Activity propagation in a network of coincidence-detecting neurons.

Guido Bugmann and John G. Taylor*

School of Computing and Mathematics, Plymouth University, Plymouth PL4 8AA.
Fax (+44) 1752 586300, Phone: (+44) 1752 586215, email: gbugmann@plymouth.ac.uk
*Centre for Neural Networks, King's College London, London WC2R 2LS United Kingdom.

24 February 2013

Keywords: Visual system, Latency, Masking, Synchronization, Jitter, Short-term memory, Sustained firing, neuron model.

Contents

Abstract

This paper presents a formal analytical description of activity propagation in a simple multilayer network of coincidence-detecting neuron models receiving and generating Poisson spike trains. Simulations are also presented.

In feedforward networks of coincidence-detecting neurons, the average firing rate decreases layer by layer, until information disappears. To prevent this, the model assumes that all neurons exhibit self-sustained firing, at a preset rate, initiated by the recognition of local features of the stimulus. Such firing can be interpreted as a form of local short-term memory. Inhibitory feedback signals from higher layers are then included in the model to minimize the duration of sustained firing, while ensuring information propagation.

The theory predict the time-dependent firing probability in successive layers and can be used to fit experimental data.

The analyzed multilayer neural network exhibits stochastic propagation of neural activity. Such propagation has interesting features, such as information delocalization, that could explain backward masking. Stochastic propagation is normally observed in simulations of networks of spiking neurons. One of the contributions of this paper is to offer a method for formalizing and quantifying such effects, albeit in a simplified system.

The mathematical analysis produces expressions for latencies in successive layers in dependence of the number of inputs of a neuron, the level of sustained firing and the onset time jitter in the first layer of the network. In this model, latencies are not caused by the neuronal integration time, but by the waiting time before a coincidence of input spikes occurs. Numerical evaluation indicates that the retinal jitter may make a major contribution to inter-layer visual latencies. This could be confirmed experimentally.

An interesting feature of the model is its potential to describe within a single framework a number of apparently unrelated characteristics of visual information processing, such as latencies, backward masking, synchronization and temporal pattern of post-stimulus histograms. Due to its simplicity, the model can easily be understood, refined and extended.

This work has its origins in the nineties, but modeling latencies and firing probabilities in more realistic biological system is still an unsolved problem.

1 Introduction

Modeling the temporal characteristics of the neuronal responses is a crucial test of our understanding of biological neural computation. In this paper we analyze and simulate a simple dynamical model of the visual system developed to identify the factors determining visual latencies. The model explains conflicting data on visual response latencies and provides testable predictions. It also offers an explanation of backward masking phenomena.

Latencies in the visual system depend on how they are measured. Using direct electrical stimulations consistently indicates latencies of less than 2 ms per neuronal relay [Ferster and Lindstrom, 1983]. However, using visual stimulation, average latencies of 10 ms per neuronal relay are commonly found values [Thorpe and Imbert, 1989]. Latencies of this length are observed, for instance, between layer 4 and layers 2-3 in the striate cortex [Best et al., 1986; Maunsell and Gibson, 1992].

These differences arise from differences between experimental conditions. Electrical stimulations produce highly synchronized input spikes [Best et al., 1986] which may raise the potential of the membrane of the neuron very rapidly to the firing threshold and cause unnaturally short latencies. It has been proposed that, with visual stimulation, input spikes arrive asynchronously and must be integrated for appropriate intervals before the neuron reaches its firing threshold [Maunsell and Gibson, 1992].

While temporal integration has been a successful model to reproduce the dependence on light intensity of the latency in retinal cells [Burgi and Pun, 1991], it would cause cortical cells to fire regularly, which is contrary to observations that spike trains are nearly random [Softky and Koch, 1993]. Softy and Koch therefore suggested that neurons somehow operated as coincidence detectors, a function that produces Poisson spike trains. Coincidence detection has some attractive features. It corresponds to a multiplicative, AND-type, function of the neuron which has been postulated for neurons in the visual area MST [Verri et al., 1992; Saito, 1993]. It is a function well suited to the binary probabilistic Random Access Memory (pRAM) neuron model developed by Gorse and Taylor [1990a]. This model operates in discrete time steps. In each step, the input spike trains form a pattern of ones and zeros. This pattern is used as the address of a memory location in a RAM. In that location is stored a *probability* of producing an output spike (a one). For instance a pRAM with two inputs can have four possible input patterns: (1,1), (1,0), (0,1) and (0,0) and four corresponding memory locations. A perfect coincidence detector has stored probabilities p(1,1)=1, p(1,0)=p(0,1)=p(0,0)=0. It only fires when each input provides a spike in the same time step. This model is theoretically attractive and powerful, but requires 2^m memory locations, a number that becomes prohibitively large for a large number m of inputs. The required memory space can be reduced by replacing one large pRAM with a pyramid of pRAMS with a smaller number of inputs. This architecture can be trained by reinforcement learning [Gorse and Taylor, 1990b]. A one-shot learning algorithm to train such structures has recently been developped [Bugmann, 2009]. The work presented here can be seen as a study into the dynamics of the operation of pyramidal networks of pRAM neurons. This is however restricted to pRAMS operating as pure coincidence detectors, a mode relevant to biological modeling.

