Synaptic Mechanisms and Network Dynamics Underlying Spatial Working Memory in a Cortical Network Model

Article *in* Cerebral Cortex · September 2000

DOI: 10.1093/cercor/10.9.910

CITATIONS
654

READS
385

4 authors, including:

Albert Compte
IDIBAPS August Pi i Sunyer Biomedical Research Institute
78 PUBLICATIONS **3,702** CITATIONS
SEE PROFILE

Nicolas Brunel
University of Chicago
159 PUBLICATIONS **10,791** CITATIONS
SEE PROFILE

Xiao-Jing Wang
New York University
141 PUBLICATIONS **16,800** CITATIONS
SEE PROFILE

Synaptic Mechanisms and Network Dynamics Underlying Spatial Working Memory in a Cortical Network Model

Albert Compte, Nicolas Brunel[1], Patricia S. Goldman-Rakic[2] and Xiao-Jing Wang

Volen Center for Complex Systems, Brandeis University, Waltham, MA 02254 and [2]Section of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA
[1]Permanent address: LPS (Laboratory associated with CNRS, Paris 6 and Paris 7 Universities), Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France

Single-neuron recordings from behaving primates have established a link between working memory processes and information-specific neuronal persistent activity in the prefrontal cortex. Using a network model endowed with a columnar architecture and based on the physiological properties of cortical neurons and synapses, we have examined the synaptic mechanisms of selective persistent activity underlying spatial working memory in the prefrontal cortex. Our model reproduces the phenomenology of the oculomotor delayed-response experiment of Funahashi *et al.* (S. Funahashi, C.J. Bruce and P.S. Goldman-Rakic, Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 61:331–349, 1989). To observe stable spontaneous and persistent activity, we find that recurrent synaptic excitation should be primarily mediated by NMDA receptors, and that overall recurrent synaptic interactions should be dominated by inhibition. Iso-directional tuning of adjacent pyramidal cells and interneurons can be accounted for by a structured pyramid-to-interneuron connectivity. Robust memory storage against random drift of the tuned persistent activity and against distractors (intervening stimuli during the delay period) may be enhanced by neuromodulation of recurrent synapses. Experimentally testable predictions concerning the neural basis of working memory are discussed.

Introduction

Dorsolateral prefrontal cortex (PFC) plays a pre-eminent role in visuospatial working memory, as demonstrated by convergent evidence from ablation and reversible lesion studies (Goldman-Rakic, 1987; Fuster, 1988), brain imaging (McCarthy *et al.*, 1994; Courtney *et al.*, 1998; Zarahn *et al.*, 1999) and primate physiological studies (Fuster, 1973; Niki and Watanabe, 1976; Funahashi *et al.*, 1989; Chafee and Goldman-Rakic, 1998; Rainer *et al.*, 1998; Sawaguchi and Yamane, 1999). In an oculomotor delayed-response task (Funahashi *et al.*, 1989), when a monkey is required to retain information of a visual cue location through a delay period (a few seconds) between the stimulus and memory-guided behavioral response, PFC neurons show location-tuned elevated activity through the entire delay period. Presumably, information about the cue location is encoded by a selective neural assembly that subserves active memory storage by virtue of its sustained firing activity. Persistent activity has also been reported for neurons in posterior parietal cortex (PPC) during delayed oculomotor response experiments (Gnadt and Andersen, 1988; Colby *et al.*, 1996; Constantinidis and Steinmetz, 1996; Chafee and Goldman-Rakic, 1998). It is thus conceivable that mnemonic activity is maintained by reverberatory loops between the PFC and PPC (Fuster, 1988; Goldman-Rakic, 1987; Chafee and Goldman-Rakic, 1998; Sarnthein *et al.*, 1998). However, location-specific persistent activity in PPC was found to be easily disrupted by intervening stimuli during the delay period, while the monkey's working memory performance was not impaired (Constantinidis and Steinmetz, 1996). An alternative mechanism is that a neural circuit within the PFC may be by itself capable of sustaining selective persistent activity. In support of that idea, recent studies have demonstrated local recurrent excitatory connections both anatomically (Levitt *et al.*, 1993; Kritzer and Goldman-Rakic, 1995) and physiologically (González-Burgos *et al.*, 2000), as well as interactions between pyramidal and non-pyramidal neurons at microcolumnar and macrocolumnar ranges (Rao *et al.*, 1999) in the PFC.

A number of theoretical models have attempted to account for selective persistent activity, based on the assumption that persistent activity is sustained by reverberatory excitation within a local recurrent neural network (Hebb, 1949; Amit, 1995). Most previous computational analyses have used firing rate models (Wilson and Cowan, 1973; Amari, 1977; Zipser *et al.*, 1993; Amit *et al.*, 1994; Seung, 1996; Camperi and Wang, 1998; Moody *et al.*, 1998; Durstewitz *et al.*, 1999). Firing rate models, however, are difficult to relate directly with the physiological data. In particular, the issue of spontaneous versus persistent activity cannot be properly analyzed. Moreover, realistic time courses of synaptic interactions between neurons are typically ignored. For these reasons, a direct dialog between models and cortical synaptic physiology has been lacking. Recently, several studies have brought models closer to experimental data. Amit and Brunel used a spiking neuron model, instead of a rate model, for object working memory (Amit and Brunel, 1997). This approach allowed them to explicitly address the question of spontaneous activity and the generation of persistent activity by a specific structured connectivity.

Lisman and co-workers proposed that the voltage sensitivity of NMDA receptors (NMDARs) at recurrent synapses could underlie the stimulus-selectivity of neuronal persistent activity (Lisman *et al.*, 1998). On the other hand, Wang found that, in order to realize a stable, low-rate, persistent activity coexisting with a stable resting state, recurrent excitation should be primarily mediated by kinetically slow synapses of the NMDA type (Wang, 1999).

In this paper, we present a PFC network model for spatial working memory that combines insights from previous modeling studies on persistent activity in recurrent circuits (Amit and Brunel, 1997; Camperi and Wang, 1998; Wang, 1999). The structure of recurrent connnectivity is consistent with a columnar organization of cortical circuitry (Levitt *et al.*, 1993; Goldman-Rakic, 1995; Kritzer and Goldman-Rakic, 1995; Mountcastle, 1997), similar to network architectures that have been proposed in other cortical network models (Ben-Yishai *et al.*, 1995; Douglas *et al.*, 1995; Somers *et al.*, 1995). The model incorporates physiological data from slice preparations on the membrane parameters and input-output transductions of pyramidal and nonpyramidal neurons (McCormick *et al.*, 1985), and on the postsynaptic current gating kinetics of AMPA receptor (AMPAR), NMDAR (Hestrin *et al.*, 1990; Jahr and Stevens, 1990;

Spruston *et al.*, 1995) and $GABA_A$ receptor ($GABA_AR$) (Salin and Prince, 1996; Xiang *et al.*, 1998) mediated synaptic transmission. By using our model to reproduce the Funahashi experiment, we investigated synaptic mechanisms and network dynamics that may account for the salient observations on neuronal correlates of working memory. We also investigated the robustness of working memory storage against distraction stimuli and noise. Predictions from our theoretical results that are testable by *in vitro* physiological studies and single-neuron recording from behaving monkeys will be discussed.

Materials and Methods

The network model represents a local circuit of the monkey dorsolateral prefrontal cortex. It includes N_E pyramidal cells, and N_I interneurons. Pyramidal cells are four times more numerous than interneurons (N_E/N_I = 4). We assume a network architecture that is consistent with a columnar organization of the monkey PFC (Goldman-Rakic, 1995; Rao *et al.*, 1999; Constantinidis *et al.*, 1999; Ó Scalaidhe and Goldman-Rakic, 1999), and similar to models of the primary visual cortex (Ben-Yishai *et al.*, 1995; Somers *et al.*, 1995; Tsodyks and Sejnowski, 1995). Neurons are spatially distributed according to the stimulus to which they are most sensitive (preferred cue in an oculomotor delayed response task) and their collaterals may differentially target neighboring (isodirectional) and distant (crossdirectional) neurons. Cells receive external synaptic inputs which are indicative of the angle of the peripheral cue during its presentation. Each model neuron is labeled by its preferred cue position (an angle), and neurons of the network cover uniformly all the angles along a circle. Therefore, the cells are spatially distributed on a ring and their position in the ring has a linear relationship with their preferred cue angle.

Both pyramidal cells and interneurons are modeled as leaky integrate and fire units (Tuckwell, 1988). Each type of cell is characterized by six intrinsic parameters: the total capacitance C_m, the total leak conductance g_L, the leak reversal potential E_L, the threshold potential V_{th}, the reset potential V_{res}, and the refractory time τ_{ref}. The values that we use in the simulations are C_m = 0.5 nF, g_L = 25 nS, E_L = −70 mV, V_{th} = −50 mV, V_{res} = −60 mV, and τ_{ref} = 2 ms for pyramidal cells; and C_m = 0.2 nF, g_L = 20 nS, E_L = −70 mV, V_{th} = −50 mV, V_{res} = −60 mV, and τ_{ref} = 1 ms for interneurons (Troyer and Miller, 1997; Wang, 1999). All cells receive external excitatory inputs from other cortical areas. This overall external input is modeled as uncorrelated Poisson spike trains to each neuron at a rate of ν_{ext} = 1800 Hz per cell (or equivalently, 1000 presynaptic Poisson spike trains at 1.8 Hz). The external input is exclusively mediated by AMPARs, with the maximum conductance $g_{ext,E}$ = 3.1 nS on pyramidal cells, and $g_{ext,I}$ = 2.38 nS on interneurons.

Neurons receive their recurrent excitatory inputs through AMPAR- and NMDAR-mediated transmission and their inhibitory inputs through $GABA_AR$s. Synaptic responses are modeled as by Wang (Wang, 1999): postsynaptic currents are modeled according to $I_{syn} = g_{syn}s(V - V_{syn})$, where g_{syn} is a synaptic conductance, s a synaptic gating variable, and V_{syn} the synaptic reversal potential (V_{syn} = 0 for excitatory synapses, V_{syn} = −70 mV for inhibitory synapses). AMPAR and $GABA_AR$ synaptic gating variables are modeled as an instantaneous jump of magnitude 1 when a spike occurs in the presynaptic neuron followed by an exponential decay with time constant 2 ms for AMPA (Hestrin *et al.*, 1990; Spruston *et al.*, 1995) and 10 ms for $GABA_A$ (Salin and Prince, 1996; Xiang *et al.*, 1998). The NMDA conductance is voltage dependent, with g_{syn} multiplied by $1/(1 + [Mg^{2+}]\exp(-0.062V_m)/3.57)$ (Jahr and Stevens, 1990), $[Mg^{2+}]$ = 1.0 mM. The channel kinetics is modeled by the following equations:

$$\frac{ds}{dt} = -\frac{1}{\tau_s}s + \alpha_s x(1-s) \qquad \frac{dx}{dt} = -\frac{1}{\tau_x}x + \sum_i \delta(t - t_i)$$

where s is the fraction of open channels, x is an intermediate gating variable, t_i are the presynaptic spike times, τ_s = 100 ms is the decay time of NMDA currents, τ_x = 2 ms controls the rise time of NMDAR channels, and α_s = 0.5 kHz controls the saturation properties of NMDAR channels at high presynaptic firing frequencies.

