EPSC}}}}. \tag{19}$$

Here I is the average current before the shift, $\nu \approx 0.8$ for the 16 Hz state. This correction decreases the ratio $\Delta I/I$, decreasing $\bar{t}$ (see equation (16)). This implies that there is reduced reliability of the network due to the average current fluctuations.

3.2. Large combinatorial space covered by small groups of neurons attempting to cross the attraction basin edge

In the simple example with three neurons, in principle, each of them can contribute to the decay process. If the network is larger the number of potentially dangerous groups grows as a binomial coefficient C_N^n, where $n = N\Delta I/I \leqslant N$ is the size of the dangerous group. The correction to N^* can be expressed as follows (appendix B):

$$N^* = \sqrt{(N_1^*)^2 + (N_2^*)^2} + N_2^*. \tag{20}$$

Here N_2^* is the combinatorial correction

$$N_2^* = \left(\frac{I}{\Delta I}\right)^3 \frac{\phi(\Delta I/I) + (\Delta I/I)\ln A}{\kappa f \tau_{\text{EPSC}}}. \tag{21}$$

Here

$$A = \frac{I}{\Delta I\sqrt{4\pi N^* \kappa f \tau_{\text{EPSC}}}}, \tag{22}$$
$$\phi(x) = x\ln[1/x] + (1-x)\ln[1/(1-x)].$$

Since the combinatorial contribution N_2^* is proportional to $1/f\tau_{\text{EPSC}}$ it is negligible compared to $N_1^* \propto 1/\sqrt{f\tau_{\text{EPSC}}}$ for large values of τ_{EPSC}. Therefore $N^* \approx N_1^*$ as claimed by equation (17) for this limit. On the other hand if τ_{EPSC} is small the main contribution to the critical number of neurons N^* comes from N_2^*.

The right-hand sides of equations (17) and (21) show a dependence on N^* through the shift in distance to the attraction edge δI and the coefficient A. This dependence is weak, however, since the former represents a very small correction and the latter depends on N^* only logarithmically. Nevertheless to generate a numerically precise prediction we iterate equations (17), (20), and (17) until a consistent value of N^* is reached. The results are shown by lower solid curves in figures 9 and 10. They are in good agreement with numerical simulations in which the synchrony of firing was suppressed by external noise.

3.3. Neuronal synchrony

The synchronization of neuronal firing can be critical for the stability of the high-frequency attractor. Our computer modelling shows that in the large network the neuronal firing pattern is highly synchronized to the order of $T = 1/f$, if the external noise is absent (for analysis of the fully synchronous state see Gerstner (1995) and Gerstner *et al* (1996)). In essence all neurons fire simultaneously and periodically with period T. During one such 'collective spike' they produce a large feedback current. Such a current can be above the firing threshold. However, by the time the next spike is generated if the duration of EPSC is shorter than T the neurons experience almost *no* feedback current. The sum of the external and the feedback current in this case is close to just the external current, i.e. is below the threshold I^* (see section 2). The delayed activity, therefore, cannot be maintained if EPSC is too short, i.e. for $f\tau_{\text{EPSC}} \ll 1$

(Wang 1999). This is consistent with the results of the computer simulations presented in figures 9 and 10.

If τ_{EPSC} is comparable or larger than T synchrony facilitates the instability by producing oscillations in the average current and thus brings the edge of the attraction basin closer. It therefore effectively decreases the distance to the edge of the attraction basin ΔI. The magnitude of this decrease is estimated in appendix B to be

$$\delta I_{syn} \approx \frac{2I}{\pi f \tau_{EPSC}} \frac{f\tau_{EPSC} - e^{-1/f\tau_{EPSC}}(f\tau_{EPSC} + 1)}{1 - e^{-1/f\tau_{EPSC}}} \tag{23}$$

and

$$\Delta I \to \Delta I - \delta I_{syn}. \tag{24}$$

Here I is the average current given by (18). This correction, which is the amplitude of the oscillations of the average current due to synchrony, is in good agreement with computer modelling (see the insert in figure 3(a)). The synchrony does not change the average current in the network significantly. Therefore the latter should not be shifted as in (18).

These corrections imply that if the original value of ΔI is equal to the sum $\delta I + \delta I_{syn}$ (see equation (19)), the effective distance to the edge of attraction basin vanishes. Delayed activity cannot be sustained under such conditions. This occurs at small values of τ_{EPSC} and determines the positions of the vertical asymptotes (dashed lines) in figures 9 and 10. This gives a quantitative meaning to the argument of the impossibility of stable delayed activity in a synchronous network at the small values of the synaptic time constant given above (see also Wang (1999)). When the synchrony is diminished by external noise, the cut-off value of τ_{EPSC} is determined only by $\Delta I = \delta I$ and is therefore much smaller. It is about 5 ms in our computer simulations.

Synchrony also affects the values of the critical number of neurons for large and small $f\tau_{EPSC}$ (N_1^* and N_2^*, respectively)

$$N_{1syn}^* = \left(\frac{I}{\Delta I}\right)^{3/2} \frac{1}{f\tau_{EPSC}} \sqrt{\frac{2(1-\kappa)\ln(t/\tau_{EPSC})}{\kappa(e^{2/f\tau_{EPSC}} - 1)}}, \tag{25}$$

$$N_{2syn}^* = \left(\frac{I}{\Delta I}\right)^3 \frac{\phi(\Delta I/I) + (\Delta I/I)\ln A}{e^{2/f\tau_{EPSC}} - 1} \frac{1-\kappa}{\kappa f^2 \tau_{EPSC}^2}. \tag{26}$$

These equations are derived in appendix B. To obtain the value of critical number of neurons N^* equation (20) should be used. Let us compare the latter equations to (17) and (21). For the limit $f \gg 1$ the expressions for N_1^* and N_{1syn}^* converge to the same asymptote $N_1^* \approx N_{1syn}^*$, whereas both N_2^* and N_{2syn}^* go to zero.

