Visual Working Memory: Neuronal Dynamics in Prefrontal Cortex

Gustavo Deco[a] and Edmund T. Rolls[b]

[a] *Institució Catalana de Recerca i Estudis Avançats (ICREA)*
Universitat Pompeu Fabra, Barcelona, Spain

[b] *University of Oxford, Dept. of Experimental Psychology, Oxford, England*

Cognitive behavior requires complex context-dependent mapping between sensory stimuli and actions. The same stimulus can lead to different behaviors depending on the situation, or the same behavior may be elicited by different cueing stimuli. Recordings have revealed the existence of neurons in the primate prefrontal cortex showing task-specific firing activity during working memory delay periods. These neurons provide a neural substrate for rule-dependent associations between particular cues and responses, i.e. for mapping stimulus and response in a flexible, context-dependent, fashion. We describe here a computational neuroscience model to explain and investigate the different types of working-memory related neuronal activity observed in the primate prefrontal cortex. Processes occurring at the AMPA, NMDA and GABA synapses are dynamically modelled in the integrate-and-fire implementation to produce realistic spiking dynamics. We assume a hierarchically organized set of different attractor network pools in the dorsolateral prefrontal cortex (PFC). The hierarchical structure is organized within the general framework of the biased competition model of attention. The effects of neuromodulation by dopamine of the synaptic processes show not only that dopamine weakens and shortens the persistent neuronal activity in the delay period, but also leads to the prediction of more response errors when dopamine is elevated because there is less different activity in the different pools of competing neurons resulting in more conflict.

1. Introduction

There is much evidence that the prefrontal cortex is involved in at least some types of working memory and related processes such as planning [1]. Working memory refers to an active system for maintaining and manipulating information in mind, held during a short period, usually of seconds. Neuronal recording studies reveal high spiking rates in prefrontal neurons during the execution of working memory tasks [2]. Recently, [3] investigated the functions of the prefrontal cortex in working memory by analyzing neuronal activity when the monkey performs two different working memory tasks using the same stimuli and responses. In a *conditional object-response (associative) task* with a delay the monkey was shown one of two stimuli, and after a delay had to make either a rightward or leftward oculomotor saccade response depending on which stimulus was shown. In another experiment, recordings were made both during the object-response task and during a *delayed spatial response task*, in which the same stimuli were used, but the rule required was different, namely to respond towards the location where the stimulus had been shown [3]. The main motivation for such studies was the fact that for real-world behavior, the mapping between a stimulus and a response is typically more complicated than a one-to-one mapping. The same stimulus can lead to different behaviors depending on the situation, or the same behavior may be elicited by different cueing stimuli. In the performance of these tasks neurons have been described that respond in the delay period to the stimulus object, the stimulus position ("sensory pools"), to combinations of the response and the stimulus object or position ("intermediate pools"), and to the response required (left or right) ("premotor pools").

The aim of the present work is to model, and therefore help to understand, the underlying mechanisms that implement the working memory-related activity observed in neurons in the primate PFC in the context-dependent stimulus-response (associative) and delayed spa-

tial response tasks investigated by [3]. The model builds on the integrate-and-fire attractor network treatment of [4] by introducing a hierarchically organized set of different attractor network pools in the dorsolateral prefrontal cortex (PFC). The hierarchical structure is organized within the general framework of the biased competition model of attention [5,6]. The general approach was developed further by [6], who introduced a neurodynamical theoretical framework for Biased Competition in the context of visual attention.

2. Methods

The basic circuit of an integrate-and-fire model consists of the cell membrane capacitance C_m in parallel with the cell membrane resistance R_m driven by a synaptic current (excitatory or inhibitory post-synaptic potential, EPSP or IPSP, respectively). If the voltage across the capacitor reaches a threshold θ the circuit is shunted and a δ-pulse (spike) is generated and transmitted to other neurons. The incoming presynaptic δ-pulse from other neurons is basically low-pass filtered first by the synaptic and membrane time constants, before it is utilized as an EPSP or IPSP in the one-compartment neuronal model. We use biologically realistic parameters. We assume for both kinds of neuron a resting potential $V_L = -70$ mV, a firing threshold $\theta = -50$ mV, and a reset potential $V_{\text{reset}} = -55$ mV. The membrane capacitance C_{m} is 0.5 nF for the pyramidal neurons and 0.2 nF for the interneurons. The membrane leak conductance g_{m} is 25 nS for pyramidal cells, and 20 nS for interneurons. The refractory period τ_{ref} is 2 ms for pyramidal cells, and 1 ms for interneurons. Consequently, the membrane time constant $\tau_{\text{m}} = C_{\text{m}}/g_{\text{m}}$ is 20 ms for pyramidal cells, and 10 ms for interneurons. More specifically, the subthreshold membrane potential $V(t)$ of each neuron evolves according to the following equation:

$$C_{\text{m}}\frac{dV(t)}{dt} = -g_{\text{m}}(V(t) - V_L) - I_{\text{syn}}(t) \quad (1)$$

where $I_{\text{syn}}(t)$ is the total synaptic current flow into the cell.