The recurrent connections between neurons in the network depend on the difference between their preferred cues. This is implemented by taking the conductance between neuron i and neuron j to be $g_{syn,ij} = W(\theta_i - \theta_j)G_{syn}$, where $W(\theta_i - \theta_j)$ is the 'connectivity footprint' normalized as

$$\frac{1}{360}\int_0^{360} W(\theta_i - \theta_j)d\theta_j = 1$$

The functional form of W is chosen to be either a constant for unstructured connections or the sum of a constant term plus a Gaussian centered at $\theta_i - \theta_j = 0$: $W(\theta_i - \theta_j) = J^- + (J^+ - J^-)\exp[-(\theta_i - \theta_j)^2/2\sigma^2]$. In this equation, the dimensionless parameter J^- represents the strength of the weak crossdirectional connections, J^+ the strength of the stronger isodirectional connections, and σ is the width of the connectivity footprint (see Fig. 1). Note that the normalization condition of W imposes a functional relationship between the three parameters defining the connectivity. Therefore, we only mention in the following two parameters, J^+ and σ. J^- is then determined using the normalization condition. In most simulations, only the excitatory-to-excitatory connectivity is structured. Parameters of the corresponding footprint are J^+_{EE}, σ_{EE}. In a few simulations (Fig. 7), the excitatory-to-inhibitory connectivity is also structured. The parameters of the corresponding footprint are J^+_{EI}, σ_{EI}. In all simulations, the inhibitory connections are unstructured, i.e. the cross- and isodirectional components of inhibitory connections are equally strong.

In order to produce a desired level of spontaneous activity, we set the values of the conductances of external synapses $g_{ext,E}$, $g_{ext,I}$, and the frequency of external inputs ν_{ext} so that each neuron receives strong suprathreshold input from external sources. We then impose that inhibition be stronger than recurrent excitation, by choosing a high enough ratio of inhibitory to excitatory conductances on each cell type, i.e. a high enough value of $G_{IE}/G_{EE} = G_{II}/G_{EI}$. Adjusting these parameters allows control of the level of spontaneous activity of both excitatory and inhibitory populations (Amit and Brunel, 1997). A structured pyramid-to-pyramid connectivity gives rise to tuned network persistent activity states. This is accomplished by a gradual increase of J^+, until the network shows a bistability between homogeneous spontaneous activity and tuned persistent activity (Amit and Brunel, 1997). An important point to make here is that the pyramid-to-pyramid footprint W_{EE} is always normalized to 1, so that an increase in the synaptic strength between neurons with similar preferred cues implies a decrease in the strength of connection between neurons with dissimilar preferred cues. This allows the preservation of the level of spontaneous activity as the connectivity footprint is varied.

In most of the simulations shown in this paper, only NMDAR channels were included at the recurrent excitatory synapses (for simplicity), since previous work suggested that dominance of the recurrent excitation by NMDARs favors network stability (Wang, 1999). In some simulations we assessed the network stability when the AMPAR contribution to the recurrent connections was included (see below). Typically (for the 'control parameter set'), N_E = 2048, N_I = 512. The recurrent excitatory synapses mediated by NMDAR channels have conductances G_{EE} =

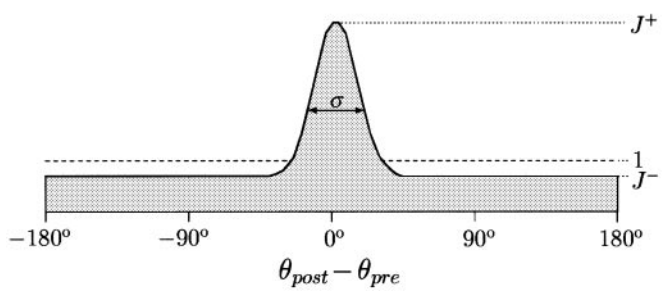

Figure 1. Structured connectivity of the model. The synaptic connection strength decreases with the difference in the preferred cues of two neurons, with strong interactions between neighboring neurons and weak interactions between more distant neurons.

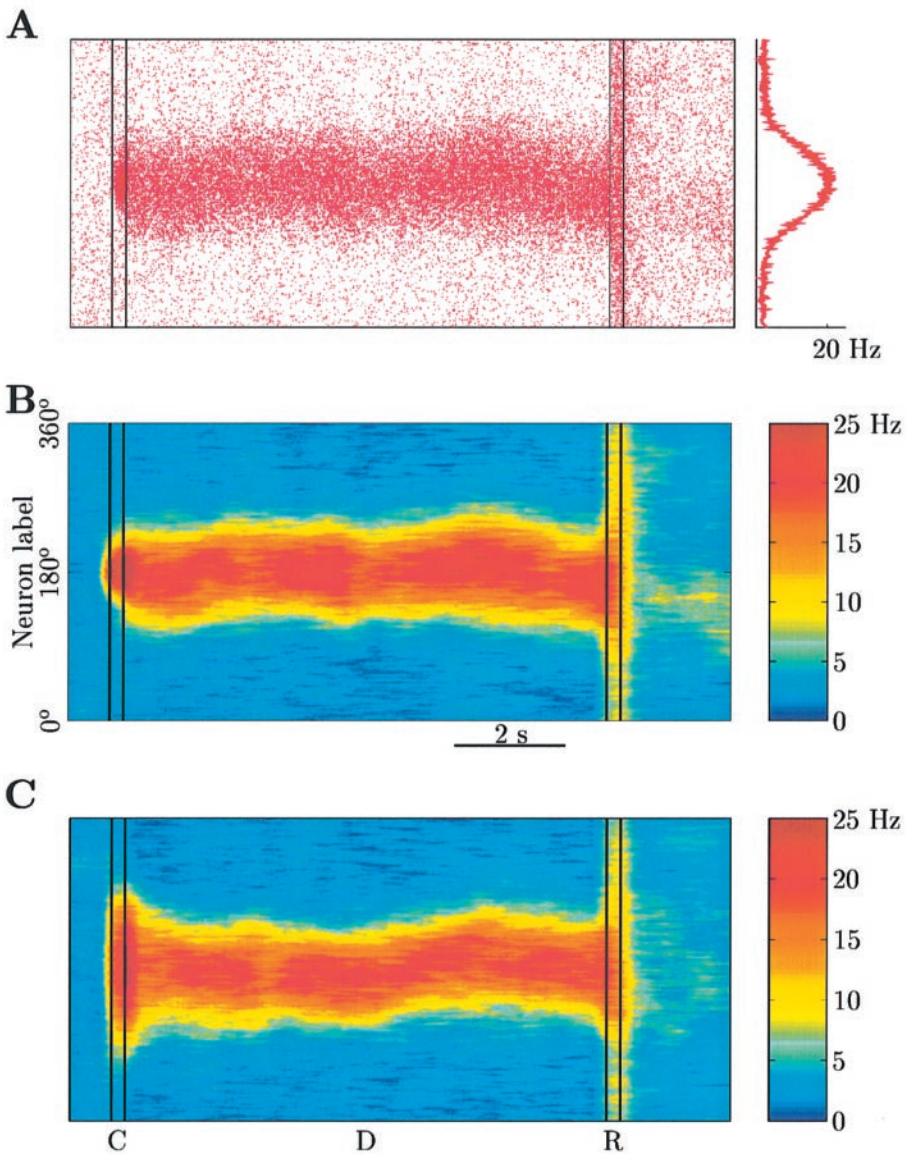

Figure 2. Working memory maintained by a tuned network activity state (a 'bump state'). C, cue period (250 ms, peak stimulus 200pA); D, delay period (8.75 s); R, response period (250 ms, external current increase 500 pA). (*A*) Pyramidal neurons rastergram. The *x* axis represents time, while the *y* axis represents neuron label according to its preferred cue. A dot in the rastergram indicates a spike of a neuron whose preferred location is at *y*, at time *x*. Note the enhanced and localized neural activity that is triggered by the cue stimulus and persists during the delay period. The population firing profile, averaged over the delay period, is shown on the right. (*B*) Color-coded spatiotemporal activity pattern (see Materials and Methods). (*C*) Same as (*B*), with less specific cue stimulation (5-fold increase in cue width). The network reaches a bump state with the same width as in (*B*) during the delay. In these simulations, inhibitory interneurons (not shown) display a spontaneous activity rate of 9 Hz and an increased delay activity rate of 13 Hz.

0.381 nS (pyramid-to-pyramid), G_{EI} = 0.292 nS (pyramid-to-interneuron); inhibitory synaptic conductances are G_{IE} = 1.336 nS (interneuron-to-pyramid), G_{II} = 1.024 nS (interneuron-to-interneuron); the connectivity footprint has characteristics σ_{EE} = 18° and J^{+}_{EE} = 1.62. Recurrent conductances are scaled inversely proportionally to N when network size is varied, to keep the total synaptic conductances unchanged. We also performed simulations with a different parameter set (Figs 4, 8), referred to as a 'modulated' parameter set. It has the same parameters as the control, except for an enhancement in the recurrent conductances: 20% increase for NMDAR-mediated synaptic transmission (G_{EE} and G_{EI}) and 40% increase for GABA synapses (G_{IE} and G_{II}). In other simulations, AMPAR-mediated synaptic transmission was introduced in the recurrent connectivity (Fig. 6). For a 67% NMDA contribution to recurrent excitatory charge entry at a holding potential of –65 mV (Fig. 6*A*) the parameters are $G_{EE,AMPA}$ = 0.251 nS, $G_{EE,NMDA}$ = 0.274 nS, $G_{EI,AMPA}$ = 0.192 nS, and $G_{EI,NMDA}$ = 0.212 nS. For a 50% NMDA contribution to recurrent connections at a holding potential of –65 mV (Fig. 6*C*), the parameters are $G_{EE,AMPA}$ = 0.393 nS, $G_{EE,NMDA}$ = 0.214 nS, $G_{EI,AMPA}$ = 0.304 nS, and $G_{EI,NMDA}$ = 0.164 nS.

The simulation protocol was chosen to mimic the protocol used in the experiment of Funahashi *et al.* (Funahashi *et al.*, 1989). In that experiment, monkeys were trained to fixate a central spot during a brief presentation (0.5 s) of a peripheral cue and throughout a subsequent delay period (1–6 s), and then to make a saccadic eye movement to where the cue had been presented in order to obtain a reward. In our simulations, cue presentation to the network is modeled through selective transient current injection to pyramidal cells whose preferred cues are close to the stimulus. During the delay period, all selective external currents are absent. After the end of the delay period, we model the effect of the motor response and reward on the network by a transient nonspecific current injection to all neurons. In several simulations, we tested the influence of distractors on the network dynamics. Distractors

are modeled as a cue stimulus (same strength, same duration) but at a different location relative to the cue stimulus.

To visualize network activity, pseudo-color spatiotemporal firing patterns were calculated. A spike time rastergram for all pyramidal neurons (or interneurons) is smoothed with a sliding window both in time (500 ms in Figs 2 and 7, 10 ms in Fig. 6*A*,*C*, 5 ms in Fig. 6*B*, and 250 ms in Fig. 8) and along the neuron population (15 neurons in all figures except Fig. 6, where it is 125 neurons). The resulting firing rate is color encoded using the Matlab package. The population vector (Georgopoulos *et al.*, 1986) was used to estimate the evolution of the peak location of the bump state (see Fig. 5). To compute the population vector during a given time interval, we first compute a vector for each pyramidal cell, with direction given by the preferred direction of that cell and the amplitude proportional to the firing rate of the cell during the corresponding time interval. The population vector is then given by the sum of all individual vectors. Its angle with respect to a reference frame is a measure of the stored positional cue in the corresponding time interval. It is a simple and convenient method to estimate the peak of the activity profile at a given time. In Figure 8 a maximum likelihood estimator was used to assess the memorized angle at a given time (angles θ_1 and θ_2 in Fig. 8). The population activity profile before the distractor was fitted to the population profile after the distractor. This method picks the position of the most salient bump present in the profile in contrast to the population vector, which computes the mean position of all bumps in the profile. 'Local field potentials' (see Fig. 6) were computed by averaging the synaptic variable s_{AMPA} across all the pyramidal neurons at each time step.

