ELSEVIER

Neurocomputing 38–40 (2001) 1525–1531

NEUROCOMPUTING

www.elsevier.com/locate/neucom

A model of the IT-PF network in object working memory which includes balanced persistent activity and tuned inhibition☆

Alfonso Renart[a,*], Ruben Moreno[a], Jaime de la Rocha[a], Nestor Parga[a], Edmund T. Rolls[b]

[a]*Department of Fisica Teorica, Universidad Autonoma de Madrid, 28049, Madrid, Spain*
[b]*Oxford University, Department of Experimental Psychology, Oxford OX1 3UD, UK*

Abstract

The properties of cells in the prefrontal cortex and inferotemporal cortex recorded in monkeys performing delayed matching-to-sample tasks with intervening visual stimuli and memory guided attention tasks are reproduced by means of a model in which two networks of leaky integrate-and-fire neurons representing the two cortical areas interact reciprocally. Each of the networks is organized in micro-columns (M-Cs) which leads naturally to a dynamic balance between excitation and inhibition within each M-C so that realistic cortical spiking statistics (low firing rates with higher or equal to one CVs) are obtained. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Working memory; Visual search; Balanced networks; Persistent activity; Competitive interactions

1. Introduction

The mechanisms used by primates to solve visual working memory tasks are not yet clear. In this type of experiments, two visual stimuli are shown separated by a period in which no explicit visual cue is shown (i.e. the delay period). After the delay, the

☆Partially supported by a Spanish Grant PB96-47 and a Medical Research Council Programme Grant PG8513790.

* Corresponding author. Present address. Center for Complex Systems, Brandeis University, Waltham, MA 02454, USA.

0925-2312/01/$ - see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0925-2312(01)00548-3

animal has to make a response which depends on a learned relationship between the two stimuli in order for which it has to maintain the first stimulus in memory during the delay period. This capacity of actively maintaining information in memory for subsequent use is usually termed working memory. The prefrontal cortex (PFC) has been implicated in working memory for several reasons: First, PF lesions impair performance in delay memory tasks and second, PF neurons show higher than baseline activity during the delay period, possibly reflecting the information being maintained (for a review see e.g. [7,6]). The PFC receives visual information from the ITC, an area specialized in high order visual processing but also affected by memory demands (see e.g. [11,13]), to which it also sends backward projections. In addition, cells in ITC also show delay period related activity in visual memory tasks. In a series of reports, Miller, Chelazzi, Desimone and co-workers found that the selectivity in the persistent activity is maintained in the PFC across different delay intervals within a single trial even when several intervening visual stimuli have to be processed between the sample and the matching stimuli, whereas this selectivity is disrupted by the intervening stimuli in the ITC [9,10]. Also, there are populations of cells in both the ITC and the PFC which respond more to a given visual stimuli when it is the match in a trial than when the same stimulus is a non-match. This effect, which they termed *match enhancement* and which could be used to detect repetitions of behaviorally relevant stimuli, was both more frequent and more intense in the PFC than in the ITC. Cells in the ITC have also been found to have memory-modulated responses in visual search (VS) tasks. These cells respond differently to a physically identical array of visual stimuli shown after the delay depending on which of the stimuli in the array was shown as a cue before the delay [3,2]. This differential response based on the contents of memory suggests that the underlying mechanism could also depend on the interaction between the ITC and PFC.

Recent evidence about the type of circuitry that might be involved in the generation of persistent activity in the PFC [12] and about the firing statistics of cells in persistent activity in this area [8] has also been obtained. In Ref. [12], the fact that tuned inhibition might contribute to the generation of persistent activity has been highlighted. In particular, adjacent putative excitatory and inhibitory neurons in the monkey dorsolateral PCF have been found to have similar response preferences during both the sensory and the delay periods of a visuospatial working memory task, whereas more distant pairs have inverted preferences [18]. This has been interpreted as evidence for a micro-columnar organization of the PFC. Evidence of a functional micro-columnar organization in some parts of the ITC also exists [5,17]. The spike trains of PFC cells in persistent activity seem to be highly irregular, with a coefficient of variation (CV) of the inter-spike-interval (ISI) in the range 1–1.5 [8]. A balance between the total excitation and inhibition afferent to cortical cells has been postulated as a possible mechanism for generating this temporally irregular activity [15,16]. Tightly coupled excitation and inhibition within PFC M-Cs might therefore contribute to provide a balanced recurrent feedback responsible for the observed temporal irregularity.