Since again, as in the asynchronous case, both N_{1syn}^* and N_{2syn}^* depend on the quantity that we are looking for N^* (however, again very weakly), an iterative procedure should be used to determine the value of N^*. Having applied this iterative procedure we obtain the upper solid curves in figures 9 and 10 and, therefore, we obtain excellent agreement with the results obtained numerically. In addition we calculate the cut-off τ_{EPSC}, below which the delayed activity is impossible in the synchronous case. For the attractor with $\bar{f} \approx 16$ Hz the cut-off value is 37 ms, while for the higher frequency attractor ($\bar{f} \approx 28$ Hz) the value is 21 ms. These values are shown in figures 9 and 10 by dashed lines. We therefore conclude that the delayed activity mediated by the AMPA receptor is impossible in the synchronous case.

4. Discussion

4.1. The cooperative character of the decay of the high-frequency attractor

The subject of this paper is related to the question of stability of the high-frequency attractor. This state is locally stable, i.e. small synaptic and other noises cannot kick the system out of the basin of attraction surrounding the state. However, this state is not guaranteed to be globally stable. After a certain period of time large fluctuations in the synaptic currents makes the system reach the edge of the attraction basin and the persistent memory state decays into the low-frequency state. Before the network reaches the edge of the attraction basin it performs an excursion into the region of parameters which is rarely visited under usual circumstances. This justifies the use of term 'instanton' for such an excursion, emphasizing the analogy with the particle travelling in the classically forbidden region in quantum mechanics. Such an analogy has been used before in application to the perception of ambiguous stimuli by Bialek and DeWeese (1995). The present study, however, is based on the microscopic picture, deriving the decay times from synaptic properties.

The use of a very simple model network allowed us to look into the nature of the global instability of delayed activity. The principal result of this paper is that the unreliability of synaptic conductance provides the most effective channel for the delayed activity decay. We propose the optimum fluctuation of the synaptic conductances leading to the loss of WM. The decay rates due to such a fluctuation, calculated analytically, agree well with the results of a numerical study.

In the most trivial scenario the network shuts itself down by not releasing the neurotransmitter in a certain fraction of synapses of *all* neurons. In such a scenario all the neurons cross the border of the attraction basin of the high-frequency state simultaneously. Such a mechanism of decay was shown to be ineffective. Instead, the network chooses to cross the border of the attraction basin by shutting the activity down in a *small subpopulation* of neurons, taking advantage of the large combinatorial space spanned by such groups. The effective mechanism of decay, therefore, involves the formation of a *critical nucleus* in the neuronal population. Once the nucleus is formed decay is unavoidable. The cooperative nature of decay in this problem is analogous to some examples of tunnelling of macroscopic objects in condensed matter physics (Larkin and Lee 1978, Levitov *et al* 1995).

4.2. NMDAR and reliability

In this paper we derive the relationship between the dynamic properties of the synaptic receptor channels and the stability of the delayed activity. We conclude that the decay of the latter is a Poisson process with the average decay time depending exponentially on the EPSC time constant. Our quantitative conclusion applied to AMPA receptors, which have a short EPSC, implies that it is not possible to sustain persistent activity in the synchronization of firing in a network. In the case of asynchronous networks one needs a large number of neurons >30 to store one bit of information with AMPA receptors. On the other hand NMDARs seem to avoid such problems, providing reliable bit of information storage with about 15 neurons for both synchronized and asynchronous cases. We therefore suggest an explanation for the experimentally observed high significance of the NMDAR for WM.

In addition to the general evolutionary importance of the NMDA channel for WM, we have a suggestion for the mechanism of the deterioration of WM in the NMDA channel blocking experiments. The blockade of the NMDA channel decreases the feedback network current and the average firing frequency. According to our analytical calculation (see equation (16)) both of these effects decrease the average WM storage time, thus impairing WM. We also predict

that with normal NMDARs the number of neurons able to store one bit of information for 20 s is about 15. Should our theory be applicable to rats performing delayed matching to position task (Cole *et al* 1993) the conclusion would be that the recurrent circuit responsible for this task contains more than 15 neurons. Of course this would imply that other sources of loss of memory, such as distraction (Camperi and Wang 1998), are not present.

One can conclude from our study that if the time constant of the NMDA channel EPSC is further increased, WM can be stored for a much longer time. Assume that τ_{EPSC} is increased by a factor of 2, for instance, by genetic enhancement (Tang *et al* 1999). If the network connectivity, firing frequency, and average currents stay the same as in the wild type, the exponential in equation (16) is increased by a factor of 10^3. This implies that WM can be stored for days instead of minutes. Alternatively the number of neurons responsible for the storage of quantum of information can be decreased by a factor of 0.7 keeping the storage time the same. This implies a higher storage capacity of the brain of genetically modified animals.

4.3. The critical size of the network

The natural question is whether our estimate $N = 15$ is relevant to the cortical circuits containing a large number of neurons. By using a network containing for example $N = 1000$ neurons the instability to noise can be made completely irrelevant. First, the decay of delayed activity is indeed observed in both psychophysically and physiologically (see Cole *et al* (1993), Funahashi *et al* (1989)). The system therefore is close to the maximum performance that can be reached by existing neural circuits. Hence noise is relevant in the existing conditions. Second, our estimate for the critical number of neurons increases if a sparsely connected network is employed. If a neuron receives inputs only from ϵN neurons involved in the storage of the same bit, the critical number of neurons should be increased by a factor $1/\sqrt{\epsilon}$. If $\epsilon \approx 0.1$, as in Brunel (2000), the critical value of N can reach 40–50. Inhibitory neurons amplify the relative noise in the input current, increasing the critical number of neurons too. We conclude therefore that a greater understanding of the architecture of the recurrent circuits is needed to make an exact claim how close the actual number of neurons is to the minimum value evaluated in the presented study. We can, however, make a comparison to the cortical sensory areas, where more experimental evidence is available. The study of correlations between monkey psychophysical performance and responses of a single neuron in area MT (Shadlen *et al* 1996) reveals that at least 100 MT neurons are engaged in the directional discrimination task. Since direction can be coded by three to five bits, this result implies 20–30 neurons/bit. On the other hand responses of neurons in temporal visual cortex to faces (Rolls *et al* 1997) show that 14 neurons code 2.77 bits of information on average. This corresponds to about 5 neurons/bit. If any reciprocity is maintained between sensory and mnemonic areas, a similar number of neurons coding for one bit should be observed in the prefrontal cortex. This makes our estimate of $N = 15$ more realistic.