The integration method used is a second-order Runge–Kutta algorithm with the firing time interpolation scheme (Hansel *et al.*, 1998b) and a time step of $\Delta t = 0.02$ ms. The code for the simulations has been written in C++. When run on a Linux 550 MHz Pentium III PC, a 6 s trial with 2048 pyramidal neurons and 512 interneurons typically takes 2 h to complete.

Results

Persistent Activity and Memory Fields

Our network model simulation used the protocol of the experiment of Funahashi *et al.* (Funahashi *et al.*, 1989), consisting of a cue presentation (C) followed by a delay period (D) then a response period (R) (see Materials and Methods for details). The network activity during any given trial was monitored by plotting its spatiotemporal firing pattern. Figure 2*A*,*B* shows two different ways of showing the temporal evolution of network activity. In both plots, the abscissa represents time, while the ordinate represents pyramidal neurons arranged according to their preferred cue directions. In Figure 2*A*, spikes of all pyramidal neurons are shown in a rastergram. In Figure 2*B*, spatiotemporal activity is smoothed (see Materials and Methods) and shown in a continuous and color-coded map. The main features of network activity can be read out from these graphs from left to right. First, before cue presentation, neurons show spontaneous activity at a few spikes per second. This activity is uniform in space: the network is untuned. Such a low spontaneous activity is an emergent property of the network. It is achieved through a combination of suprathreshold external inputs representing background activity in other brain areas, and of a strong feedback inhibition in the network.

Second, during the cue period (C), a pattern of increased activity develops around the location of the cue (180°). This increased activity is due to the external input to the subpopulation of neurons with preferred cues closest to the cue stimulus (approximately those with preferred cues between 162° and 198°, see Materials and Methods).

Third, in the delay period (D), the network initially localized response widens and stabilizes. The elevated persistent activity remains restricted to a selective neural subpopulation throughout the delay period. This is quantified by the peaked network profile of the averaged delay-period activity (right panel in Fig. 2*A*). The enhanced persistent rates (~20 Hz) are achieved through the strong excitatory feedback between cells sharing similar tuning properties. In Figure 2*C*, the same simulation is repeated, but with a more broadly tuned cue stimulus (cells between 90° and 270° are activated by the stimulus). Thus, the network response during the cue period is more widespread. However, the network persistent activity eventually evolves to the same profile during the delay period as with a more specific cue stimulation (Fig. 2*B*). Therefore, the tuned persistent activity profile is independent of the precise shape (or intensity) of the cue stimulus. Such a 'bump state' is an *attractor* of the network dynamics.

Finally, during the response period (R), a transient and overall increase of external inputs to the whole network leads to a transient increase of neuronal firing, which turns off the persistent activity (Funahashi *et al.*, 1991; Goldman-Rakic *et al.*, 1990). This 'switching off' of persistent activity by excitation is due to the strong inhibitory feedback. A global excitatory drive to the network increases the firing rate of inhibitory cells in a way that is strong enough to effectively wipe out persistent activity and refresh the short-term memory.

An example of a single neuron's selective persistent activity is shown in Figure 3 for a pyramidal cell with preferred cue at 260°. During a delayed-response simulation in which the cue location was 270°, the spike trains were recorded for several trials (Fig. 3*A*). The cell shows a low rate (3.5 Hz) of spontaneous activity. In response to a transient cue at 270° (250 ms), the cell displays an enhanced persistent activity during the delay period (8.75 s). Both spontaneous and delay-period firing activities are quite irregular in time (Fig. 3*A*). The cell is switched back to the spontaneous state during the response period (250 ms). With eight cue presentations, the average firing rate of the neuron's persistent activity for each cue is computed and the resulting tuning curve is shown in Fig. 3*B*. The persistent activity is tuned to a memory field around 270°. Note that the maximum persistent activity rate is ~20 Hz, within the physiological range of PFC neurons. Moreover, the firing rate of delay period activity for a non preferred cue is *lower* than the spontaneous rate, as is often observed experimentally (Funahashi *et al.*, 1989). The tuning curve can be fitted by a Gaussian function. It is significantly broader (with a width of ~40°) than the width of the recurrent excitatory connectivity (18°, see Materials and Methods). The width of the tuning curve (or the size of the neuron's *memory field*) depends also on a variety of network parameters, such as the relative strengths of recurrent synaptic excitation and inhibition, as we shall see below.

The persistent activity of a single neuron is sustained by synaptic excitation from the rest of the network. To dissect various contributions to the synaptic drive to a cell, we calculated separately the components of the synaptic inputs to the cell during both spontaneous state and delay period (Fig. 3*C*). The time-average over the delay period of the different types of synaptic inputs to the cell (external excitatory, recurrent excitatory and recurrent inhibitory inputs) are plotted as a function of the cue stimulus. Several important features of the network activity can be seen from the graph. It shows that the overall recurrent (excitatory plus inhibitory) input from the other cells in the network is negative, so that the net effect is hyperpolarizing, and the neuron is restrained from firing at the high rates that would otherwise be imposed by external inputs alone. This is a consequence of the inhibition dominance of the recurrent interactions, which is crucial to the network function.

Another important point can be made from Figure 3*C* by comparing the persistent activity state with the spontaneous state. Even though there is a significant increase in both excitatory and inhibitory recurrent currents in the persistent activity state, the summated total recurrent input remains approximately unchanged with respect to the spontaneous state. Thus, in the present network, excitation and inhibition balance each other dynamically. A balance between excitation and inhibition has also been suggested to account for the irregularity of interspike intervals of cortical cells *in vivo* (Shadlen and Newsome, 1994; Tsodyks and Sejnowski, 1995; Amit and Brunel, 1997; van Vreeswijk and Sompolinsky, 1998). For cues within the neuron's memory field, neighboring cells with similar preferred cues show enhanced firing and send increased lateral excitation to each other, so that the total recurrent input to the cell is higher than during spontaneous activity, leading to an enhanced persistent activity. On the other hand, when the cue location is very different from the preferred cue, cells with enhanced persistent activity are far away from (and thus send little excitation to) the recorded cell, and the total recurrent input shows actually a slight decrease during the delay period, which explains why the firing rate is lower than spontaneous activity in that situation (Fig. 3*B*).

As mentioned above, the size of the memory field depends on the interplay between recurrent excitation and inhibition. Memory fields are thus under modulatory control of various neurotransmitters via their action on synaptic transmissions. In particular, a concomitant enhancement of recurrent synaptic excitation (20%) and inhibition (40%) (modulated parameter set, see Materials and Methods) leads to a sharper tuning of persistent activity (in Fig. 4*A*, the width of the tuning curve is 30°. Compare with 40° in Fig. 3). Stronger inhibition reduces the spontaneous activity, whereas stronger excitation leads to a higher mnemonic activity for the preferred cue, thereby increasing the separation between the firing rates of the two states (the signal-to-noise ratio). We show in Figure 4*A* the rastergrams of a pyramidal neuron with a preferred cue at 270° for eight cue stimuli. For comparison, the data from a neuron in the principal sulcus recorded by Funahashi *et al.* (Funahashi *et al.*, 1989, Figs 3 and 9) is plotted in Figure 4*C*. Note the similarities between the simulated cell and the real cell. Indeed, both cells show a low spontaneous activity, a cue-selective delay activity after stimulus presentation, a high degree of variability of spike trains during both spontaneous and persistent states, and a transient increase in the firing rate during the response period before the cell returns to its spontaneous state. Activity during the cue period is controlled by the intensity of the external stimulation during that period and it is not directly related to the delay activity, which is intrinsically set by the network synaptic conductances. Finally, the tuning curves of both simulated and real cells during the delay period show comparable shape and rates (Fig. 4*B*,*D* respectively).

Random Drift of Memory

In the model, a network persistent activity can be peaked at an angle anywhere between 0° and 360°, which encodes the memory of a cue stimulus in a graded fashion. Thus, there is a continuum of such structured activity profiles ('bump states'), which is realized by the circular symmetry of the network. A particular bump state is selected by the cue during the stimulus presentation. However, after the cue is withdrawn, no external input is present to constrain the peak's location of the network activity profile during the delay period. This raises the question

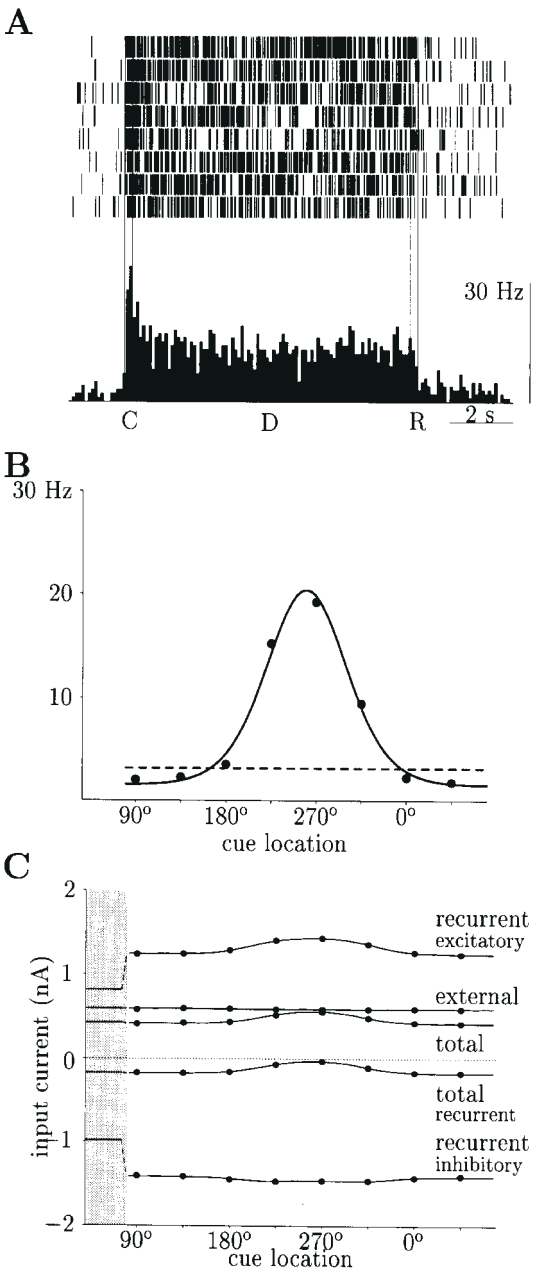

Figure 3. Synaptic mechanisms underlying spontaneous and persistent activity ('control' parameter set, see Materials and Methods). (*A*) Rastergram and average discharge rate vs time of a cell in trials in which a cue is shown close to the preferred angle of the cell. (*B*) Tuning curve of the cell in the delay period. The standard deviation of the Gaussian fit is 40°. (*C*) Components of the synaptic inputs of the cell. The shaded area shows the average inputs during spontaneous activity. To the right, we show the average synaptic components of the cell as a function of cue position: recurrent excitatory, recurrent inhibitory and external excitatory (labelled 'external'). The sum of both recurrent inputs is labelled 'total recurrent' and the sum of all independent inputs is labelled 'total'. Depolarizing inputs are positive and hyperpolarizing inputs negative. Note the increase in both recurrent excitatory and inhibitory inputs during the delay. Because of the balance between excitation and inhibition, the sum of the two (total recurrent) remains close to the level of the spontaneous state.

of the stability of one particular bump against random fluctuations that can move it to another adjacent bump, in which case the memory of the cue location would be lost.