The synaptic current flows into the cells are mediated by three different families of receptors. The total synaptic current is given by the sum of glutamatergic excitatory components (NMDA and AMPA) and inhibitory components (GABA, I_{G}). We consider that external excitatory contributions are produced through AMPA receptors (I_{Ae}), while the excitatory recurrent synapses are produced through AMPA and NMDA receptors (I_{A} and I_{N}). The total synaptic current is therefore given by:

$$I_{\text{syn}}(t) = I_{\text{Ae}}(t) + I_{\text{A}}(t) + I_{\text{N}}(t) + I_{\text{G}}(t) \quad (2)$$

where

$$I_{\text{Ae}}(t) = g_{\text{Ae}}(V(t) - V_E) \sum_{j=1}^{N_e} s_j^{\text{Ae}}(t) \quad (3)$$

$$I_{\text{A}}(t) = g_{\text{A}}(V(t) - V_E) \sum_{j=1}^{N_E} w_j s_j^{\text{A}}(t) \quad (4)$$

$$I_{\text{N}}(t) = \frac{g_{\text{N}}(V(t) - V_E)}{(1 + C_{\text{Mg}++} e^{(\frac{-0.062V(t)}{3.57})})} \sum_{j=1}^{N_E} w_j s_j^{\text{N}}(t) \quad (5)$$

$$I_{\text{G}}(t) = g_{\text{G}}(V(t) - V_I) \sum_{j=1}^{N_I} s_j^{\text{G}}(t) \quad (6)$$

In the preceding equations $V_E = 0$ mV and $V_I = -70$ mV. The fractions of open channels s are given by:

$$\frac{ds_j^{\text{Ae}}(t)}{dt} = -\frac{s_j^{\text{Ae}}(t)}{\tau_{\text{A}}} + \sum_k \delta(t - t_j^k) \quad (7)$$

$$\frac{ds_j^{\text{A}}(t)}{dt} = -\frac{s_j^{\text{A}}(t)}{\tau_{\text{A}}} + \sum_k \delta(t - t_j^k) \quad (8)$$

$$\frac{ds_j^{\text{N}}(t)}{dt} = -\frac{s_j^{\text{N}}(t)}{\tau_{\text{N,d}}} + \alpha x_j(t)(1 - s_j^{\text{N}}(t)) \quad (9)$$

$$\frac{dx_j(t)}{dt} = -\frac{x_j(t)}{\tau_{\text{N,r}}} + \sum_k \delta(t - t_j^k) \quad (10)$$

$$\frac{ds_j^{\mathrm{G}}(t)}{dt} = -\frac{s_j^{\mathrm{G}}(t)}{\tau_{\mathrm{G}}} + \sum_k \delta(t - t_j^k) \qquad (11)$$

where the sums over k represent a sum over spikes emitted by presynaptic neuron j at time t_j^k. The value of $\alpha = 0.5\mathrm{ms}^{-1}$.

The values of the conductances (in nS) for pyramidal neurons were: $g_{\mathrm{Ae}} = 2.08$, $g_{\mathrm{A}} = 0.052$, $g_{\mathrm{N}} = 0.164$ and $g_{\mathrm{G}} = 1.13$; and for interneurons: $g_{\mathrm{Ae}} = 1.62$, $g_{\mathrm{A}} = 0.0405$, $g_{\mathrm{N}} = 0.129$ and $g_{\mathrm{G}} = 0.87$. We consider that the NMDA currents have a voltage dependence that is controlled by the extracellular magnesium concentration, $[\mathrm{Mg}^{++}] = 1$ mM. We neglect the rise time of both AMPA and GABA synaptic currents, because they are typically extremely short (< 1 ms). The rise time for NMDA synapses is $\tau_{\mathrm{N,r}} = 2$ ms. The decay time for AMPA synapses is $\tau_{\mathrm{A}} = 2$ ms, for NMDA synapses $\tau_{\mathrm{N,d}} = 100$ ms, and for GABA synapses $\tau_{\mathrm{G}} = 10$ ms.