Since noise due to synaptic transmission failures is uncorrelated from cell to cell its effect decreases when the size of the network increases. In large networks the uncorrelated component of noise, which may be due to, for example, distractors (Durstewitz *et al* 2000), becomes relevant. The transitions caused by distractors therefore become important when the number of neurons in the network exceed a certain value N_d. Assume that the distribution of input current due to distractors is Gaussian with standard deviation, σ. The critical number of neurons below which our theory takes place N_d is obtained from equation (16)

$$N_d \approx \frac{1}{\sigma}\sqrt{\frac{I^3}{\kappa f \tau_{EPSC} \Delta I}}. \tag{27}$$

Here we ignore any logarithmic factors. Our theory therefore has a region of validity if $N^* \lesssim N_d$ (see equation (17)), i.e. if the distance to the edge of the attraction basin is larger then the standard deviation of the input current due to distractors: $\Delta I > \sigma$. The latter condition seems plausible, since otherwise the persistent activity is destroyed by distractors almost instantaneously. Assuming that our theory has a region of validity, i.e. the latter condition is satisfied, the following three regimes are possible depending on the value of N. (i) $N \lesssim N^*$, the persistent activity is highly unstable with respect to intrinsic noise and is destroyed almost instantaneously (in about 100 ms). (ii) $N^* \lesssim N \lesssim N_d$, our theory is valid, the decay time is given by equation (16), and is of the order of a few minutes. (iii) $N_d \lesssim N$, the intrinsic noise is irrelevant and decay is caused by distractors. More experimental evidence is needed to establish what regime the real cortical networks operate in. It is worth noting, however, that the latter regime seems unlikely, since in this case the number of neurons can be reduced at least to N_d without any loss in reliability. Our theory may therefore be relevant in establishing the evolutionary factors determining the architecture of such circuits.

4.4. The impact of synchrony

Our study shows that in the absence of noise the firing of cells in recurrent circuits is strongly synchronized both with fast AMPAR and slow NMDAR. Thus the feedback current experiences oscillations at the firing frequency shown in figure 3(*a*) and described by equation (B.14). This is partially in agreement with conclusions of Wang (1999), who suggests that a fast AMPA current may lead to synchronous firing. In addition, we suggest that even a slow NMDAR mediated feedback current may lead to synchrony (as in figure 3(*a*)). However, the impact of synchrony is drastically different in the cases of AMPAR and NMDAR mediating the feedback. In the former case the oscillations of the current, given by equation (B.14), are very large and the feedback current is virtually unimportant. Thus the synchronized high-frequency state in the case of AMPAR is *locally* unstable. In the case of NMDAR, however, the synchronous state can be made to be locally stable. The instability of the synchronous high-frequency state in this case is much softer and is reflected by the reduction of its lifetime with respect to the asynchronous state. This is in line with the conclusion of Wang (1999), suggesting that slow NMDAR channels may be required for the maintenance of a persistent state in the case of synchronous firing.

We also studied the decay process in the presence of uncorrelated noise in the afferent inputs. The study suggests that the effect of such a noise on WM storage reliability is not monotonic. The addition of small white noise (<10% of the total external current) increases the reliability by destroying synchronization and is therefore beneficial for WM storage. A further increase in noise (>12%) destroys WM, producing transitions between low- and high-frequency states. We conclude therefore that there is an optimum amount of afferent noise, which on one hand smoothes the synchrony and on the other hand does not produce a decay in delayed activity.

If the afferent noise is not too large the neurons fire in synchrony. The natural consequences of synchrony are a decrease in the coefficient of variation of the interspike interval for single neurons and the increase in the crosscorrelations between neurons. The latter prediction is consistent with the findings of some multielectrode studies in monkey prefrontal cortex (see Dudkin *et al* (1997a), Funahashi (1998)). On the other hand synchrony may also be relevant to the phenomenon of temporal binding in visual cortical areas (Engel *et al* 1999, Roskies 1999). We argue therefore that temporal binding with a precision of many milliseconds can be accomplished by recurrent neural networks.

4.5. The high firing rate problem

We suggest a solution to the high firing frequency problem by using a more precise model of the spike generation mechanism, i.e. the leaky integrator model with time varying resting potential and integration time. The minimum firing frequency for the recurrent network based on such a model is determined by the rate of variation of the potential and time-constant and is within the range of physiologically observed values. Further experimental work is needed to see if the model is applicable to other types of neurons, such as inhibitory cells.

4.6. Conclusion

In conclusion we have studied the stability of delayed activity in recurrent neural network subjected to the influence of noise. We conclude that the global stability of the persistent activity is affected by properties of the synaptic receptor channel. The NMDA channel, having a long EPSC duration time, is a reliable mediator of the delayed response. On the other hand the AMPA receptor is much less reliable, and in the case of synchronized firing cannot be used to sustain response. The effect of the NMDA channel blockade on the WM task performance is discussed.

Acknowledgments

The author is grateful to Thomas Albright, Michael DeWeese, Jean-Mark Fellous, Paul Tiesinga, and Tony Zador for discussions and numerous helpful suggestions. This paper was supported by the Alfred P Sloan Foundation.