The simulation of Figure 5 shows clearly that the bump, once elicited by a cue, is not completely stationary. Rather, it slowly

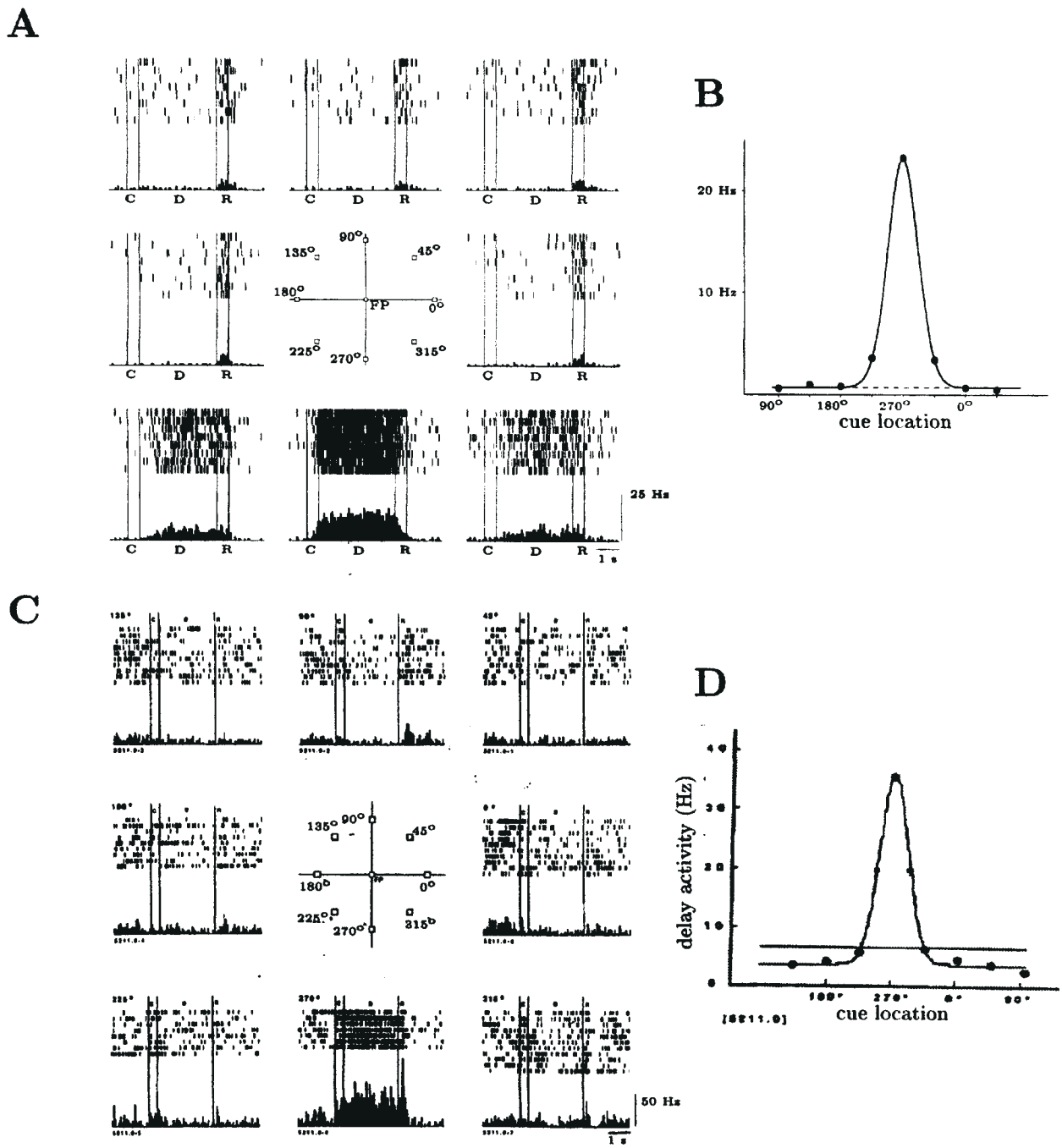

Figure 4. Single cell recording in experiment (Funahashi *et al.*, 1989) and in the simulation reveals directional delay period activity. (*A*,*B*) Network simulation with the modulated parameter set (see Materials and Methods); (*C*,*D*) Experiment of Funahashi *et al.* [(Funahashi *et al.*, 1989), see their Figs 3 and 9]. Each rastergram represents the response of the cell when the cue was shown in one of the eight locations indicated in the center diagram. Both cells respond vigorously in the delay only for one direction (270°), and are suppressed relative to inter-trial spontaneous activity in the upper visual field. The delay period tuning curves (*B*,*D*) show the average discharge rate during the delay period (circles), together with a Gaussian fit of the data. The horizontal line indicates average inter-trial spontaneous activity. Note the similarity between experiment and simulation.

drifts in a seemingly random fashion. Such a drift is due to the random inputs that the network continuously receives from outside. This random bombardment has a small effect on the bump location, so that even though the shape of the bump is stable, its location is only marginally stable as a result of the translational invariance along the circle. To quantify this drift behavior, we have estimated the instantaneous location of the bump (the peak of the network activity profile) using the population vector (see Materials and Methods). The time evolution of the population vector is shown for different trials and network sizes in Figure 5. For a given network size, the population vector drifts away in any single trial from the cue location during the delay period, indicating a slow deterioration of memory for the cue location. Note that the network has equal probabilities of drifting up and down, due to the circular symmetry (Fig. 5*A*). The three panels in Figure 5*A* clearly show

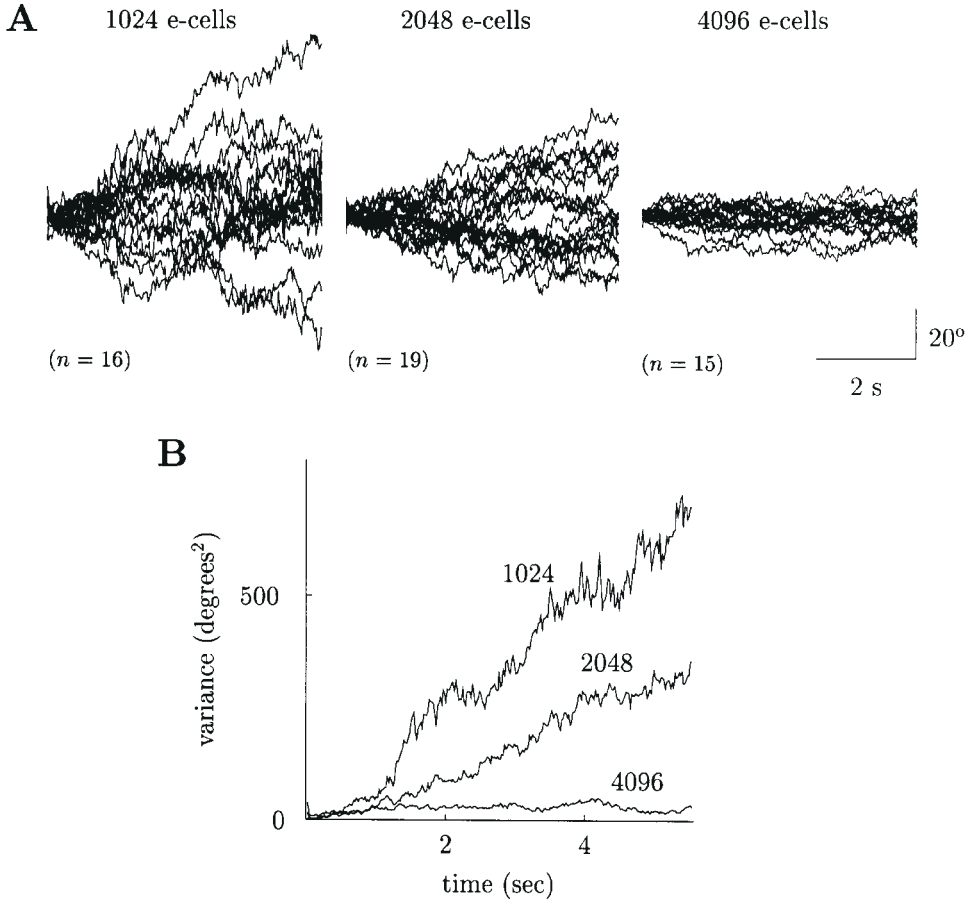

Figure 5. The network pattern of persistent activity drifts randomly in time due to noise, but memory storage is robust in large networks and weaker recurrent synaptic connections. (*A*) Population vector position versus time for different runs and different network sizes. (*B*) Variance of the population vector position around the initial stimulation point averaged across trials and plotted versus time for each of the three network sizes studied. Note the linear trend, similar to a diffusion process.

that the drift effect is smaller for larger network sizes, indicating that the memory of the cue location is more robust with a larger neural population in the network. The variance of the population vector as a function of time is plotted for different network sizes in Figure 5*B*, showing an approximate linear trend. This linear trend is consistent with a diffusion process (Berg, 1983). The figure shows also that the slope of the variance dramatically decreases with network size. Indeed, after a 4 s delay, the location has drifted on average by 20° in a network of 1024 neurons; by ~15° in a network of 2048 neurons; and by <10° in a network of 4096 neurons. Further simulations with the modulated parameter set of Figure 4*A* showed a remarkably smaller degree of drift at each network size compared to Figure 5 (data not shown), indicating that the memory drift can be controlled by neuromodulation of recurrent synapses.

Stability, Synchrony and the NMDA to AMPA Ratio

As can be seen in Figure 2*A*, neural discharges appear quite asynchronous both during the spontaneous state and the delay period. The lack of synchronicity is due to the predominance of NMDAR-mediated transmission at the recurrent excitatory synapses. To show this, we performed simulations with different relative contributions of the AMPAR and NMDAR to the recurrent synaptic excitation. This was done by varying the contribution of the NMDAR to the total charge entry into a cell by a unitary EPSC at recurrent connections. The neuronal firings are essentially asynchronous, as long as NMDAR currents contribute at least 75% to the total charge entry mediated by excitatory recurrent inputs (at a holding potential of –65 mV). With less NMDA contribution, neurons become partially synchronized, as shown in Figure 6A. On a large scale (Fig. 6*A*), the structure of the network persistent activity profile remains similar to that in Figure 2*A*. On a fine temporal scale, however (Fig. 6*B*), the temporal structure of the network activity has dramatically changed from asynchronous behavior to pronounced synchronized oscillations at ~40 Hz (see power spectrum in the right panel). These oscillations are due to the fact that when the contribution of AMPAR channels to recurrent connections becomes large, these synaptic inputs with a fast time constant tend to produce surges of activity, that are later dampened by the slower synaptic inhibition. Thus, an oscillatory behavior emerges. The network behavior remains irregular due to external noise. Another important characteristic of the network is that firing rates of persistent activity tend to be higher as the NMDA contribution decreases. This can be explained by the lack of saturation of the steady-state, AMPAR-mediated synaptic response at physiological firing rates, while NMDARs saturate at rather low presynaptic frequencies. For still smaller NMDA contribution to total charge entry, oscillations become more pronounced. Fluctuations in the network dynamics eventually destroy persistent activity (Fig. 6*C*).