Figure 1 shows schematically the synaptic structure assumed in the prefrontal cortical network. The network is composed of N_E (excitatory) pyramidal cells and N_I inhibitory interneurons. In our simulations, we use $N_E = 1600$ and $N_I = 400$. The neurons are fully connected. There are different populations or pools of neurons in the prefrontal cortical network. Each pool of excitatory cells contains fN_E neurons, where f the fraction of the neurons in any one pool was set to be 0.05. There are four excitatory pools, namely: sensory, task or rule-specific, premotor, and nonselective. The sensory pools encode information about objects, or spatial location. The premotor pools encode the motor response (in our case the leftward or rightward oculomotor saccade). The intermediate pools are task-specific and perform the mapping between the sensory stimuli and the required motor response. The intermediate pools respond to combinations of the sensory stimuli and the response required, e.g. to object 1 requiring a left oculomotor saccade. The intermediate pools receive an external biasing input that reflects the current rule (e.g. on this trial when object 1 is shown make the left response after the delay period). The remaining excitatory neurons do not have specific sensory, response or biasing inputs, and are in a nonselective pool. All

the inhibitory neurons are clustered into a common inhibitory pool, so that there is global competition throughout the network.
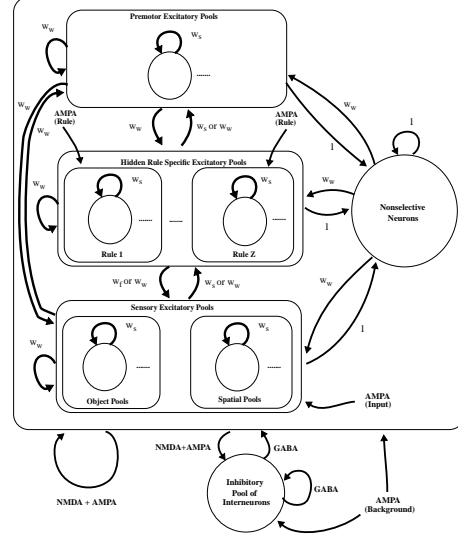

Figure 1: Prefrontal Cortical Module.

We assume that the synaptic coupling strengths between any two neurons in the network are established by Hebbian learning. As a consequence of this, neurons within a specific excitatory pool are mutually coupled with a strong weight $w_s = 2.1$. Neurons in the inhibitory pool are mutually connected with an intermediate weight $w = 1$. They are also connected with all excitatory neurons with the same intermediate weight $w = 1$. The connection strength between two neurons in two different specific excitatory pools is weak and given by $w_w = 1 - 2f(w_s - 1)/(1 - 2f)$ (=0.8778) unless otherwise specified. Neurons in a specific excitatory pool are connected to neurons in the nonselective pool with a feedforward synaptic weight $w = 1$ and a feedback synaptic connection of weight w_w. The connections between the different pools are set up to achieve the required mapping from the sensory input pools thorough the intermediate pools to the premotor pools, assuming Hebbian learning based on the activity of individual pools while the different tasks are being performed. The forward connections (input

4

to intermediate to output pools) are $w_s = 2.1$. The corresponding feedback synaptic connections are slightly weaker ($w_f = 1.7$ for the feedback synapses between rule-specific and sensory pools, and w_w for the feedback synapses between the premotor and rule-specific pools).

Each neuron (pyramidal cells and interneurons) receives $N_{\text{ext}} = 800$ excitatory AMPA synaptic connections from outside the network. These connections provide three different type of external interactions: 1) a background noise due to the spontaneous firing activity of neurons outside the network; 2) a sensory related input; and 3) a rule or context-related bias input that specifies the task. The external inputs are given by a Poisson train of spikes. In order to model the background spontaneous activity of neurons in the network [4], we assume that Poisson spikes arrive at each external synapse with a rate of $\nu_{\text{ext}} = 3$ Hz, consistent with the spontaneous activity observed in the cerebral cortex.The sensory input is encoded by increasing the external input Poisson rate ν_{ext} to $\nu_{\text{ext}} + \lambda_{\text{input}}$ to the neurons in the appropriate specific sensory pools. We used $\lambda_{\text{input}} = 100$ Hz. Finally, the biasing specification of the context, i.e. which rule is active, is modelled by assuming that each neuron in each of the pools in the group of intermediate pools associated with the active task, receives external Poisson spikes with an increased rate from ν_{ext} to $\nu_{\text{ext}} + \lambda_{\text{rule}}$ throughout the trial. We use $\lambda_{\text{rule}} = 120$ Hz. This external top-down rule-specific input probably comes from the external prefrontal neurons which directly encode abstract rules [7], which in turn are influenced by the reward system (in the orbitofrontal cortex and amygdala) to enable the correct rule to be selected during for example reversal. During the last 100 ms of the response period, the external rate to all neurons is increased by a factor 1.5 in order to take into account the increase in afferent inputs due to behavioral responses and reward signals [4].