Appendix A. Single-neuron transduction function

In this appendix we calculate the single-neuron transduction function. The first step is to find the membrane voltage as a function of time. From equation (1) using the variation of integration constant we obtain

$$V(t) = V_{\text{reset}}\mathrm{e}^{-S(t)} + \int_{t_0}^{t} \mathrm{d}t'\, \mathrm{e}^{S(t')-S(t)}[I(t') + E(t')/\tau(t')]. \tag{A.1}$$

Here

$$S(t) = \int_{t_0}^{t} \mathrm{d}t'/\tau(t'), \tag{A.2}$$

and we introduced the refractory period both for generality and to resolve the peculiarity at $t = 0$. For constant current and functions given by (2) and (3) the expression for the voltage can be further simplified

$$V(t) = V_{\text{reset}}\mathrm{e}^{-S(t)} + E_0[1 - \mathrm{e}^{-S(t)}] + \left(I - \frac{\Delta E}{\tau_0}\right)\frac{\tau_0}{\alpha}\frac{L_{1/\alpha}(\mathrm{e}^{\alpha t/\tau_0})\mathrm{e}^{-S(t)}}{(\mathrm{e}^{\alpha t_0/\tau_0} - 1)^{1/\alpha}}, \tag{A.3}$$

where

$$S(t) = \frac{1}{\alpha}\ln\left(\frac{\mathrm{e}^{\alpha t/\tau_0} - 1}{\mathrm{e}^{\alpha t_0/\tau_0} - 1}\right) \tag{A.4}$$

and $L_n(x)$ is defined for an integer n by

$$L_0(x) = \ln(x), \qquad L_n(x) = \frac{1}{n}(x-1)^n - L_{n-1}(x), \tag{A.5}$$

and is obtained for fractional n by analytical continuation. For example

$$\begin{aligned} L_1(x) &= x - 1 - \ln(x), \\ L_2(x) &= (x-1)^2/2 - x + 1 - \ln(x), \\ L_3(x) &= (x-1)^3/3 - (x-1)^2/2 + x - 1 - \ln(x), \\ &\cdots. \end{aligned} \tag{A.6}$$

Solving the equation $V(t) = \theta$ produces the interspike interval t and frequency $f = 1/t$. The solution cannot be done in the closed form but some asymptotes can be calculated, however. The calculation depends on the value of the parameter $\xi = \alpha t/\tau_0$.

(i) $\xi = \alpha t/\tau_0 \ll 1$. For this limit

$$e^{-S(t)} \approx \frac{(e^{\alpha t_0/\tau_0} - 1)^{1/\alpha}}{\xi + \xi^2/2}, \tag{A.7}$$

is very small and can be neglected everywhere, except when multiplied by the large factor $L_{1/\alpha}$, this gives

$$\frac{L_{1/\alpha}(e^{\alpha t/\tau_0})e^{-S(t)}}{(e^{\alpha t_0/\tau_0} - 1)^{1/\alpha}} \approx \frac{\alpha}{1+\alpha}\left(\xi + \frac{\xi^2}{2}\right) - \frac{\alpha}{1+2\alpha}\xi^2. \tag{A.8}$$

The equation on the interspike interval is

$$\theta = E_0 + (\tau_0 I - \Delta E)\left[\frac{\xi}{1+\alpha} - \frac{\xi^2}{2}\frac{1}{(1+\alpha)(1+2\alpha)}\right]. \tag{A.9}$$

Solving this quadratic equation with respect to ξ we obtain the asymptote given by equation (4).

(ii) $\xi \gg 1$. In this case the solution can be found directly from equation (1) by assuming $\dot{V} = 0$. Solving the resulting algebraic equation for t we obtain equation (7).

Appendix B. Decay of the high-frequency attractor

B.1. Asynchronous case: derivation of N_1^* and N_2^*

As shown in the main text the decay of the high-frequency state occurs through the formation of the critical nucleus of size $n = N\Delta I/I$ and involves therefore a small subpopulation of the entire network. Hence, the reduction in the average current in the network during the decay is small. For the asymptotic limit $\Delta I \ll I$ one can neglect the variation of the average network current when calculating the probability of the formation of the critical nucleus. Once it is formed decay is unavoidable. Thus the transition probability associated with this mechanism is independent of the network dynamics and has a universal form.

We first calculate the probability distribution of the input current for a single neuron. When doing so we can neglect the variations in the average network current as indicated above. The average and standard deviation of the input current of each neuron, given by equations (9) and (10) are

$$\bar{I} = N\kappa f \tau_{\text{EPSC}} j_0 \tag{B.1}$$

and

$$\overline{\delta I_n^2} = \overline{(I_n - \bar{I})^2} = N\kappa f \tau_{\text{EPSC}} j_0^2/2. \tag{B.2}$$

The distribution of I_n according to the central limit theorem is therefore

$$p(I_n) = \frac{1}{\sqrt{2\pi\overline{\delta I_n^2}}} \exp\left[-\frac{(I_n - \bar{I})^2}{2\overline{\delta I_n^2}}\right]. \tag{B.3}$$

The probability that the input current is below the threshold I^*, i.e. the neuron does not fire, is given by the error function derived from distribution (B.3)

$$P = \int_{-\infty}^{I^*} \mathrm{d}I_n \, p(I_n) \approx A \mathrm{e}^{-\kappa f \tau_{\mathrm{EPSC}} N \Delta I^2 / \bar{I}^2}, \tag{B.4}$$

where $A = \bar{I}/\Delta I \sqrt{4\pi\kappa f \tau_{\mathrm{EPSC}} N}$, and $\Delta I = \bar{I} - I^*$. In the derivation of (B.4) we used the asymptotic expression for the error function

$$\int_{-\infty}^{y} \mathrm{d}x \, \frac{\mathrm{e}^{-x^2/2\sigma^2}}{\sigma\sqrt{2\pi}} \approx \frac{\sigma}{|y|\sqrt{2\pi}} \mathrm{e}^{-y^2/2\sigma^2}, \tag{B.5}$$

when $y \ll -\sigma$.