Therefore, our simulations show that (i) NMDARs are

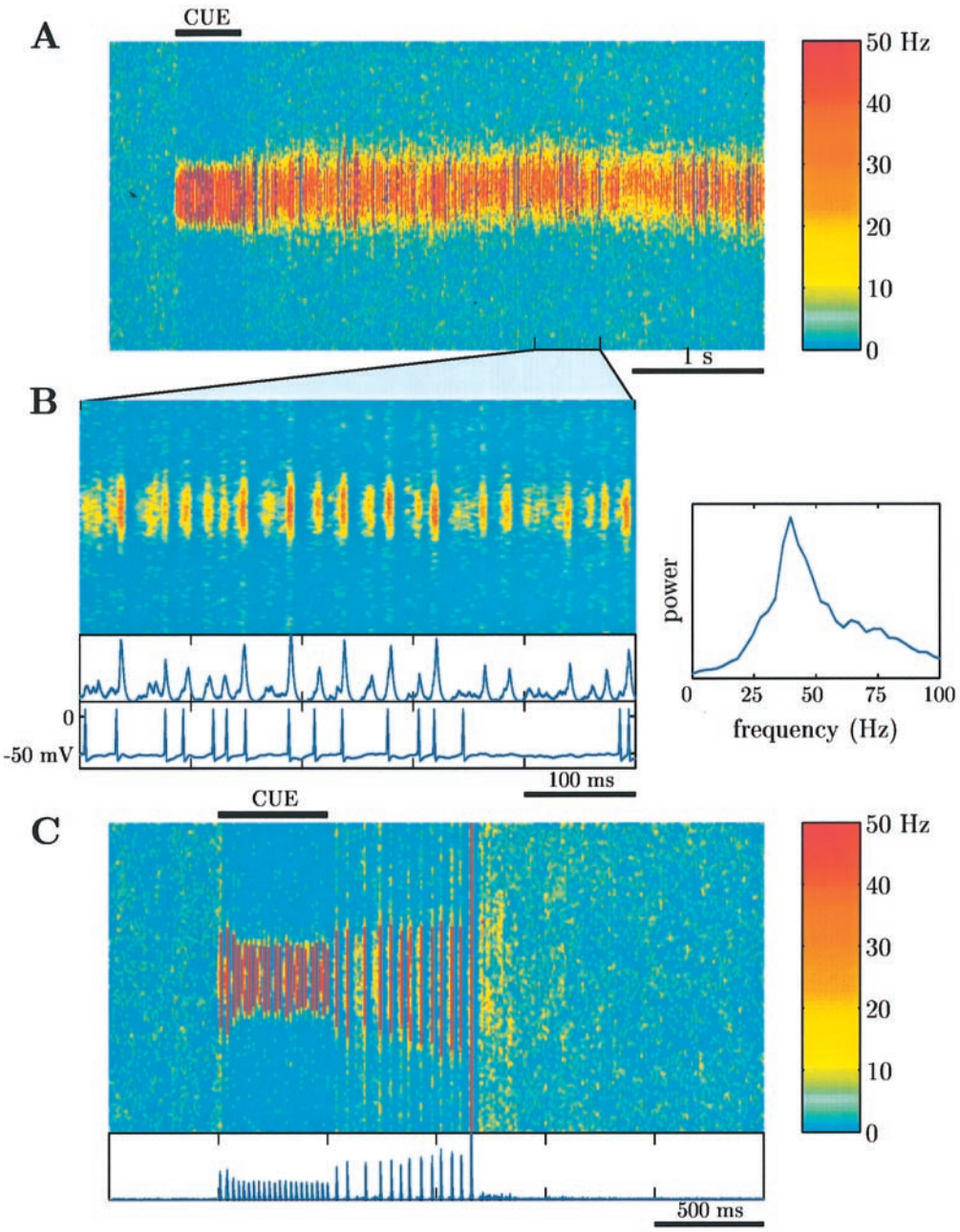

Figure 6. A decrease of the NMDAR channel contribution to recurrent synapses gives rise to oscillations in the delay period. (*A*) Network spatiotemporal firing pattern with moderate AMPA component in recurrent interactions. Here the NMDAR channels contribute 67% to the total recurrent excitatory charge entry at a holding potential of –65 mV (NMDA:AMPA ratio 0.038 in terms of the peak elicited EPSC). (*B*) 500 ms blowup of the upper panel to show the AMPA-induced oscillations, the local field potential and the membrane potential of a single neuron. On the right is shown the power spectrum of the local field, demonstrating a large peak at ~40 Hz. (*C*) Example of the disruptive effect of weak NMDA component in the delay period activity of the network. The NMDA contribution to the recurrent excitatory charge entry is here 50% at a holding potential of –65 mV (NMDA:AMPA ratio 0.019 in terms of peak EPSC). Note the eventual recruitment of the whole network that abolishes persistent activity. The network then goes back to the asynchronous spontaneous activity. The local field shows a very strong oscillation before disruption of persistent activity.

necessary for sustaining a dynamically stable persistent activity; and (ii) rhythmic oscillations in the gamma frequency range (20–80 Hz) readily occur in such a strongly recurrent network, if a substantial component of recurrent excitation is mediated by AMPA receptors. This is a phenomenon produced by the interplay between a fast positive feedback followed by a slower negative feedback (Wang, 1999).

Tuning in Inhibitory Cells

Recently, recordings from putative inhibitory neurons in the PFC have been reported in monkey experiments using the same oculomotor delayed-response paradigm (Rao *et al.*, 1999). It was found that some PFC fast-spiking interneurons display tuned persistent activity, similar to pyramidal cells. In our network simulation of Figure 2, only excitatory cells show tuning to the

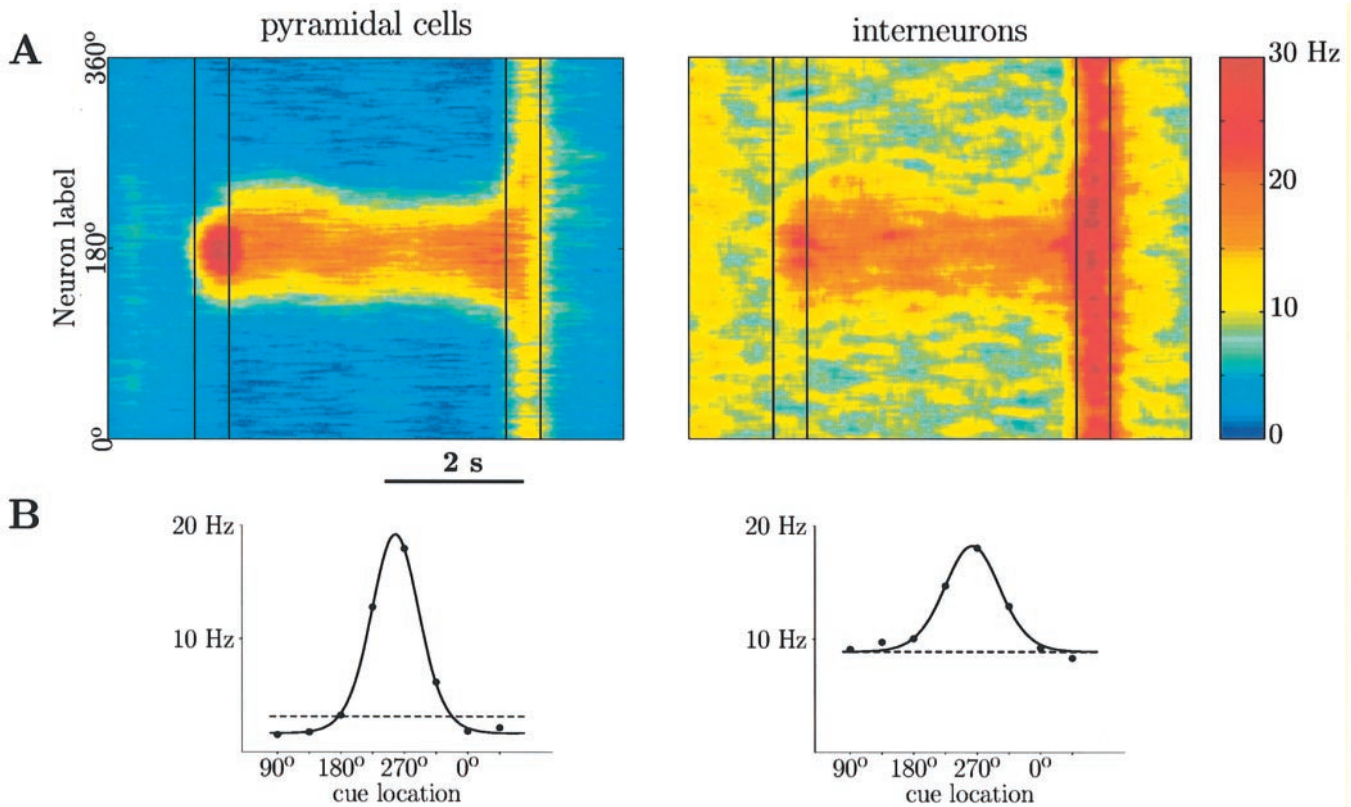

Figure 7. Memory fields of interneurons with structured pyramid-to-interneuron connections. Left panels show pyramidal neurons, and right panels show inhibitory neurons. (*A*) Spatiotemporal firing patterns showing a 'bump state' during the delay period both in the pyramidal neurons and in the interneurons. (*B*) Tuning curves for a single pyramidal neuron and a single neighboring interneuron (preferred cue 260°), tested with eight cues.

cue. This is due to the fact that in that simulation, only connections between excitatory cells were structured. Tuning in inhibitory cells can be realized by introducing a structured pyramid-to-interneuron connections. Network simulations were performed with the parameter set of Figure 2, except for σ_{EI} = 18° and J^{+}_{EI} = 1.25. Results are shown in Figure 7. In that figure, the population of inhibitory cells shows a peaked activity profile ('bump state') during the delay period, similar to excitatory cells (Fig. 7*A*). The tuning of inhibitory cells is less pronounced (Fig. 7*B*) because the pyramid-to-interneuron connectivity was chosen to be more weakly modulated than the pyramid-to-pyramid connectivity. Sharpening further the pyramid-to-interneuron connectivity results in higher inhibitory rates that eventually destabilize persistent activity. Similar to the experimental data (Rao *et al.*, 1999), in our model a pair of pyramidal cell and interneuron show similar tuning if they are close to each other; and orthogonal tuning if they are far apart.

Distractors

If information about a cue stimulus is required for a delayed behavioral response, it is important that the memory of the cue is maintained in spite of possible distraction inputs from outside of the memory network. We tested the resistance of the network memory storage to distractors presented as intervening stimuli during the delay period. The simulation protocol in the presence of distractors is shown in Figure 8. First, a cue stimulus is shown at angle θ_S. It elicits a bump state that stores the memory of that stimulus, up to a small drift. The bump state peaks at an angle θ_1 (close to θ_S) just prior to the presentation of a distractor at angle θ_D. A distractor has identical characteristics as a cue stimulus (in particular, *it has the same intensity and duration*), except that it is presented at a different location and during the delay period. We then measure the effect of the distractor by measuring the peak location of the bump state after distraction (angle θ_2). The effect of distraction is quantified by the difference in the peak location of the bump state before and after the distractor, $\theta_2 - \theta_1$. Distraction stimuli are presented at various cue positions, close to or far away from the original cue, in separate trials.