3. Results

We consider now the theoretical analyses of the single-cells recording of PFC neurons of [3], in which monkeys were trained to perform ei-

ther an object-response task with a delay, or a delayed spatial response task.The conditional object-response task with a delay is defined by the association of the identity of an object (O1 or O2) with a saccade response (L or R), independently of the location of the cue object. On the other hand, the spatial delayed response task required the monkey to make a saccade (L or R) response after a delay towards the location at which the cue object was presented (S1 or S2). Under this second condition the monkey had to ignore the feature characteristics of the object, and allocate its attention and memory to process the spatial location of the stimulus. Figure 2 presents the experimental results of [3]. The curves plot the averaged firing activity over time for three neurons beginning 500 ms before the cue presentation (which lasted for 500 ms) and during a delay period of 1000 ms. The left panel corresponds to the delayed spatial response task condition and the right panel to the conditional object-response condition. Each picture plots two curves corresponding to the two possible response directions, ipsilateral or contralateral. Row (a) shows a neuron that was response direction-selective in both tasks. Row (b) shows a neuron that was direction-selective in only the spatial task. Row (c) shows a neuron that was direction-selective in only the object-response task. These results demonstrated that the information represented by neurons in the lateral PFC of primates is not limited to discrete sensory events or motor responses, but instead that the behavioral context in which the animals were engaged had a decisive influence on the activity of some of the neurons. In particular, some populations of neurons respond to combinations of the task being performed and the response that is required. In this way, PFC neurons provide a neural substrate for responding appropriately on the basis of an abstract rule or context.
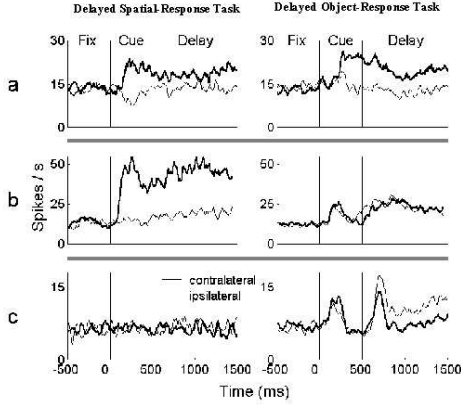
Figure 2: Experimental results of Assad et al. 2000.

We perform numerical simulations of the experiment of [3] by means of a prefrontal cortical architecture that includes two premotor pools of response neurons, one corresponding with leftward saccade responses (L) and the other corresponding to rightward saccade responses (R); four sensory pools (two object-selective neuronal pools, one corresponding to object O1 and the other corresponding to object O2; and two pools with selectivity for the spatial location of the stimulus, one corresponding to location S1 and the other corresponding to the location S2); and four intermediate neuronal pools, one for each of the four possible stimulus-response combinations. The intermediate pools are considered as being in two groups, one for the object-response associative task, and the other for the delayed spatial response task. Figure 3 plots the temporal evolution of the averaged population activity for three neural pools, namely the premotor pool 'L', the intermediate spatial pool 'S1-L', and the intermediate associative pool 'O1-L'. Cue, response and selective context-specific associative activity is explicitly maintained during the short-term memory related delay period by the recurrent connections. As in Figure 2, the left panel corresponds to the delayed spatial response condition and the right panel to the conditional object-response associative task condition. Each graph shows two curves corresponding to the two possible response directions (blue corresponds to L and red to R). The first row shows activity in the

premotor pool 'L' which was response direction-selective in both tasks. The second row shows activity in the intermediate spatial pool 'S1-L' that was response direction-selective (to the L, blue curve) in only the delayed spatial response task. The third row shows activity in the intermediate associative pool 'O1-L' that was direction-selective in only the conditional object-place associative task. All the three types of neurons found experimentally by [3] can be identified with pools in our prefrontal network.
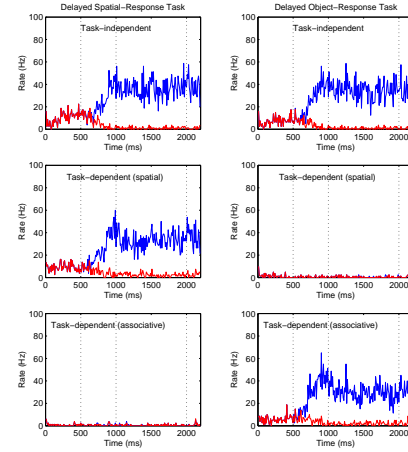