The reverberations of the current will be impossible if $n = N\Delta I/\bar{I}$ neurons are below the threshold. When this occurs the feedback current to each neuron is reduced with respect to the average current by ΔI, i.e. is below the threshold, and further delayed activity is impossible. The probability of this event is given by the binomial distribution

$$p_0 \sim C_N^n P^n. \tag{B.6}$$

Here the binomial coefficient $C_N^n = N!/(N-n)!n!$ accounts for the large number of groups of neurons that can contribute to the decay.

Since the input currents stay approximately constant during time interval τ_{EPSC}, we break the time axis into windows with duration $\Delta t \sim \tau_{\mathrm{EPSC}}$ and denote by p the average probability of decay of the high-frequency activity during such a little window. Assuming that k windows have been passed since the delayed activity commenced. The probability that the decay occurs during the kth time window, i.e. exactly between $t = k\Delta t$ and $t = (k+1)\Delta t$, is

$$\rho(t)\Delta t = (1-p)^k p. \tag{B.7}$$

Here $(1-p)^k$ is the probability that the decay did not happen during either of k early time intervals. After simple manipulations with this expression we obtain for the density of probability of decay as a function of time

$$\rho(t) = \mathrm{e}^{-pt/\Delta t}/p\Delta t. \tag{B.8}$$

This is a Poisson distribution and decay is a Poisson process. The reason for this is that the system retains the values of input currents during the time interval $\Delta t \sim \tau_{\mathrm{EPSC}} \sim 100$ ms. Therefore all processes separated by longer times are independent. Since the decay time is of the order of 10–100 s, the attempts of the system to decay at various times can be considered independent. We finally notice that since the values of the current are preserved on the scales of τ_{EPSC} we can conclude that

$$p \sim p_0 \tag{B.9}$$

given by equation (B.6). The average decay time can then be estimated using (B.8):

$$\frac{\bar{t}}{\tau_{\mathrm{EPSC}}} \sim p_0^{-1}. \tag{B.10}$$

In the limit $f\tau_{\mathrm{EPSC}} \gg 1$ the minimum size of the network necessary to maintain the delayed activity is small. We can therefore assume N to be small for this limit. The same assumption can be made for $n < N$. Thus for an effective network (using the neurons sparingly) we can assume the combinatorial term in equation (B.6) to be close to unity. Since in the expression for $p_0 \sim C_N^n A^n \exp(-n\kappa f \tau_{\mathrm{EPSC}} N \Delta I^2/\bar{I}^2)$ it is also compensated by the small prefactor of the exponential $A^n < 1$ (see equation (B.4)), we can assume $C_N^n A^n \sim 1$ for the effective network. Therefore for the limit $f\tau_{\mathrm{EPSC}} \gg 1$ the main dependence of the parameters of the

model is concentrated in the exponential factor. This set of assumptions in combination with equation (B.10) leads to the expression (16) in the main text. If more precision is needed equation (B.10) can be used directly.

To find the critical number of neurons N^* we calculate the logarithm of both sides of equation (B.10) and obtain

$$\ln(\bar{t}/\tau_{\text{EPSC}}) \approx -\ln C_{N^*}^{n} - n \ln p_0. \tag{B.11}$$

Using the asymptotic expression for the logarithm of the binomial coefficient, which can be derived from the Stirling's formula

$$\begin{aligned} &\ln C_N^n \approx N\phi(n/N); \qquad N, n \gg 1; \\ &\phi(x) = x \ln[1/x] + (1-x)\ln[1/(1-x)], \end{aligned} \tag{B.12}$$

we derive the quadratic equation for N^*

$$\left(\frac{\Delta I}{I}\right)^3 \kappa f \tau_{\text{EPSC}} (N^*)^2 - N^* \left[\phi\left(\frac{\Delta I}{I}\right) + \frac{\Delta I}{I} A\right] - \ln \frac{\bar{t}}{\tau_{\text{EPSC}}} \approx 0. \tag{B.13}$$

Solving this equation we obtain the set of equations (17), (20), and (21) in the main text.

B.2. Synchronous firing case

In this section we assume that all the neurons in the network fire simultaneously. The important quantities which will be studied are the time dependence of the averaged over the network current $\bar{I}(t)$ and the fluctuations of the input current into a single neuron $\overline{\delta j_n^2}$. The former is responsible for the shift in the distance to the edge of the attraction basin (equation (24)), the latter determines the average decay time, as follows from the previous section.

Assume that the neuron fired at $t = 0$. The average number of EPSCs arriving to the postsynaptic terminal is κN due to the finite probability of the synaptic vesicle release κ and the all-to-all topology of the connections. The average current at times $0 < t < T = 1/f$ is contributed by the spikes at $t = 0$ and by all previous spikes at times Tk, $k = 1, 2, \ldots$

$$\begin{aligned} \bar{I}(t) &= I_0 \mathrm{e}^{-t/\tau_{\text{EPSC}}} + I_0 \mathrm{e}^{-(t+T)/\tau_{\text{EPSC}}} + \cdots \\ &= I_0 \mathrm{e}^{-t/\tau_{\text{EPSC}}} / (1 - \mathrm{e}^{-1/f\tau_{\text{EPSC}}}), \end{aligned} \tag{B.14}$$

where we introduced the notation $I_0 = \kappa N j_0$. The average over time value of this averaged over neurons current is $\langle \bar{I} \rangle = I_0 f \tau_{\text{EPSC}}$, where by angular brackets we denote the time average: $\langle A \rangle = \int_0^T A(t)\,\mathrm{d}t/T$. The average current $\bar{I}$ therefore experiences oscillations between $\bar{I}_{\max} = \bar{I}(0)$ and $\bar{I}_{\min} = \bar{I}(T)$. The minimum value of $\bar{I}$ is below the average current $\langle \bar{I} \rangle$. This minimum value according to the logic presented in the main text, determines the shift of the distance to the edge of the attraction basin. This shift is therefore $\langle \bar{I} \rangle - \bar{I}_{\min}$. Our computer modelling, however, shows that the actual amplitude of the current oscillations is consistently about 60% of the value predicted by this argument. This was tested at various values of parameters. The explanation of this 60% factor is as follows. In the case where all the neurons fire simultaneously the shape of the dependence of the average current versus time is saw-like. However, the spikes do not fire absolutely simultaneously. The uncertainty in the spiking time is of the order of T. This uncertainty, which is intrinsic to the recurrent synaptic noises, and therefore difficult to calculate exactly, smears out the saw-like dependence of the average current of time (see inset in figure 3(*a*)). This smearing can approximately be accounted for by dumping the higher harmonics of the saw-like dependence. When only the principal harmonic remains, the amplitude of the saw-like curve is reduced by a factor