We studied how the behavior of the network is affected by distractors in the 'control' case (network parameters of Fig. 2), and with a modulatory enhancement of both NMDA (by 20%) and GABA (by 40%) recurrent synapses, resulting in an enhanced 'signal-to-noise' ratio (modulated parameter set as in Fig. 4*A*). We stress that in all cases, cue and distractor stimulation amplitudes were identical. If the stimulation amplitude is sufficiently large, the distractor is powerful enough to overcome the intrinsic dynamics of the recurrent circuit, and the network is always perturbed to a location close to the intervening stimulus (see sample trial in Fig. 8*A*, and red points in Fig. 8*C*). However, with a lower stimulus intensity (blue points in Fig. 8*C*), or with an enhanced signal-to-noise ratio (sample trial in Fig. 8*B* or orange trace in Fig. 8*D*), the network was found to be much more resistant to an intervening stimulus. If the distractor is close to the initial cue, the amount of distraction increases approximately linearly with the distance, reaching a maximum around $\theta_D - \theta_S$ = 90°. At larger distances (if the distractor and the initial cue are separated by >90°), the distraction becomes very small (<10°),

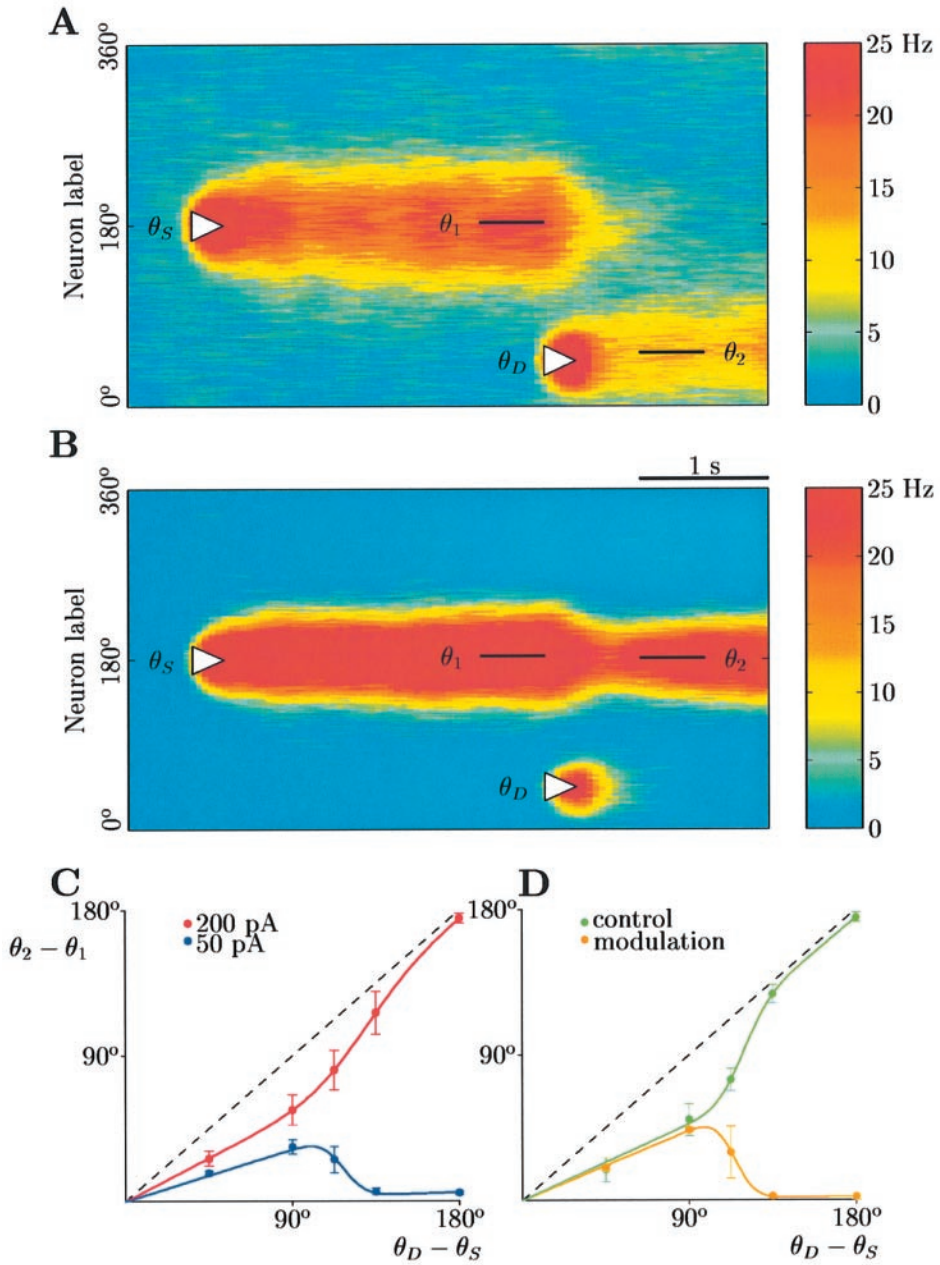

Figure 8. The network resists to distractors when the stimulus intensity is low. (*A*,*B*) Network spatiotemporal firing pattern The cue is presented initially for 250 ms at θ_S, triggering a tuned persistent activity. After a 2.5 s delay, a distractor stimulus is presented at θ_D, with the same intensity and duration as the cue stimulus. The population vector is computed in a window of 500 ms just before the distractor (θ_1) and 500 ms after the distractor (θ_2). (*A*) A case of complete distraction for the control network parameter set (see Materials and Methods) and strong stimulation (200 pA). (*B*) A case of perfect robustness to distraction for the modulated parameter set (see Materials and Methods) and moderate stimulation (100 pA). (*C*) Dependence of network distraction on the distance between the cue and distractor and on the stimulation intensity. The 'distracted' angle $\theta_2 - \theta_1$ is plotted versus distraction angle $\theta_D - \theta_S$ for several distractor cues. The dashed line indicates perfect distraction (as in *A*) while points on the *x*-axis show absence of distraction (as in *B*). Stimulation intensity is 50 pA (red) and 200 pA (blue) (control parameter set). (*D*) Comparison between control and modulated cases for a given stimulation paradigm (250 ms duration, 120 pA intensity). Modulation of both NMDAR- and $GABA_A$R-mediated synaptic transmission enhances dramatically the network's resistance to distractors, particularly at high distraction angles. $N_E = 4096$, $N_I = 1024$.

which shows that the network is essentially unaffected by distractors far from the cue location. The resistance to distractors can be understood by the fact that inhibition is much stronger in the persistent activity state. Thus, it is harder to elicit a new bump during the delay period than from the spontaneous activity state. This explains why in Figure 8*B* the distractor just elicits a transient increase in cells that receive direct inputs from the intervening stimulus. Inhibition dominance of the synaptic circuitry underlies the network's ability to ignore distractors, as long as the external inputs are not too strong to overrule the recurrent network dynamics. This resistance can be facilitated by an increased signal-to-noise ratio, which in our network can be brought about by concomitant modulation of recurrent conductances. On the other hand, if the inputs are very strong, the network is no longer resistant to intervening stimuli. In this case, the network can be reset by every new transient stimulus, and retains a memory of the last stimulus in the form of a refreshed selective persistent activity state.

Discussion

We have used a recurrent cortical network model that incorporates physiological properties of cortical neurons and synapses, to decipher the neuronal mechanisms underlying persistent activity in a spatial working memory circuit. In the paper we have focused on spatial working memory in PFC, in order to compare the model with physiological data from behaving monkeys, and thus be able to draw specific experimental predictions. However, we believe that the synaptic mechanisms identified in this study could be applicable to other types of mnemonic persistent activity observed in PFC as well as in other cortical areas.

Neuronal Mechanisms of Spatial Working Memory

Since the emergence of persistent activity requires sufficiently strong recurrent synaptic excitation, one may ask how a spontaneous activity state with low firing rates can be realized, and how the firing rates of persistent activity can be controlled within a physiological range (20–40 Hz), in spite of such an explosive positive feedback. We found that the dynamic stability of both states depends critically on the predominant contribution of NMDARs to the recurrent synaptic excitation, and on a strong inhibition that overall dominates the recurrent synaptic circuit. Strong recurrent excitation between nearby cells (with similar preferred cues), in interplay with recurrent inhibition, produces a structured network activity profile of persistent activity, which gives rise to 'memory fields' in individual neurons. Neuronal firing properties in both spontaneous and selective persistent states are found to be in agreement with single-neuron recording data from the PFC of the behaving monkey (Funahashi *et al.*, 1989; Chafee and Goldman-Rakic, 1998; Rao *et al.*, 1999).

The study of pattern formation in neural models has a long history (Wilson and Cowan, 1972, 1973; Amari, 1977; Ben-Yishai *et al.*, 1995; Skaggs *et al.*, 1995; Somers *et al.*, 1995; Tsodyks and Sejnowski, 1995; Redish *et al.*, 1996; Seung, 1996; Zhang, 1996; Bressloff and Coombes, 1998; Camperi and Wang, 1998; Hansel and Sompolinsky, 1998) [reviewed recently by Ermentrout (Ermentrout, 1998)]. In most of these studies, spatially tuned activity patterns ('bump states') appear through a continuous ('Turing') bifurcation; therefore they do not coexist with the resting state. On the other hand, to fulfil a working memory function, a PFC network should display bistability (or multistability) between the resting state and structured activity states, so that the network can be switched on and off between the two by transient inputs (Amit and Brunel, 1997; Camperi and Wang, 1998). The main conceptual novelty of the present work is to build a network of spiking neurons that shows bistability between two different types of active states: a resting state with spontaneous firing rates of a few Hertz, and a spatially structured state with firing rates of ~20–30 Hz (comparable to the physiological data). This property is mainly brought about by the dominance of recurrent synaptic inputs by the GABAergic contribution; and the network is stabilized by NMDARs at the recurrent synapses. Both features were not present in previous pattern formation studies.

A Mechanism for Switching Off Working Memory

Persistent activity is usually turned off following a transient increase of neuronal firing during the response period [see e.g. Fig. 3 of (Funahashi *et al.*, 1989), and Fig. 16 of (Goldman-Rakic *et al.*, 1990)]. Our simulations show that a simple way to turn off persistent activity is to increase transiently the external excitatory inputs to a large neural population of the network. These transient inputs increase the firing rates of pyramidal cells as well as interneurons. The increase in recurrent inhibitory inputs switches off the bump state. Such a mechanism is plausible given the available data. However, our model does not address the specific neuronal source of the input signal for memory erasure.

In a model of one-population spiking neurons without noise, Laing *et al.* also use an excitatory pulse to switch off a bump to a silent state (Laing *et al.*, 2000). However, this is achieved in their network through a quite different mechanism, i.e. by synchronizing all cells so that the persistent activity destabilizes.

Physiological data show that activity during the response period in PFC has two distinct phases (Funahashi *et al.*, 1991; Rao *et al.*, 1999). Earlier activity (in the 'pre-saccadic' period) is primarily tuned to the direction of the cue, while later activity (in the 'post-saccadic' period) is in many cases tuned to the opposite direction (especially for interneurons). Our model accounts for the tuning properties observed during the pre-saccadic period, since the tuned network activity takes some time, of the order a few hundred of milliseconds, to vanish. On the other hand, it does not account for the inversion of tuning reported in experiments during the post-saccadic period. A detailed modeling of the PFC activity during the saccade is outside the scope of the present paper.