Four year after Softky and Koch's paper, it was shown that partial reset of the membrane potential after the production of a spike was the likely cause for the irregular firing [Troyer and Miller, 1997; Bugmann et al., 1997]. In that model, the first spike of a response relies on inputs raising the potential by the full value of the firing threshold. However, subsequent input spikes face a much smaller threshold and are more likely to produce output spikes. Spike trains produced in this way have a Poisson distribution of inter-spike intervals if the correct level of partial reset is used. Functionally, the production of the first spike is the critical process determining the selectivity of the neuron. Celebrini et al. [1990] observed that the first spike of a response was already well tuned.

Such tuning reflects the co-operative contribution of a large fraction of the input neurons in its receptive field. It excludes the possibility for a neuron to be slowly brought to threshold by a small number of very active inputs. The fastest way to detect that a given number m of inputs are active is through a mechanism detecting the coincidence of m single spike produced by each one of the inputs. Given that

neurons fire at frequencies up to 200Hz in LGN [Lessica and Stanley, 2004], up to 200Hz in V1 [Reich et al, 2001] and up to 100Hz in IT [Logothetis et al., 1995; Oram, 2010], the integration of single input spikes from different source neurons must be completed within less than 5ms in V1 and its targets, and less than 10ms in IT. In V1, average decay time constants of 18ms have been reported [Boudreau and Ferster, 2005] corresponding to a 25% loss of potential within 5ms. Thus, with appropriately set synaptic weights, the first output spike can only be produced by spikes from many inputs arriving within a time window of that order. Results by Roy and Alloway [2001] also point to a coincidence time window of 6-8ms in the somato-sensory cortex.This is not a strict within-one-time-step coincidence as analyzed in the theory in this paper, but leads validity, at least qualitative, to the results produced here. It also supports the approximation in this paper that only the first spike is the result of a coincidence, while subsequent ones are generated as a Poisson process.

In this paper we consider a multilayer network of very simple neuron models that only fires in response to the coincidence, within one time step (typically $1ms$), of one spike from each of their inputs.

A problem with the generation of output spikes only upon coincidences of input spikes is that the output firing frequency is always lower, and even much lower if the number of inputs is large, than the level of the input frequency [Bugmann, 1991, 1992a]. Therefore, neural information may rapidly vanish after propagating through a few cortical layers [Bugmann, 1992b].

When the model analyzed in this paper was originally developed, in the early 90s, partial reset was unknown. Therefore, to explain the rather constant level of firing in various cortical areas, some kind of firing-rate-boosting mechanism had to be imagined. In analog electronics, self-feedback connections are used to control the gain of an operational amplifier. The same method was used to control the frequency of random spike trains [Bugmann and Taylor, 1993a]. Neurophysiologically, this would correspond to the use of autapses (synapses made by the axon of a neuron in its own dendritic tree) [Van der Loos and Glaser, 1972], or a local excitatory feedback circuit described by [Douglas and Martin, 1991] and analyzed earlier by [Grossberg, 1982], possibly using local-circuit synapses producing giant $10mV$ EPSPs described by [Thomson and Deuchars, 1994]. When local inhibitory inter-neurons are disabled, such a circuit can cause a prolonged sustained firing [Douglas and Martin, 1991], which a self-feedback loop would cause in the high-gain limit. In our model, the effect of the local excitatory circuit is extremely simplified: neurons enter in a state of sustained firing as soon as a coincidence is detected. Sustained response have since been observed in many cortical areas, and are nowadays an accepted concept. See e.g. [McCormick et al., 2003].

Indefinite sustained firing however causes a new problem. It locks the visual system in a state where new information cannot be processed. Although probabilistic synaptic transmission theoretically prevents indefinite sustained firing from occurring, it can last long enough (e.g. several tens of seconds [Fuster et al., 1981]) to make a network unusable. Therefore, a resetting mechanism is incorporated in the model: when a neuron detects a coincidence, it sends a resetting signal to its input neurons in the previous layer. This feature reduces to a strict minimum the duration of the sustained firing while ensuring information propagation. Similar feedback inhibition schemes are used in models of speech production [Houghton, 1990; Burgess and Hitch, 1992],

olfactory recognition [Granger et al., 1990] and visual search [Humphrey and Muller, 1993]. There is no physiological evidence that interlayer feedback projections play the role of controlling the duration of the sustained activity. However, experiments on short-term memory show that the duration of the sustained firing is somehow controlled by external cues [Funahashi et al., 1989].

The question of whether self-sustained firing is needed in the presence of partial reset is answered positively in the discussion (section 5.3).

