

# A Model for Delay Activity Without Recurrent Excitation

Marc de Kamps

Chair for Robotics and Embedded Systems, Institut für Informatik, TU München,  
Boltzmannstrasse 3, D-85748 Garching bei München, Germany

[kamps@in.tum.de](mailto:kamps@in.tum.de)

<http://www6.in.tum.de/~kamps>

**Abstract.** Delay activity (DA) is the increased firing rate of a cortical population, which persists when the stimulus that induced it is removed. It is believed to be the neural substrate for working memory, and as such highly relevant for theories of cognition. The cortex is highly recurrent, mainly excitatory, and finding stable attractors for DA at low firing rates for realistic neuronal parameters has proven to be hard. Most models for DA use recurrent excitation. Here a model with recurrent disinhibition is presented, which is manifestly stable. This model requires a cortical circuit that is slightly more complex than circuits in models using recurrent excitation, but circuits of comparable complexity have been found in cortex. Since delay attractors can not be observed directly, it is important to consider all theoretical possibilities.

## 1 Introduction

Delay activity (DA) is the increased firing rate with respect to baseline of a population of neurons, which is caused by a stimulus and which persists once the stimulus is removed. DA is believed to be the neural substrate of working memory and therefore a good model of DA is of prime importance for models of higher cognition. Experimental results indicate that DA is in the order of 10-20 Hz, whereas the normal cortical background rate is in the order of 1-10 Hz. Since the cortex is a highly recurrent network, which consists primarily of excitatory neurons, this is remarkable and it turns out to be challenging to create realistic models of cortical dynamics of DA that remain stable at rates which are far below maximum firing rates. The problem was clearly defined by Amit and Brunel [1] and their model produced stable rates for a small stimulus sensitive excitatory population, which was embedded in a larger local pool of excitatory neurons, and which were both controlled by a local inhibitory pool. Delay activity was sustained by a higher potentiation of efficacies between neurons in the stimulus sensitive population, together with the contribution from the non-stimulus specific excitatory background, i.e. by recurrent excitatory feedback. Recently Latham and Nirenberg [2] have shown that the model of Amit and Brunel (and others based on similar modeling assumptions) produces reasonable rates, for biologically plausible neuronal and network parameters, but that these rates are extremely sensitive to the choice

of these parameters. To solve this problem Latham and Nirenberg extended the original analysis from [1] considerably, and they went beyond the sparse coding limit. In the sparse coding limit, the stimulus sensitive group of neurons is assumed to make out such a small fraction of the local excitatory background, that it does not influence the background populations significantly. Latham and Nirenberg found that low rate, stable DA is possible beyond the sparse coding limit, i.e. when DA influences the local cortical background rates significantly. Network parameters were chosen such that the local background state is effectively inhibitory and an increase in DA is matched by an increase in the inhibitory background, which ensures the systems stability.

This is a strong assumption: an attractor that corresponds to a specific working memory state may involve many DA populations, and the implication of the model of Latham and Nirenberg is that the local background is increased significantly (and is effectively inhibitory) in the entire cortical area that sustains the attractor. This may actually be what happens, but as Latham and Nirenberg point out, attractors cannot be observed directly and inferences must be made about their existence by comparing experimental data with model predictions. It is therefore important to consider possible alternatives.

In this paper, I will argue that it is possible to implement delay activity be recurrent disinhibition, rather than recurrent excitation and that the only excitatory activity necessary in this model is feedforward. A cortical circuit that is implemented in this way is manifestly stable at low and plausible firing rates, but is slightly more complex than the ones considered in [1] and [2]. In the next section, I will introduce the model, and in the last section I will discuss the differences between the various models for DA and their implications for experimental data.

## 2 The Model

The main modeling assumptions are the same as in [1] and [2]: cortical neurons receive a large and unspecific background from remote cortical areas, a local excitatory and a local inhibitory population. The stationary population firing rate of population  $i$  is given by:

$$\nu_i = \phi_i(\mu_i, \sigma_i), \quad (1)$$

where:

$$\phi_i(\mu_i, \sigma_i) \equiv \left\{ \tau_{ref,i} + \sqrt{\pi} \tau_i \int_{\frac{V_{reset,i} - \mu_i}{\sigma_i}}^{\frac{\mu_i - \theta_i}{\sigma_i}} du [1 + \operatorname{erf}(u)] e^{u^2} \right\}^{-1} \quad (2)$$