Random Drift of Memory

Persistent network activity that encodes an analog quantity typically displays random drifts in time, because the activity pattern is marginally stable (Ben-Yishai *et al.*, 1995; Seung, 1996; Zhang, 1996; Lee *et al.*, 1997; Camperi and Wang, 1998). This has the consequence that the memory of the cue will become less and less precise as time goes by. Our simulations show that though this effect can be important in small networks, it becomes less pronounced in large networks. However, these simulations were performed under two assumptions. First, with an increased network size the recurrent coupling is normalized by the number of cells, so as to maintain a fixed average recurrent drive to the cells. As a result, the signal-to-noise ratio of the input to a cell decreases with the network size. This would not be the case for a sparsely connected network where, when network size is varied, both the number of synapses per cell and the strength of each individual synapse could remain constant. Second, we assumed that noise in the input is uncorrelated from cell to cell. If significant correlations are present between noise signals in different cells, drifts are likely to occur even in a very large network.

The magnitude of random drifts in a realistic working memory circuit storing an analog variable is therefore still an open question. Our study predicts that, independently of the amplitude of the random drift, the variance of the distance between the bump location and the stimulus location increases linearly with time, as for a diffusion process. This implies that, in a visuospatial delayed-response task, the variance of the distance between the cue and the eye position following the saccade to the memorized cue position should increase linearly with the delay time interval. In a psychophysical study using a visuospatial delayed-response task, White *et al.* plotted the scatters of eye positions following the saccade versus delay time in monkeys [(White *et al.*, 1994), see their Fig. 5*B*]. The data in this figure (squared in order to get the saccade error variance) can be fitted by a straight line, similarly to the drifting mechanism that occurs in our model, for delay times up to 4 s. Similar data are also

available for humans (Ploner *et al.*, 1998), showing also a linearly increasing saccade error variance up to 20 s delay times. It is thus possible that such a gradual loss of accuracy in memory-guided saccade is a manifestation of slow random drifts of the persistent activity in PFC during the delay period.

NMDA Contribution to Recurrent Synapses: Implications for Synchronous Network Behavior and Stability of Working Memory

Our simulations show that the network dynamics are critically dependent on the ratio of NMDAR and AMPAR channels at recurrent synapses. When NMDAR channels dominate, persistent activity is stable at physiological rates (20–40Hz), and the network dynamics are essentially asynchronous. With a substantial AMPA component of the recurrent excitation, the network displays coherent oscillations; if the AMPAR-mediated recurrent excitation is too large, persistent activity is abolished. This result was initially obtained with a spatially unstructured network model (Wang, 1999). Here, we found that the same conclusion holds for a spatially structured network as well. Our working memory model requires that at recurrent synapes the NMDA receptors should contribute >65% of the charge entry by a unitary EPSC at –65 mV. However, the precise value of the required NMDA:AMPA ratio is likely to depend on the details of the model, as well as the type of neuron models (e.g. integrate-and-fire model versus compartmental conductance-based model).

The relative contributions by NMDARs and AMPARs to charge entry of a unitary EPSC remain unknown for intrinsic synapses of the PFC. Estimates from other cortical areas vary considerably. For example, the NMDA component contributes 17% to the EPSP integral (at –60 mV) for pyramidal cells in layer 5 somatosensory cortex (Markram *et al.*, 1997), and 65% of EPSC's charge entry (at –65 mV) for hippocampal pyramidal cells of the young rat (Spruston *et al.*, 1995). At intrinsic synapses of the layer 4 somatosensory cortex, NMDA receptors contribute 39% to the EPSP integral (at –60 mV) in the young rat (Feldmeyer *et al.*, 1999), and >90% of EPSC's charge entry (at –70 mV) in the mouse (Fleidervish *et al.*, 1998). Further studies of this issue would be worthwhile.

The crucial features of NMDAR-mediated transmission for stable persistent activity in our model are its slow synaptic kinetics [for stability with respect to synchronized oscillations (Wang, 1999; Laing *et al.*, 2000)] and its saturation properties [for robust low persistent rates, see also (Wang, 1999)]. On the other hand, the voltage-dependence of the NMDA conductance due to magnesium block is not crucial here. However, the voltage-gating of the NMDA current could conceivably contribute to selectivity of persistent activity in a neural assembly. This is because during a cue presentation the cells that are tuned to the cue stimulus are more active, and their membrane potential is more depolarized, than those that are not tuned to the cue. Therefore, the NMDA conductance should be differentially unblocked in those cells that are excited by the cue stimulus (Lisman *et al.*, 1998).

Thus, the model predicts that the long decay time constant of NMDAR-mediated synaptic transmission is critically important to the persistent activity underlying working memory function of PFC. This conclusion is supported by behavioral experiments with rats performing a spatial delayed alternation task, where it was found that systemic administration (Verma and Moghaddam, 1996) or microinjection into the PFC (Romanides *et al.*, 1999) of NMDAR antagonists in PFC impaired working memory. However, in these studies, pharmacological manipulation has not been combined with physiological recordings from PFC neurons. Thus, changes in the mnemonic neuronal dynamics caused by the NMDAR blockade are presently unknown. Furthermore, evidence suggests that dysfunction of NMDAR-mediated synaptic transmission may lead to working memory deficits similar to those observed in schizophrenia (Javitt and Zukin, 1991; Krystal *et al.*, 1994; Akbarian *et al.*, 1996). Our model study identified a candidate mechanism through which working memory relies on NMDARs at recurrent synapses of PFC, namely asynchronous firing resulting in persistent activity stability. Moreover, according to the hypothesis that dopamine differentially modulates NMDAR-mediated synaptic transmission (Cepeda *et al.*, 1992), a malfunction of the dopaminergic innervation of PFC would also give rise to working memory deficits, as has been shown by many studies in humans and animals (Sawaguchi *et al.*, 1990; Sawaguchi and Goldman-Rakic, 1991; Goldman-Rakic, 1994; Okubo *et al.*, 1997; Arnsten, 1998).

Tuning in Inhibitory Cells

Recently, Rao and co-workers reported memory fields in interneurons as well as in pyramidal cells (Rao *et al.*, 1999). In our model, a structured connectivity from pyramidal cells to interneurons leads to a persistent state with tuned interneurons as well as pyramidal cells. This result suggests that the selectivity of persistent activity in putative inhibitory neurons observed experimentally (Rao *et al.*, 1999) may be explained by a structured pyramid-to-interneuron connectivity that is similar to the pyramid-to-pyramid connectivity. However, we have been unable to stabilize a network state in which the tuning is identical between interneurons and pyramidal cells. Thus, we would predict that pyramidal cells are more sharply tuned than interneurons in the delay period.

Maintaining Working Memory in the Face of Distractors

An important property of working memory is its ability to resist distractions. Neurophysiological studies of working memory in the associative cortex of monkeys have shown that delay activity is typically resistant to distractors in the PFC (Miller *et al.*, 1996), though it is not in areas which are closer to primary sensory areas such as PPC (Constantinidis and Steinmetz, 1996) or inferotemporal cortex (Miller *et al.*, 1996). A main result of the present study is that a network model of spatial working memory with strong recurrent inhibition is intrinsically resistant to distractors, provided stimulus intensities are low. On the other hand, persistent activity is disrupted by distractors when stimuli are strong. Moreover, the network is less distractable with enhanced NMDAR-mediated recurrent excitation and feedback inhibition. Thus, we propose two candidate factors that may explain the areal differences observed with respect to response to distractors. In areas close to primary sensory areas, stimulus intensities are probably stronger, as indicated by the magnitude of visual responses during cue presentation, and delay activity would be easily disrupted by distractors. In the PFC, which is further away from sensory areas, direct afferent inputs are likely to be weaker, hence persistent activity of PFC neurons is more resistant to distractors. Alternatively, the PFC circuit may be uniquely equipped with an optimal balance between NMDAR-mediated excitation and recurrent inhibition, while other cortical areas are not optimal in that respect. This would endow the PFC with an exceptional ability to hold behaviorally relevant information on-line, in spite of external distractions.

Experimentally Testable Predictions

The present work raises a number of mechanistic issues about working memory processes that could be addressed experimentally.

(i) *In vitro* slice experiments can be carried out to investigate whether NMDARs indeed dominate recurrent synaptic excitation in PFC, as suggested by the model.

(ii) In experiments with behaving animals, a combination of pharmacology with single-neuron recordings, similar to the iontophoresis experiments of other workers (Williams and Goldman-Rakic, 1995; Rao *et al.*, 2000), would elucidate how working memory performance is affected by modulation or blockade of NMDARs in PFC; and what are the concomitant changes in the neuronal persistent activity.

(iii) An open question is whether neuronal firings in a persistent activity state during the delay period are asynchronous. Alternatively, neurons may be partially synchronized and/or display coherent oscillations. This question can be addressed with simultaneous recordings from multiple neurons, and using local field recordings from the monkey PFC to probe population activity in a delayed-response task. Moreover, we predict that the propensity of a network to display coherent oscillations is higher if the relative contribution of NMDAR channels to recurrent excitation is smaller.

(iv) As mentioned above, memory-guided saccadic response has been found to be less accurate with longer delay periods (White *et al.*, 1994; Ploner *et al.*, 1998). It would be interesting to study more systematically whether the loss of accuracy in the saccade is correlated with slow random drifts of persistent neuronal activity in the PFC during the delay period.

(v) Distractor experiments similar to those of Miller (Miller *et al.*, 1996) could be performed in combination with pharmacological manipulation of synaptic transmission. We predict that a cooperative facilitation of recurrent inhibition and NMDAR-mediated recurrent excitation would enhance the network's ability to resist distractors and preserve the memory of behaviorally relevant information in spite of intervening stimuli.

Notes

This work was supported by NSF (IBN-9733006), the A.P. Sloan Foundation, the W.M. Keck Foundation and CNRS. We thank J. Tegnér for useful discussions.

Address correspondence to Xiao-Jing Wang, Volen Center for Complex Systems, MS 013, Brandeis University, 415 South Street, Waltham, MA 02254–9110, USA. Email: xjwang@brandeis.edu.

References

Akbarian S, Sucher NJ, Bradley D, Tazzofoli A, Trinh D, Hetrick WP, Potkin SG, Sandman C,A Bunney WE, Jones EG (1996) Selective alterations in gene expression for NMDA receptor subunits in prefrontal cortex of schizophrenics. J Neurosci 16:19–30.

Amari, S (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. Biol Cybernet 27:77–87.

Amit DJ (1995) The Hebbian paradigm reintegrated: local reverberations as internal representations. Behav Brain Sci 18:617.

Amit DJ, Brunel N (1997) Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. Cereb Cortex 7:237–252.

Amit D, Brunel N, Tsodyks MV (1994) Correlations of cortical Hebbian reverberations: experiment vs theory. J Neurosci 14:6435–6445.

Arnsten AFT (1998) Catecholamine modulation of prefrontal cortical cognitive function. Trends Cogn Sci 2:436–447.

Ben-Yishai R, Lev Bar-Or R, Sompolinsky H (1995) Theory of orientation tuning in visual cortex. Proc Natl Acad Sci USA 92:3844–3848.

Berg HC (1983) Random walks in biology. Princeton, NJ: Princeton University Press.

Bressloff PC, Coombes S (1998) Spike train dynamics underlying pattern formation in integrate-and-fire oscillator networks. Phys Rev Lett 81:2384–2387.