$2/\pi \approx 0.6$. This is consistent with the numerical result. We therefore conclude that the shift of ΔI is

$$\delta I_{\text{syn}} \approx 0.6[\langle \bar{I} \rangle - \bar{I}(T)] \tag{B.15}$$

which leads to equation (23) in the text.

We now turn to the calculation of the standard deviation of the input current into one neuron. Similar to (B.14), using the central limit theorem we obtain

$$\delta I_n^2(t) = \sum_{k=0}^{\infty} j_0^2 \kappa(1-\kappa) N \mathrm{e}^{-(2t+2Tk)/\tau_{\text{EPSC}}}. \tag{B.16}$$

This quantity is most important at $t \approx T$ when the average current reaches its minimum and the neuron stops firing with maximum probability. Performing the summation we therefore obtain

$$\delta I_n^2(T) = \frac{\kappa(1-\kappa) I^2}{(\kappa f \tau_{\text{EPSC}})^2 N} \frac{1}{\mathrm{e}^{2/f\tau_{\text{EPSC}}} - 1}. \tag{B.17}$$

Substituting this value into equation (B.4) in the previous section and repeating the subsequent derivation we obtain equations (25) and (26) in the main text.

B.3. Fluctuations of the average current

In this subsection we calculate the fluctuations of the average network current. We consider the asynchronous case for simplicity. The conclusions are perfectly good for the synchronous case, for the reasons that will become clear later in this section, and agree well with computer simulations.

The average current satisfies the linearized equation similar to the linearized version of equation (12)

$$\tau_{\text{EPSC}} \Delta \dot{\bar{I}}(t) = (\nu - 1)\Delta \bar{I}(t) + \xi(t). \tag{B.18}$$

Here $\Delta \bar{I}(t)$ is the deviation of the average network current from the equilibrium value and $\xi(t)$ is the noise. The unitless network feedback coefficient $\nu < 1$ is defined by (14). As evident from this equation $\Delta \bar{I}(t)$ has a slow time constant $\tau_{\text{EPSC}}/(1-\nu)$ (since $1/(1-\nu) > 5$ in our simulations). On the other hand the noise $\xi(t)$ is determined by synapses and has a correlation time that is relatively small ($\sim \tau_{\text{EPSC}}$).

The correlation function of noise in the average current can be found from equation (9):

$$\Xi(t) \equiv \langle \xi(t)\xi(0) \rangle = \frac{\overline{\delta I_n^2}}{N} \exp(-|t|/\tau_{\text{EPSC}}). \tag{B.19}$$

Here angular brackets imply averaging over time and the value of $\Xi(t=0)$ follows from equation (B.2) and the central limit theorem (dispersion of the average is equal to the dispersion of each of the homogeneous constituents I_n divided by the number of elements N). We conclude therefore that

$$\Xi(t=0) = \frac{I^2}{2f\tau_{\text{EPSC}} N^2}. \tag{B.20}$$

The correlation function of $\Delta \bar{I}$ can then be easily found from equation (B.18) using the Fourier transform. With $C(t) \equiv \langle \Delta \bar{I}(t) \Delta \bar{I}(0) \rangle$ then

$$C(\omega) = \langle |\Delta \bar{I}(\omega)|^2 \rangle = \frac{\Xi(\omega)}{(1-\nu)^2 + (\omega \tau_{\text{EPSC}})^2}. \tag{B.21}$$

Since

$$\Xi(\omega) = \frac{2\tau_{\text{EPSC}}\,\Xi(t=0)}{1+(\omega\tau_{\text{EPSC}})^2}, \tag{B.22}$$

the expression for $C(t)$ is readily obtained by inverting the Fourier transform (B.21). In the limit $1-\nu \ll 1$, which holds in our simulations the answer is

$$C(t) = \frac{\Xi(t=0)}{1-\nu}\exp[-(1-\nu)|t|/\tau_{\text{EPSC}}]. \tag{B.23}$$

The value of C taken at $t=0$ determines the standard deviation of the average current (19).

The fluctuations described by the correlation function (B.23) have a large correlation time compared to the firing frequency. We conclude therefore that synchrony should not affect the long range component of the correlation function.

Appendix C. Transitions from the low-frequency state to the high-frequency state

In this appendix we derive the transition rate for the spontaneous 'creation' of delayed activity. We consider a network similar to the one used in the rest of the paper. There are two specifics in this problem: first, the transition occurs from the state where there is virtually no recurrent feedback current. The transition is therefore induced mostly by the noise in the afferent inputs. Second, since the recurrent feedback is negligible, the transition involves the entire network and is therefore described by the mean-field approximation. The universality pertinent to the non-mean-field approach is lost here and transition probabilities can be derived only approximately.