$$\mu_i = \tau_i \sum_j J_{ij} N_{ij} \nu_j, \quad (3)$$

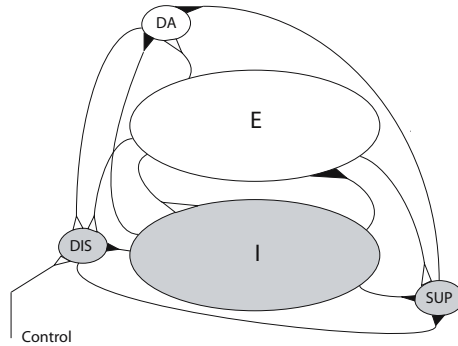
$$\sigma_i = \sqrt{\sum_j \tau_i J_{ij}^2 N_{ij}}. \quad (4)$$

$\tau_i$ ,  $\tau_{ref,i}$  are the membrane time constant and the absolute refractory period, respectively, in s,  $\theta_i$  and  $V_{reset,i}$  the threshold potential and the reset potential, respectively, in V, all for neurons in population  $i$ .  $N_{ij}$  is the effective number of neurons from population  $j$  seen by a neuron in population  $i$  and  $J_{ij}$  the effective efficacy from a spike in population  $j$  on a neuron in population  $i$  in V. These equations form a closed system which can be solved in  $\nu_i$ . In practice, one does this by introducing a pseudo-dynamics:

$$\tau_i \frac{d\nu_i}{dt} = -\nu_i + \phi(\mu_i, \sigma_i), \tag{5}$$

and selecting initial values  $\nu_i(0)$ .

The circuit has a structure as shown in Fig. 1. The neuronal, network and connectivity parameters are given in Table 1. Although the number of parameters is quite large, many are already familiar from [1]:  $g$ ,  $x$ ,  $J_{EE}$ ,  $J_{IE}$ ,  $J_{EI}$  and  $J_{II}$  are chosen such that an unspecific cortical background rate  $\nu_{ext}$ , which is input to all populations, is replicated in the local excitatory pool **E**, while the local inhibitory pool **I** fires at a slightly higher rate. About half of the input of any given neuron comes from the cortical background ( $x = 0.5$ ). **E** and **I** are stimulus insensitive, to a first approximation. A small number of neurons are distinguished by the fact that they receive slightly more potentiated input from the other neurons in **E** and they constitute a subset of **E**, denoted by **DA**. Normally, they would fire at a higher rate than neurons in **E**, but it is assumed that these neurons receive more potentiated input from a subset of **I**, denoted by **SUP** and therefore they typically fire at the same rate as neurons in **E**. Neurons in **SUP** can be inhibited rather strongly by **I** and therefore do not influence **SUP** under non-delay conditions. Hence, **SUP** will fire at the same rate as **I**, and **DA** will fire at the same rate as **E**.



**Fig. 1.** Local circuit for delay activity. Excitatory populations are white, inhibitory populations are grey. White (black) triangles indicate excitatory (inhibitory) efficacies. The relative sizes give a rough indication of the number of neurons involved.

**Table 1.** The circuit parameters.  $\tau_{exc}$  is the membrane time constant for the excitatory parameters: **E** and **DA**,  $\tau_{inh}$  for the inhibitory populations: **I**, **SUP** and **DIS**. All populations receive an extra cortical background rate  $\nu_{ext}$ , with connection parameters  $\{xC_E, J_{EE}\}$ .