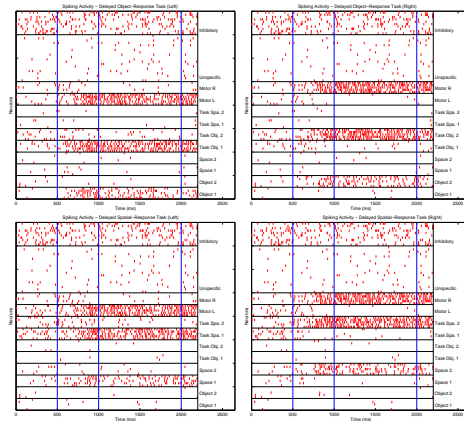Figure 3: Simulation results corresponding to the experimental paradigm of Assad et al. 2000.

Figure 4: Rastergrams for pools in the PFC network

Figure 4 plots the rastergrams of randomly selected neurons for each pool in the network. The

spatio-temporal spiking activity shows that during the short term memory delay period only the relevant sensory cue, associated future oculomotor response, and intermediate neurons maintain persistent activity and build up a stable global attractor in the network. The underlying biased competition mechanisms are very explicit in this experiment. Note that neurons in pools for the irrelevant input sensory dimension (location for the object-response associative task, and object for the delayed spatial response task), are inhibited during the sensory cue period and are not sustained during the delay short term memory period. Only the relevant single pool attractors, given the rule context, that are suitable for the cue-response mapping survive the competition and are persistently maintained with high firing activity during the short-term memory delay period. This suppression effect has been recently observed by by recording the activity of prefrontal neurons in monkeys carrying out a focused attention task [8] . In their spatial cueing task, they observed strong filtering of the PFC response to unattended targets. These attentional modulation effects are well-known in posterior areas of the visual system [5]. Our previous [6] and present computational simulations suggest that, in the PFC, filtering of ignored inputs may reach a level commensurate with the strong, global effects of selective attention in human behavior, and that this selection in the prefrontal cortex is the basis of the attentional modulation found in more posterior sensory cortical areas, implemented through backprojections from the prefrontal cortex to the more posterior cortical areas.

4. Discussion

We have presented a detailed theoretical neurodynamical analysis of the spiking and synaptic mechanisms underlying behavior that requires complex context-dependent mapping between sensory stimuli and actions. Our neurodynamical architecture of the prefrontal cortex (PFC) unifies attentionally biased competitive mechanisms and recurrent excitatory mechanisms that support short-term memory-related

neuronal activity. Overall, the network has the architecture of a single attractor network with multiple activated populations or pools of neurons. These different pools engage in competitive interactions, are organized with some hierarchy imposed by the asymmetrically strong forward and backward connections, and receive biasing inputs to influence the relative activity of the different pools, thus implementing attention-based or rule-based mapping from sensory inputs to motor outputs. Even more, the integrate-and-fire implementation of the network enables us to make explicit predictions of the effect of neuromodulation by manipulation of the dopamine level on the conditional object-response and delayed spatial response tasks. A decrease in NMDA-related conductances produced by an increase in D2 receptor activation or a decrease in D1 receptor activation weakens and shortens the persistent neuronal activity in short term memory periods. Additionally, we predict more response errors as a consequence of the more similar level of neuronal firing in the competing neuronal pools. This fact provides insight into why schizophrenic patients have difficulties in working memory tasks, in paying and maintaining attention selectively, and thus more generally in executive functions.

REFERENCES

1. Fuster, J., Exp. Brain Res. 133 (2000) 66.
2. Funahashi, S., Bruce, C. and Goldman-Rakic, P., J. of Neurophys. 61 (1993) 331.
3. Assad, W., Rainer, G. and Miller, E., J. of Neurophys. 84 (2000) 451.
4. Brunel, N. and Wang, X., J. of Comp. Neurosc. 11 (2001) 63.
5. Chelazzi, L., Psych. Res. 62 (1999) 195.
6. Rolls, E. T. and Deco, G.. Computational Neuroscience of Vision. Oxford University Press (2002).
7. Wallis, J., Anderson, K. and Miller, E., Nature 411 (2001) 953.
8. Everling, S., Tinsley, C., Gaffan, D. and Duncan, J, Nature Neurosc. 5 (2002) 671.