Camperi M, Wang X-J (1998) A model of visuospatial short-term memory in prefrontal cortex: recurrent network and cellular bistability. J Comput Neurosci 5:383–405.

Cepeda C, Radisavljevic Z, Peacock W, Levine MS, NA Buchwald (1992) Differential modulation by dopamine of responses evoked by excitatory amino acids in human cortex. Synapse 11:330–341.

Chafee MV, Goldman-Rakic PS (1998) Neuronal activity in macaque prefrontal area 8a and posterior parietal area 7ip related to memory guided saccades. J Neurophysiol 79:2919–2940.

Colby CL, Duhamel JR, Goldberg ME (1996) Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. J Neurophysiol 76:2841–2852.

Constantinidis C, Steinmetz MA (1996) Neuronal activity in posterior parietal area 7a during the delay periods of a spatial memory task. J Neurophysiol 76:1352–1355.

Constantinidis C, Franowicz MN, Goldman-Rakic PS (1999) Multiple electrode analysis of local circuitry in the primate prefrontal cortex during spatial working memory. Soc Neurosci Abstr 25:44.1.

Courtney SM, Petit L, Maisog JM, Ungerleider LG, Haxby JV (1998) An area specialized for spatial working memory in human frontal cortex. Science 279:1347–1351.

Douglas RJ, Koch C, Mahowald M, Martin KM, Suarez HH (1995) Recurrent excitation in neocortical circuits. Science 269:981–985.

Durstewitz D, Kelc M, Güntürkun O (1999) A neurocomputational theory of the dopaminergic modulation of working memory functions. J Neurosci 19:2807–2822.

Ermentrout GB (1998) Neural networks as spatio-temporal pattern-forming systems. Rep Progr Phys 61:353–430.

Feldmeyer D, Egger V, Lubke J, Sakmann B (1999) Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single 'barrel' of developing rat somatosensory corte. J Physiol 521:169–190.

Fleidervish IA, Binshtok AM, Gutnick MJ (1998) Functionally distinct NMDA receptors mediate horizontal connectivity within layer 4 of mouse barrel cortex. Neuron 21:1055–1065.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 61:331–349.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1991) Neuronal activity related to saccadic eye movements in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 65:1464–1483.

Fuster JM (1973) Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. J Neurophysiol 36:61–78.

Fuster JM (1988) The prefrontal cortex. New York: Raven Press.

Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. Science 233:1416–1419.

Gnadt JW, Andersen RA (1988) Memory related motor planning activity in posterior parietal cortex of macaque. Exp Brain Res 70:216–220.

Goldman-Rakic PS (1987) Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In: Handbook of physiology, vol. V: The nervous system (Plum F, Mountcastle V, eds), chap. 9, pp. 373–417. Bethesda, MD: American Physiological Society.

Goldman-Rakic PS (1994) Working memory dysfunction in schizophrenia. J Neuropsychol Clin Neurosci 6:348–357.

Goldman-Rakic PS (1995) Cellular basis of working memory. Neuron 14:477–485.

Goldman-Rakic PS, Funahashi S, Bruce CJ (1990) Neocortical memory circuits. In: Cold Spring Harbor Symposia on Quantitative Biology, vol. LV, pp. 1025–1038.

González-Burgos G, Barrionuevo G, Lewis DA (2000) Horizontal synaptic connections in monkey prefrontal cortex: an *in vitro* electrophysiological study. Cereb Cortex 10:82–92.

Hansel D, Sompolinsky H (1998) Modeling feature selectivity in local cortical circuits. In: Methods in neuronal modeling, 2nd edn (Koch C, Segev I, eds). Cambridge, MA: MIT Press.

Hansel D, Mato G, Meunier C, Neltner (1998) On numerical simulations of integrate-and-fire neural networks. Neural Comput 10:467–483.

Hebb DO (1949) Organization of behavior. New York: Wiley.

Hestrin S, Sah P, Nicoll R (1990) Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. Neuron 5:247–253.

Jahr CE, Stevens CF (1990) Voltage dependence of NMDA-activated

macroscopic conductances predicted by single-channel kinetics. J Neurosci 10:3178–3182.
Javitt DC, Zukin SR (1991) Recent advances in the phencyclidine model of schizophrenia. Am J Psychiat 148:1301–1308.
Kritzer MF, Goldman-Rakic PS (1995) Intrinsic circuit organization of the major layers and sublayers of the dorsolateral prefrontal cortex in the rhesus monkey. J Comp Neurol 359:131–143.
Krystal JH, Karper LP, Seibyl JP, Freeman GK, Delaney R, Bremner JD, Heninger GR, Bowers MB, Charney DS (1994) Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. Arch Gen Psychiat 51:199–214.
Laing CR, Chow CC, Ermentrout GB (2000) Stationary bumps in networks of spiking neurons. Preprint, University of Pittsburgh.
Lee DD, Reis BY, Seung HS, Tank DW (1997) Nonlinear network models of the oculomotor integrator. In: Computational neuroscience (Bower J, ed.). New York: Plenum Press.
Levitt B, Lewis DA, Yoshioka T, Lund J (1993) Topography of pyramidal neuron intrinsic connections in macaque monkey prefrontal cortex (areas 9 and 46). J Comp Neurol 338:360–376.
Lisman JE, Fellous J-M, Wang X-J (1998) A role for NMDA-receptor channels in working memory. Nature Neurosci 1:273–275.
Markram H, Lubke J, Frotscher M, Roth A, Sakmann B (1997) Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. J Physiol 500:409–440.
McCarthy G, Blamire AM, Puce A, Nobre AC, Bloch G, Hyder F, Goldman-Rakic PS, Shulman RG (1994) Functional magnetic resonance imaging of human prefrontal cortex activation during a spatial working memory task. Proc Natl Acad Sci USA 91:8690–8694.
McCormick D, Connors B, Lighthall J, Prince D (1985) Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons in the neocortex. J Neurophysiol 54:782–806.
Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci 16:5154–5167.
Moody SL, Wise SP, di Pellegrino G, Zipser D (1998) A model that accounts for activity in primate frontal cortex during a delayed matching-to-sample task. J Neurosci 18:399–410.
Mountcastle VB (1997) The cortical organization of the neocortex. Brain 120:701–722.
Niki H, Watanabe M (1976) Prefrontal unit activity and delayed response: relation to cue location versus direction of response. Brain Res 105:79–88.
Okubo Y, Suhara T, Suzuki K, Kobayashi K, Inoue O, Terasaki O, Someya Y, Sassa T, Sudo Y, Matsushima E, Iyo M, Tateno Y, Toru M (1997) Decreased prefrontal dopamine D1 receptors in schizophrenia revealed by PET. Nature 385:634–636.
Ó Scalaidhe SP, Goldman-Rakic PS (1999) Functional organization of prefrontal cortical columns in the macaque. Soc Neurosci Abstr 25:620.2.
Ploner CJ, Gaymard B, Rivaud S, Agid Y, Pierrot-Deseilligny C (1998) Temporal limits of spatial working memory in humans. Eur J Neurosci 10:794–797.
Rainer G, Assad WF, Miller EK (1998) Memory fields of neurons in the primate prefrontal cortex. Proc Natl Acad Sci USA 95:15008–15013.
Rao SG, Williams GV, Goldman-Rakic PS (1999) Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in PFC. J Neurophysiol 81:1903–1916.
Rao SG, Williams GV, Goldman-Rakic PS 2000 Destruction and creation of spatial tuning by disinhibition: $GABA_A$ blockade of prefrontal cortical neurons engaged by working memory. J Neurosci 20:485–494.
Redish AD, Elga AN, Touretzky DS (1996) A coupled attractor model of the rodent head direction system. Network 7:671–685.
Romanides AJ, Duffy P, Kalivas PW (1999) Glutamatergic and dopaminergic afferents to the prefrontal cortex regulate spatial working memory in rats. Neuroscience 92:97–106.
Salin PA, Prince DA (1996) Spontaneous $GABA_A$ receptor mediated inhibitory currents in adult rat somatosensory cortex. J Neurophysiol 75:1573–1588.
Sarnthein J, Petsche H, Rappelsberger P, Shaw GL, von Stein A (1998) Synchronization between prefrontal and posterior association cortex during human working memory. Proc Natl Acad Sci USA 95: 7092–7096.
Sawaguchi T, Goldman-Rakic PS (1991) D1 dopamine receptors in prefrontal cortex: involvement in working memory. Science 251: 947–950.
Sawaguchi T, Yamane I (1999) Properties of delay-period neuronal activity in the monkey dorsolateral prefrontal cortex during a spatial delayed matching-to-sample task. J Neurophysiol 82:2070–2080.
Sawaguchi T, Matsumara M, Kubota K (1990) Catecholaminergic effects on neuronal activity related to a delayed response task in monkey prefrontal cortex. J Neurophysiol 63:1385–1400.
Seung HS (1996) How the brain keeps the eyes still. Proc Natl Acad Sci USA 93:13339–13344.
Shadlen MN, Newsome WT (1994) Noise, neural codes and cortical organization. Curr Opin Neurobiol 4:569.
Skaggs WE, Knierim JJ, Kudrimoti HS, McNaughton BL (1997) A model of the neural basis of the rat's sense of direction. In: Advances in neural information processing systems 7 (Tesauro G, Touretzky DS, Leen TK, eds). Cambridge, MA: MIT Press.
Somers DC, Nelson SB, Sur M (1995) An emergent model of orientation selectivity in cat visual cortical simple cells. J Neurosci 15:5448–5465.
Spruston N, Jonas P, Sakmann B (1995) Dendritic glutamate receptor channel in rat hippocampal CA3 and CA1 pyramidal neurons. J Physiol 482:325–352.
Troyer TW, Miller KD (1997) Physiological gain leads to high ISI variability in a simple model of a cortical regular spiking cell. Neural Comput 9:971–983.
Tsodyks MV, Sejnowski T (1995) Rapid state switching in balanced cortical network models. Network 6:111–124.
Tuckwell HC (1988) Introduction to theoretical neurobiology. Cambridge: Cambridge University Press.
van Vreeswijk C, H Sompolinsky (1998) Chaotic balanced state in a model of cortical circuits. Neural Comput 10:1321–1371.
Verma A, Moghaddam B (1996) NMDA receptor antagonists impair prefrontal cortex function as assessed via spatial delayed alternation performance in rats: modulation by dopamine. J Neurosci 16:373–379.
Wang X-J (1999) Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. J Neurosci 19:9587–9603.
White JM, Sparks DL, Stanford TR (1994) Saccades to remembered target locations: and analysis of systematic and variable errors. Vis Res 34:79–92.
Williams GV and Goldman-Rakic PS (1995) Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. Nature 376: 572–575.
Wilson HR, Cowan JD (1972) Excitatory and inhibitory interactions in localized populations of model neurons. Biophys J 12:1–24.
Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik 13:55–80.
Xiang Z, Huguenard JR, Prince DA (1998) $GABA_A$ receptor mediated currents in interneurons and pyramidal cells of rat visual cortex. J Physiol 506:715–730.
Zarahn E, Aguirre GK, D'Esposito M (1999) Temporal isolation of the neural correlates of spatial mnemonic processing with fMRI. Brain Res Cogn Brain Res 7:255–268.
Zhang K (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensembles: a theory. J Neurosci 16:2112–2126.
Zipser D, Kehoe B, Littlewort G, Fuster J (1993) A spiking network model of short-term active memory. J Neurosci 13:3406–3420.