We begin by evaluating the fluctuations of the average network current, which is done similarly to the previous appendix

$$\delta I^2 = \delta I_{\text{R}}^2 + \delta I_{\text{E}}^2. \tag{C.1}$$

Here δI_{R}^2 and δI_{E}^2 are the fluctuations due to recurrent and external noise respectively. The former is given by equations (19) and (B.23):

$$\delta I_{\text{R}}^2 = \frac{I_{\text{R}}^2}{2\kappa_{\text{R}} f_{\text{R}} \tau_{\text{R}} N_{\text{R}}^2 (1-\nu)}, \tag{C.2}$$

with the quantities labelled by subscript R referring to the recurrent connections. The latter fluctuation δI_{E}^2 is calculated similarly to (B.23). When doing this we keep in mind that external currents are most likely transmitted by the AMPA receptor, since the receptor has to be voltage-independent to act in the low-frequency state. Therefore the correlations of noise in the external current are more short-lived: $\Xi_{\text{E}}(t) = \Xi_{\text{E}}(t=0)\exp(-|t|/\tau_{\text{E}})$, with $\Xi_{\text{E}}(t=0) = I_{\text{E}}^2/2 f_{\text{E}}\tau_{\text{E}}N_{\text{E}}^2$. Here the subscript 'E' applies to quantities describing the external current with $\tau_{\text{E}} \sim 10$ ms. Repeating the derivation from the previous appendix we obtain

$$\delta I_{\text{E}}^2 = \frac{I_{\text{E}}^2}{2\kappa_{\text{E}} f_{\text{E}} \tau_{\text{R}} N_{\text{E}}^2 (1-\nu)}. \tag{C.3}$$

Here we assume that the time constant of the recurrent connections is large: $\tau_{\text{R}} \sim 100$ ms $\gg$ $\tau_{\text{E}} \sim 10$ ms.

Let us compare the fluctuations due to the external and recurrent currents. We first do the estimate for the high-frequency state. Assuming that $\kappa_{\text{E}} \sim \kappa_{\text{R}}$, $f_{\text{E}} \sim f_{\text{R}}$, $N_{\text{E}} \sim 10^4$, $N_{\text{R}} \sim 10^2$, and $I_{\text{R}} \sim 0.1 I_{\text{E}}$ we obtain $\delta I_{\text{E}}/\delta I_{\text{R}} \sim 0.1$. Thus fluctuations in the external current are smaller than in the recurrent one. If however the recurrent current is reduced by an order of magnitude, like in the low-frequency state, the fluctuations are of similar strengths. We therefore keep both terms in the expression for the rate of transitions from the low-frequency state.

The mean life-time of the low-frequency state is now given by

$$\bar{t} \sim \tau_R / p_0, \tag{C.4}$$

where p_0 is the probability that the average network current reaches the threshold, derived in a similar way to the expressions in the previous appendix

$$p_0 = \int_{I^*}^{\infty} dI\ p(I) \sim \frac{\delta I}{\Delta I} e^{-\Delta I^2/2\delta I^2}. \tag{C.5}$$

Here ΔI is the distance from the low-frequency state to the edge of the attraction basin I^*. The value of the dispersion of the input current is given by equation (C.1). The latter expression has an approximate character, since the tails of the distribution of the average current around the stable state can be affected by the variation of the firing rate.

If the recurrent currents in the low-frequency state are negligible, then a simple estimate can be found for the transition time

$$\bar{t} \sim \tau_R e^{(1-\nu)\kappa_E f_E \tau_R N_E^2 (\Delta I / I_E)^2}. \tag{C.6}$$

On comparison to the non-mean-field expression (16) we notice that the transition time is increased due to the second power of the factor $\Delta I / I < 1$ in the exponential. This is due to a non-cooperative character of decay of the low-frequency state.

References

Abbot L F and Van Vreeswijk C 1993 Asynchronous states in networks of pulse-coupled oscillators *Phys. Rev.* B **48** 1483–90

Adler C M, Goldberg T E, Malhotra A K, Pickar D and Breier A 1998 Effects of ketamine on thought disorder, working memory, and semantic memory in healthy volunteers *Biol. Psychiatry* **43** 811–6

Amit D J and Brunel N 1997 Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex *Cereb. Cortex* **7** 237–52

Amit D J, Tsodyks M V 1991a Quantitative study of attractor neural network retrieving at low spike rates: I. Substrate-spikes, rates and neuronal gain *Network: Comput. Neural Syst.* **2** 259–73

——1991b Quantitative study of attractor neural network retrieving at low spike rates: II. Low-rate retrieval in symmetric networks *Network: Comput. Neural Syst.* **2** 275–94

Aura J and Riekkinen P Jr 1999 Blockade of NMDA receptors located at the dorsomedial prefrontal cortex impairs spatial working memory in rats *Neuroreport* **10** 243–8

Bialek W and DeWeese M 1995 Random switching and optimal processing in the perception of ambiguous signals *Phys. Rev. Lett.* **74** 3077–80

Brunel N 2000 Dynamics of sparsely connected networks of excitatory and inhibitory neurons *J. Comput. Neurosci.* **8** 183–208

Camperi M and Wang X J 1998 A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability *J. Comput. Neurosci.* **5** 383–405

Clements J D, Lester R A, Tong G, Jahr C E and Westbrook G L 1992 The time course of glutamate in the synaptic cleft *Science* **258** 1498–501

Cole B J, Klewer M, Jones G H and Stephens D N 1993 Contrasting effects of the competitive NMDA antagonist CPP and the non-competitive NMDA antagonist MK 801 on performance of an operant delayed matching to position task in rats *Psychopharmacology (Berl.)* **111** 465–71

Cotman C W, Monaghan D T, Ottersen O P and Storm-Mathisen J 1987 Anatomical organization of excitatory amino acid receptors and their pathways *Trends Neurosci.* **10** 273–80

Dobrunz L E and Stevens C F 1997 Heterogeneity of release probability, facilitation, and depletion at central synapses *Neuron* **18** 995–1008

Dudkin K N, Kruchinin V K and Chueva I V 1997a Synchronization processes in the mechanisms of short-term memory in monkeys: the involvement of cholinergic and glutaminergic cortical structures *Neurosci. Behav. Physiol.* **27** 303–8

——1997b Effect of NMDA on the activity of cortical glutaminergic structures in delayed visual differentiation in monkeys *Neurosci. Behav. Physiol.* **27** 153–8

Durstewitz D, Seamans J K and Sejnowski T J 2000 Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex *J. Neurophysiol.* **83** 1733–50