2. Model

Only a schematic description of the architecture of the model is provided here. A more detailed exposition will be given elsewhere. The global network consists of two modules, representing local circuits in the ITC and PFC, which are connected reciprocally. In each module there is a large number of excitatory and inhibitory neurons which are sparsely connected and organized in M-Cs. The M-Cs are defined by having a denser connectivity, so that the probability of connection between neurons in the same M-C is ~ 5 times larger than between neurons in different M-Cs. The excitatory-excitatory (e-e) connections are plastic, so that synapses to an excitatory neuron from excitatory afferents within its own M-C are potentiated, whereas the rest are depressed. In the absence of potentiation, the baseline synaptic efficacies J are such that inhibition dominates over excitation, with $J_{e\leftarrow i} \sim 4J_{e\leftarrow e}$, $J_{i\leftarrow i} \sim 4J_{i\leftarrow e}$ and $J_{i\leftarrow e} \sim J_{e\leftarrow e}$. A large background excitatory input is also present from outside each module. The IT and PF modules interchange only excitatory signals, the feed-forward projections being stronger than the feed-back ones. These long-range connections are random but plastic, so that each M-C in one module receives (sends) stronger synapses from (to) an associated pair in the other module. The amount of depression is chosen so that the net afferent synaptic efficacy with or without plasticity is the same.

We use the leaky integrate-and-fire (LIF) model with no synaptic dynamics for the single neurons. All neurons with the same statistical properties are grouped into a single sub-population, and each sub-population is characterized by the mean rate and CV of its constituent neurons. Our theory, which is an extension of [1], assumes that spike trains in the network can be described as renewal processes, with an ISI distribution close to an exponential, so that the statistics obtained be similar to Poisson, as is observed experimentally. The main steps followed to construct a self-consistent description of the mean firing rates and CV of the ISI are: (1) By means of the ISI distribution and using the renewal assumption we calculate the auto-correlogram (two-point correlation function) of the individual spike trains, which is the sum of a delta function at zero lag, plus an exponential term and which depends on the rate ν, the CV and the time-scale of the exponential correlations τ_c. (2) We then use that, since the network is sparse, the neurons will essentially be uncorrelated, so that the mean and two-point correlation of the total afferent current are calculated by adding all the individual contributions. (3) We approximate the resulting total afferent current by a Gaussian process with the same mean and two-point correlation, and calculate the first correction to the white-noise case of the mean rate and CV of a LIF neuron charged with a Gaussian current with finite, but small temporal correlations, using τ_c/τ_v as perturbative constant (where τ_v is the membrane time constant). This first correction will be a good approximation if the spike trains are close to Poisson. (4) The final result is that the output mean rate and CV are identical to those that would be obtained if all the pre-synaptic neurons where Poisson, (the total afferent current being then a white noise) but with a renormalized input variance which takes into account the CV of the pre-synaptic neurons. The mean and variance of this "effective" white noise current to neurons in, e.g.

sub-population β are:

$$\mu_\beta = \sum_\gamma C_{\beta\gamma} \langle J_{\beta\gamma} \rangle_J \nu_\gamma, \tag{1}$$

$$\sigma_\beta^2 = \sum_\gamma C_{\beta\gamma} \langle J_{\beta\gamma}^2 \rangle_J \nu_\gamma CV_\gamma^2, \tag{2}$$

where β and γ label the post- and pre-synaptic neuronal sub-populations, respectively, $C_{\beta\gamma}$ is the number of synaptic connections that a neuron in β receives from neurons γ and $\langle J_{\beta\gamma} \rangle_J$ and $\langle J_{\beta\gamma}^2 \rangle_J$ are the first two moments of the distribution of the synaptic strength of those connections. Both the mean rate and the CV are treated as dynamical variables and are, therefore, self-consistent in the stationary states.