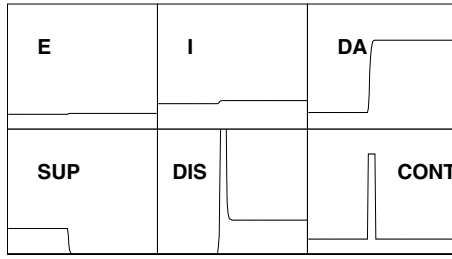
Neuronal parameters						
$\tau_{exc} = 20$ ms, $\tau_{inh} = 10$ ms, $\tau_{ref} = 2$ ms, $V_{reset} = 0$ mV, $\theta = 20$ mV						
Network parameters						
$C_E = 20000$ , $C_I = 2000$ , $x = 0.5$ , $g = 5$ , $\nu_{ext} = 3$ Hz, $\nu_E = 3$ Hz, $\nu_I = 5.1$ Hz						
$x_{DA} = 0.02$ , $\gamma_{DA} = 1.25$ , $x_{SUP} = 0.02$ , $\gamma_{SUP} = 3$ , $x_{SUP,DA} = 0.07$ , $\gamma_{SUP,DA} = 3$						
$x_{DIS} = 0.1$ , $\gamma_{DIS} = 2.5$ , $x_C = 0.03$ , $\gamma_C = 1.9$ , $x_{DA,DIS} = 0.1$ , $\gamma_{2,2}$ , $\gamma_{I,DIS} = 2.2$						
Connectivity table						
$i$	<b>E</b>	<b>I</b>	<b>DA</b>	<b>SUP</b>	<b>DIS</b>	<b>CONT</b>
$N_{Ei}$	$x(1 - x_{DA})C_E$	$(1 - x_{SUP})C_I$	$xx_{DA}C_E$	$x_{SUP}C_I$	0	0
$J_{Ei}$	$\theta/193.4$	$gJ_{EE}$	$J_{EE}$	$J_{EI}$	0	0
$N_{Ii}$	$x(1 - x_{DA})C_E$	$(1 - x_{SUP})C_I$	$xx_{DA}C_E$	$x_{SUP}C_I$	0	0
$J_{Ii}$	$\theta/120$	$gJ_{IE}$	$J_{IE}$	$J_{II}$	0	0
$N_{DAi}$	$x C_E$	$(1 - x_{SUP})C_I$	0	$x_{SUP,DA}C_I$	0	0
$J_{DAi}$	$\gamma_{DA}J_{EE}$	$J_{EI}$	0	$\gamma_{SUP,DA}J_{EI}$	0	0
$N_{SUPi}$	$x(1 - x_{DA})C_E$	$C_I$	$xx_{DA}C_E$	0	$x_{DIS}C_I$	0
$J_{SUPi}$	$J_{IE}$	$J_{II}$	$J_{IE}$	0	$\gamma_{DIS}J_{II}$	0
$N_{DISi}$	$C_E x(1 - x_C - x_{DA,DIS})$	$C_I$	$xx_{DA,DIS}C_E$	0	0	$x_C C_E$
$J_{DISi}$	$J_{IE}$	$\gamma_{I,DIS}J_{II}$	$\gamma_{DA,DIS}J_{IE}$	0	0	$\gamma_C J_{IE}$

Now consider the situation where **DIS** is stimulated rather strongly by an external control signal and therefore will inhibit **SUP**, which results in turn in the disinhibition of **DA**, which responds by a higher firing rate. Moreover, **DA** excites **DIS**. Under no-delay conditions **DA**'s firing rate, which is equal to that of **E**, is not able to overcome the hard inhibition of **I** on **DIS**. But, if it is higher, disinhibited, firing rate is able to keep **DIS** active, even when the control stimulus is removed, then the elevated firing rate of **DA** will persist and is delay activity. This state will only return to the original one if **DIS** will be inhibited again, for instance due to another control signal, or if one of the populations that fire at an elevated rate (**DA**, **DIS**) is affected by adaptation.

### 3 Results

In Fig. 2 we see this happen: at  $t = 0.40$  s, the rate of **CONT** is raised by an external stimulus, which lasts for 0.05 s. The result is that for a brief while **DIS** is firing at a high rate (approximately 70 Hz), which inhibits **SUP**. The inhibition of **SUP**, allows the higher potentiated **DA** population to fire at a higher rate of approximately 18 Hz. Its excitation of **DIS** is strong enough to keep it firing at approximately 7 Hz, which is strong enough to keep **SUP** inhibited.

For this mechanism to function, the **SUP** neurons must inhibit the **DA** neurons rather specifically. This is reflected in the relative high values of  $x_{SUP,DA}$



**Fig. 2.** For each histogram, the horizontal axis represents time  $t$ :  $0 < t < 1$  s. The vertical axis represents firing rate  $f$ ,  $0 < f < 25$  Hz. The peak in **DIS** briefly extends to 70 Hz.

and  $\gamma_{DA}$ . The values for  $x_{SUP,DA}$  and  $\gamma_{SUP,DA}$  are chosen such that **DA** fires at rate  $\nu_{ext}$  in the non-delay condition, i.e. at the same rate as the rest of module **E**. It is assumed that there is not much interaction between the **SUP** and the **I** module. Other than the specific connections from **SUP** to **E** and from **DIS** to **SUP**, there is no distinction between these neurons and neurons of **I**, and **SUP** is assumed to fire at the same background rate as **I**,  $\nu_I$  in the non-delay condition. **DIS** neurons must inhibit the **SUP** neurons rather specifically, shown by the relatively high values of  $x_{DIS}$  and  $\gamma_{DIS}$ . Importantly, the **DA** neurons must not be able to overcome the inhibition on **DIS** at the normal firing rates, but must inhibit **DIS** when firing at delay rates. The values of the other  $x$  and  $\gamma$  parameters is uncritical, and might have been taken zero instead. They have been chosen to demonstrate that small interactions between other populations than the ones described above, which are to be expected, do not disrupt the mechanism.