Engel A K, Fries P, Konig P, Brecht M and Singer W 1999 Temporal binding, binocular rivalry, and consciousness *Consc. Cogn.* **8** 126–51

Funahashi S 1998 Dynamic changes in interactions among prefrontal neurons in relation to working memory processes *Soc. Neurosci. Abstr.* **24** 1426

Funahashi S, Bruce C J and Goldman-Rakic P S 1989 Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex *J. Neurophysiol.* **61** 331–49

Fuster J M 1995 *Memory in the Cerebral Cortex* (Cambridge, MA: MIT Press)

Gerstner W 1995 Time structure of the activity in neural network models *Phys. Rev.* E **51** 738–58

——2000 Population dynamics of spiking neurons: fast transients, asynchronous states, and locking *Neural Comput.* **12** 43–89

Gerstner W and van Hemmen J L 1992a Associative memory in a network of 'spiking' neurons *Network: Comput. Neural Syst.* **3** 139–64

——1992b Universality in neural networks: the importance of the 'mean firing rate' *Biol. Cybern.* **67** 195–205

——1993 Coherence and incoherence in a globally coupled ensemble of pulse-emitting units *Phys. Rev. Lett.* **71** 312–5

Goldman-Rakic P S 1994 Working memory disfunction in schizophrenia *J. Neuropsychiatry Clin. Neurosci.* **6** 348–57

——1995 Cellular basis of working memory *Neuron* **14** 477–85

Goldman-Rakic P S, Funahashi S and Bruce C J 1990 Neocortical memory circuits *Cold Spring Harbor Symposia on Quantitative Biology* vol 55 (Cold Spring Harbor, NY: Cold Spring Harbor Press) pp 1025–38

Hertz J, Krogh A and Palmer R G 1991 *Introduction to the Theory of Neural Computation* (Redwood City, CA: Addison-Wesley)

Javitt D C, Steinschneider M, Schroeder C E and Arezzo J C 1996 Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia *Proc. Natl Acad. Sci. USA* **93** 11 962–7

Javitt D C and Zukin S R 1991 Recent advances in the phencyclidine model of schizophrenia *Am. J. Psychiatry* **148** 1301–8

Kistler W M, Gerstner W and van Hemmen J L 1997 Reduction of the Hodgkin–Huxley equations to a single-variable threshold model *Neural Comput.* **93** 1015–45

Kogo N and Ariel M 1999 Response attenuation during coincident afferent excitatory inputs *J. Neurophysiol.* **81** 2945–55

Koulakov A A 1999 Instantons in WM: implications for schizophrenia webpage *Los Alamos preprint archive* http://xxx.lanl.gov/abs/cond-mat/9911006

Krystal J H, Karper L P, Seibyl J P, Freeman G K, Delaney R, Bremner J D, Heninger G R, Bowers M B Jr and Charney D S 1994 Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses *Arch. Gen. Psychiatry* **51** 199–214

Larkin A I and Lee P A 1978 Tunneling of solitons and charge-density waves through impurities *Phys. Rev.* B **17** 1596–600

Lester R A, Clements J D, Westbrook G L and Jahr C E 1990 Channel kinetics determine the time course of NMDA receptor-mediated synaptic currents *Nature* **346** 565–7

Levitov L S, Shytov A V and Yakovets A Y 1995 Quantum breaking of elastic string *Phys. Rev. Lett.* **75** 370–3

Lisman J E, Fellous J M and Wang X J 1998 A role for NMDA-receptor channels in working memory *Nature Neurosci.* **1** 273–5

Mainen Z F, Malinow R and Svoboda K 1999 Synaptic calcium transients in single spines indicate that NMDA receptors are not saturated *Nature* **399** 151–5

Miyashita Y 1988 Neuronal correlate of visual associative long-term memory in the primate temporal cortex *Nature* **335** 817–20

Miyashita Y and Chang H S 1988 Neuronal correlate of pictorial short-term memory in the primate temporal cortex *Nature* **331** 68–70

Pontecorvo M J, Clissold D B, White M F and Ferkany J W 1991 N-methyl-D-aspartate antagonists and working memory performance: comparison with the effects of scopolamine, propranolol, diazepam, and phenylisopropyladenosine *Behav. Neurosci.* **105** 521–35

Rolls E T, Treves A and Tovee M J 1997 The representational capacity of the distributed encoding of information provided by populations of neurons in the primate temporal visual cortex *Exp. Brain Res.* **114** 149–62

Roskies A 1999 The binding problem *Neuron* **24** 7–9

Sakai K and Miyashita Y 1991 Neural organization for the long-term memory of paired associates *Nature* **354** 152–55

Scherzer C R, Landwehrmeyer G B, Kerner J A, Counihan T J, Kosinski C M, Standaert D G, Daggett L P, Velicelebi G, Penney J B and Young A B 1998 Expression of N-methyl-D-aspartate receptor subunit mRNAs in the human brain: hippocampus and cortex *J. Comp. Neurol.* **390** 75–90

Shadlen M N, Britten K H, Newsome W T and Movshon J A 1996 A computational analysis of the relationship between neuronal and behavioral responses to visual motion *J. Neurosci.* **16** 1486–510

Stevens C F and Zador A M 1998a Novel integrate-and-fire-like model of repetitive firing in cortical neurons *Proc. 5th Joint Symp. on Neural Computation*

——1998b Input synchrony and the irregular firing of cortical neurons *Nat. Neurosci.* **1** 210

Tang Y P, Shimizu E, Dube G R, Rampon C, Kerchner G A, Zhuo M, Liu G and Tsien J Z 1999 Genetic enhancement of learning and memory in mice *Nature* **401** 63–9

Tuckwell H C 1988 *Introduction to Theoretical Neurobiology* vol 1 (Cambridge: Cambridge University Press)

Wang X J 1999 Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory *J. Neurosci.* **19** 9587–603

Wilson H R and Cowan J D 1972 Excitatory and inhibitory interactions in localized populations of model neurons *Biophys. J.* **12** 1–24