It is well known that dopamine (DA) affects the biophysical properties of neurons in the PFC and modulates their delay period activity (see e.g. [4] and references therein). Dopaminergic neurons in the mid-brain, are also known to respond to salient, behaviorally relevant stimuli (see e.g. [14]). Although the DA modulation of PFC activity is non-selective, recent detailed computational studies have found that the net DA effect on active neurons is excitatory and reinforcing, whereas its effect on neurons in spontaneous activity is the opposite [4]. We have schematically mimicked the DA modulation of the PFC by an increase in the e-e recurrent synaptic strength of the PFC M-C coding for the behaviorally relevant sample stimulus on each trial. Since the modulation only operates on the time scale of a single trial, it is not taken into account in the balance of the total afferent potentiation and depression. The effects of this schematic manipulation go qualitatively in the same direction as those obtained with more the detailed biophysical models [4].

3. Results

There exists a range of parameters in the model in which a dynamic balance between the excitatory and inhibitory sub-populations within each M-C is achieved. This balance holds independently of the firing rate of the neurons, and is thus maintained during spontaneous, persistent and stimulus-driven activity. As a result of this balance, the mean current to the neurons becomes essentially constant and independent of the pre-synaptic rates and only the size of the afferent fluctuations change, determining the output rate. These fluctuations provoke temporally irregular activity (the irregularity actually increasing with the rate) so that large CVs are obtained ($\sim$1–1.3).

The model also reproduces almost quantitatively the firing rate modulations observed in the DMS tasks [9,10] (Fig. 1). The baseline level of potentiation (without the DA modulation) in both modules has been set below the critical value needed for bi-stability. Only the PFC M-C modulated by DA (a $\sim$30% increase in the e-e synaptic efficacy starting after the presentation of the sample) has a persistent elevated rate, which corresponds to its only stable state in this conditions. In this sense, the DA

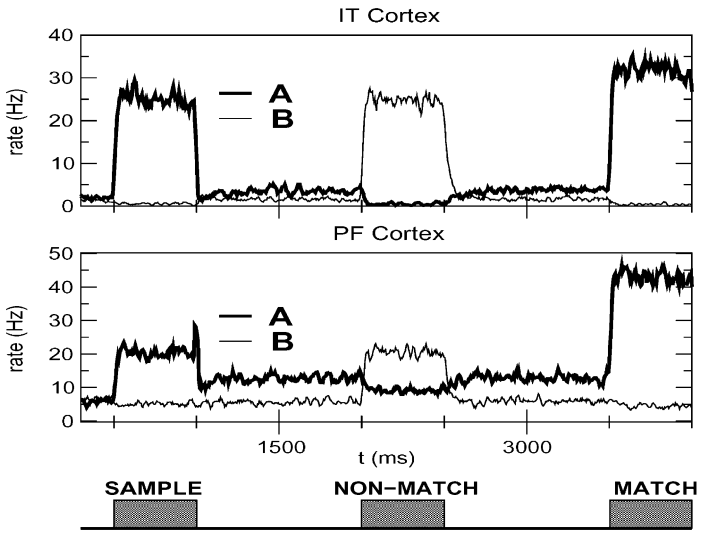

Fig. 1. DMS task with an intervening stimulus in the IT (top) and PF (bottom) networks. Thick and thin lines represent PSTHs of the excitatory sub-population of the M-Cs which are selective for the sample stimulus A and the intervening stimulus B, respectively. The PSTHs are obtained by generating Poisson trains with instantaneous firing rates as predicted by the model. This is justified since the CV in both sub-populations remains within the range 1–1.2 the whole trial (data not shown in the figure).

modulation in the model acts as a switch on the dynamic properties of the PFC network. The match enhancement effect occurs in the PF module as a result of the larger amplification by the recurrent excitation of the external signal in the DA-modulated M-C. The match enhancement in the IT module is due to the back-projected signal from the PFC, and as a result is smaller in magnitude. The very low rate delay activity in the IT module also a reflection of PF activity.

In the VS tasks, one of the two stimuli in the array is effective (good) and the other is ineffective (poor) in driving the cells response. As can be seen in Fig. 2, the stationary response to the *same* array stimulus is much larger when the good stimulus was shown as a cue than when the cue was the poor stimulus. This is a result of the fact of the feedback signal form the PF module targets cells selective to the cue stimulus in the array, biasing the competition established between the M-Cs selective to the good and the poor stimuli in the IT module. This large *non-target suppression* effect achieved in the model is directly related to the large competition between different IT M-Cs, which is an inherent property of our network of balanced M-Cs. The top-down bias is also evident in the slight selectivity of the delay period activity in the IT module, consistently with the experimental observations [2,3].