It turns out that parameter space is large and other reasonable values for firing rates can easily be found: for example if one chooses  $\gamma_{DA} = 1.13$ ,  $\gamma_{SUP,DA} = 1.6$  and  $\gamma_{DA,DIS} = 3.5$  one finds delay rate at 10 Hz, rather than at 18 Hz, with the other rates close to the ones shown in Fig. 2. The explanation is simple: the lower value for  $\gamma_{DA}$  leads to a lower firing rate in case **SUP** is inhibited.  $\gamma_{SUP}$  must be decreased, so as to keep **DA** firing at the background rate  $\nu_E$  in the non-delay condition, and the lower delay firing rate of **DA** must be compensated by an increased potentiation  $\gamma_{DA,DIS}$ .

The parameter space is substantially enlarged by inhibiting **DIS**. This basically decouples **DA** from **DIS** and **SUP** in the non-delay condition. If **DIS** were firing at background rate  $\nu_I$ , changes in the rate of **DA** will be fed back to **DA** via **DIS** and **SUP** and it becomes harder to find parameters that give a desired delay rate.

## 4 Discussion

This model uses almost the same modeling assumptions as [1]. Particularly important is the idea that every neuron receives a large number of input spikes, even if it is not directly stimulated and only receives baseline rates from other

neurons. I use this idea as well, but in a different way: the values in Table 1 show that a moderately higher potentiation of a relatively small fraction of its input synapses can lead to a firing rate which is significantly higher than baseline activity. Hence, it is possible that spontaneous firing rates significantly above baseline could emerge in such a population, if there is no compensating extra potentiation of its inhibitory inputs. This possibility is crucial for the model described here. A large number of parameters is necessary to describe local pools **E**, **I**, which are firing at stable and low firing rates, but this part of the model is the same as in [1] and [2].

The crucial departure from these models consists in the assumption that there is structure in the **I** population, which functions as a disinhibition circuit. Such disinhibition circuits have been shown to exist in cortex [3]. Although at first sight this model involves a more complicated structure than [1] and [2], it is simpler in dynamical terms. *This is because there is only feedforward excitation in delay conditions*, and in non-delay conditions the **DA** and **SUP** are integral parts of the **E** and **I** populations, respectively. Moreover, it is not necessary for **DA** neurons to couple to themselves. Finally, the inhibition of **DIS** under non-delay conditions prevents even indirect feedback of **DA** onto itself.

To play a part in cognitive processes, working memory must be controlled: it must be selected, information must be stored into it and retrieved from it, and it is likely that the control of such operations is performed by gating circuits which are very much like the ones described in this model. It is also well known that inhibition plays a role in working memory and that disruptions of its function can result in substantial cognitive impairment. This suggests that inhibition plays a more important role than just rate control.

Earlier experience with large scale cortical modeling [4] has shown that stability of dynamics in local circuits is essential to ensure stability in a large network. This has been the prime motivation for this model, but in the end experiment decides. The predictions of this model for experiments that involve **DA** are clearly different from [2]: in our model only specific subsets of neurons fire at elevated rates in the delay condition, whereas in [2] the entire area that sustains the attractor is involved. In the former case it should be easy to find neurons that fire at baseline activity, whereas this should be more difficult in the latter.

## References

1. Amit, D., Brunel, N. Model of Global Spontaneous Activity and Local Structred Activity During Delay Periods in the Cerebral Cortex. *Cerebral Cortex* **7** (1997) 237-252
2. Latham, P. E., Nirenberg, S. Computing and Stability in Cortical Networks *Neural Computation* **16** (2004) 1385-1412
3. Gonchar, Y. and Burkhalter, A. Connectivity of GABAergic Calretinin-immunoreactive Neurons in Rat Primary Visual Cortex. *Cerebral Cortex* **9** (1999) 683-696
4. van der Velde, F., de Kamps, M. Neural Blackboard Architectures of Combinatorial Structures in Cognition. To appear in: *Behavioral and Brain Sciences*.