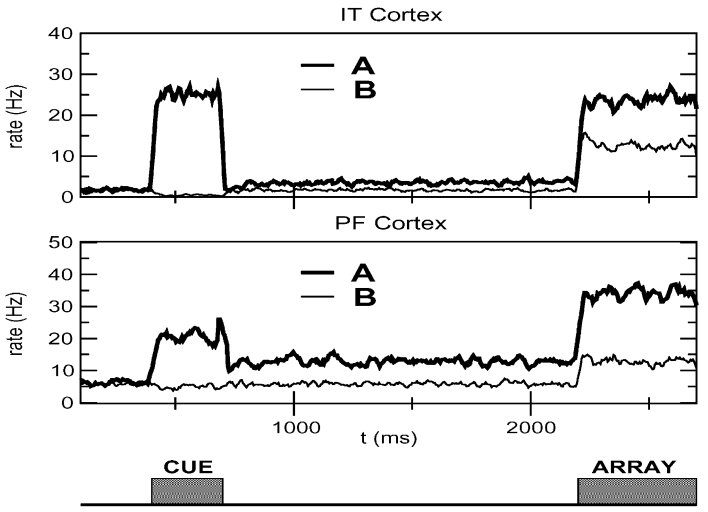

Fig. 2. Visual search experiment. The thick line represents the instantaneous mean firing rate of a subpopulation of cells when the stimulus to which the respond (good stimulus) is presented as the cue in the trial. The thin line represents the same thing when an ineffective (poor) stimulus is presented as cue. Note the large suppression in response to the array in the IT module when the cue is the poor stimulus. This is due to the large competition between representations present in the M-C network.

4. Discussion

We have proposed a model in which a dynamic excitation-inhibition balance is achieved through a M-C architecture. This leads naturally to temporally irregular activity [8] and to tuned inhibitory responses [12]. The M-Cs compete strongly for activation, which leads to the strong suppression effects observed in Refs. [2,3]. With the aid of an schematic representation of the DA-modulation in the PFC, we are also able to produce resistant selective delay period activity in the PF module and enhanced responses to the matching stimuli in both modules, with a larger enhancement in the PF module [9,10].

References

[1] D.J. Amit, N. Brunel, Cereb. Cortex. 7 (1997) 237.
[2] L. Chelazzi, J. Duncan, E.K. Miller, R. Desimone, J. Neurophysiol. 80 (1998) 2918.
[3] L. Chelazzi, E.K. Miller, J. Duncan, R. Desimone, Nature 363 (1993) 345.
[4] D. Durstewitz, J.K. Seamans, T.J. Sejnowsky, J. Neurophysiol. 83 (2000) 1733.
[5] I. Fujita, K. Tanaka, M. Ito, K. Cheng, Nature 360 (1992) 343.
[6] J.M. Fuster, Memory in the Cerebral Cortex, MIT Press, Cambridge, MA, 1995.
[7] P.S. Goldman-Rakic, Neuron 14 (1995) 477.
[8] Y.-H. Liu, Ph.D. Thesis, Brandeis University, 2000.
[9] E.K. Miller, R. Desimone, Science 263 (1994) 520.

[10] E.K. Miller, C.A. Erickson, R. Desimone, J. Neurosci. 16 (1996) 5154.
[11] Y. Miyashita, Annu. Rev. Neurosci. 16 (1993) 245.
[12] S.G. Rao, G.V. Williams, P.S. Goldman-Rakic, J. Neurophysiol. 81 (1999) 1903.
[13] E.T. Rolls, A. Treves, Neural Networks and Brain Function, Oxford University Press, Oxford, 1998.
[14] W. Schultz, J. Neurophysiol. 80 (1998) 1.
[15] M.N. Shadlen, W.T. Newsome, Curr. Opin. Neurobiol. 4 (1994) 569.
[16] M.N. Shadlen, W.T. Newsome, J. Neurosci. 15 (1998) 3870.
[17] G. Wang, K. Tanaka, M. Tanifuji, Science 272 (1996) 1665.
[18] F.A. Wilson, S.P. O'Scalaidhe, P.S. Goldman-Rakic, Proc. Natl. Acad. Sci. 91 (1994) 4009.