The input layer of the model represents an aggregate of the Lateral Geniculate Nucleus (LGN) and the Retina. The LGN is generally assumed to be a gating relay controlling the transmission of retinal spike trains. The LGN receives a large number of feedback projections from the primary visual area (V1) and other visual areas [Sherman and Koch, 1986]. The role of these feedbacks is still unclear. Disabling the feedbacks from V1 to the LGN has sometimes led to an increase and sometimes to depression of the responses in LGN [Garey et al., 1991 and reference therein]. As feedbacks form synapses mainly on inhibitory inter-neurons [Garey et al., 1991 and reference therein] we will simply assume that a subset of the feedback connections has an inhibitory effect and stops the transmission when active.

Retinal ganglion cells respond to brief flashes of light with a burst of spikes lasting several tens of ms. The burst starts after a certain delay (retinal latency) and with some jitter in onset time [Levick, 1973]. The standard deviation of the onset time is typically 8 ms in cat. In the model, the retina and LGN are represented by a single node exhibiting jitter in onset-time and indefinite sustained firing which can be silenced by feedback signals. The retinal latency, which depends on the light intensity, is not modeled.

In section 2 we present the mathematical description of the model and calculate the time-dependent firing probabilities in successive layers. Feedback is not considered in the mathematical description because it does not affect the velocity of the feed-forward activation propagation. Temporal effects of feedback will however be examined trough numerical simulation of the network. These include temporal features of postsynaptic histograms and synchronization (see results and discussion in section 5.4.)

In section 3, expressions for the latencies are extracted from the firing probabilities and compared to latencies measured on simulations of the model. Latencies depend on the probability of a coincidence, that depends on the probability of inputs neurons to be in a state of sustained firing. Latencies also depend the onset jitter in the first layer of the model. The discussion (section 5.1) will present arguments based on physiological data suggesting that the jitter in onset time of retinal ganglion cells is a major contributor to visual latencies.

In section 4, theoretical latency expressions are used to fit two sets of experimental data. First, a set of latencies data by [Kawano etal, 1994] in area MST were used. Good fits were obtained suggesting that some early physiological responses may reflect pure propagation delays.

Secondly, visual backward masking data by [Bergen and Julesz, 1983] were used. One of the property of our model is the delocalization or spread in time of the visual activity. Although the model operates in a purely feed-forward way, and the stimulus is presented at the same time to all input neurons, stochastic propagation leads to neurons in different layers to be responding at the same time to the stimulus (as illustrated in

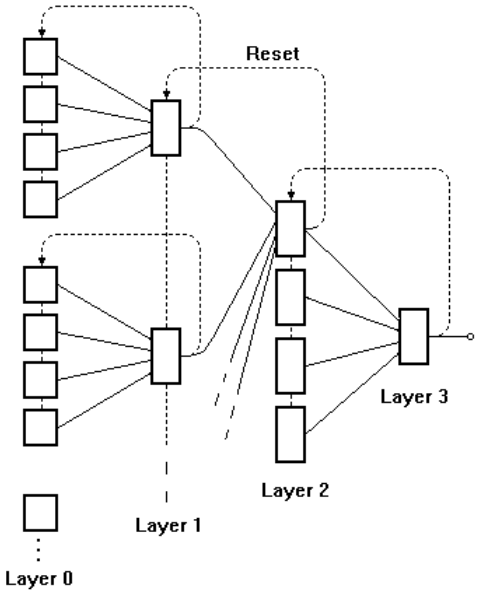


Figure 1: Architecture of the neural network used to model a subset of active neurons in the visual system. The pyramidal network has 64 neurons in its input layer (Layer 0), divided in 16 groups of 4 neurons connected to 16 neurons in layer 1. These neurons are divided into 4 groups of 4 neurons connected to 4 neurons in layer 2. These 4 neurons are connected to a single neuron in layer 3.

section 3.3). Such a property led us to investigate the application of our model to visual backward masking phenomena. Good fits are produced, suggesting that stochastic activity propagation may constitute a basis for explaining both the backward making phenomena and the probabilistic response behavior of subjects.

Preliminary and partial results were presented at a workshop [Bugmann and Taylor, 1994a] and as a conference paper [Bugmann and Taylor, 1993b]. The theory is described here for the first time.

2 Theory of activity propagation

2.1 Summary of the model

Our purpose is not to model the whole visual system but only the subnetwork which is involved in processing a given stimulus [1]. Therefore, in our model, all neurons must be activated at some time to ensure information propagation. Each neuron receives converging activation from a number of neurons in the previous layer. Several layers may constitute a cortical area. Due to a roughly equal number of neurons in each layer, neurons are likely to share some of the input neurons in the previous layer. We will not consider such a sharing in our model and use a pure pyramidal architecture shown in figure 1.

The model can be summarized as follows:

[1] Such sub-network can be built using learning rules described in [Bugmann, 2012]

1) Neurons in the visual cortex are part of a pyramidal multilayer network where each neuron receives spikes from m distinct neurons in the preceding layer.

2) Neurons operate in discrete time steps and act as coincidence detectors, firing only if all m inputs provide a spike within a given time window (one time step).

3) Neurons receive excitatory feedback from local-circuit neurons and stay in a state of persistent activity ("ON" state) after producing their first spike. During the ON-state, the neuron produces a spike with a probability p_1 at each time step.

4) Neurons in the self-sustained ON-state are silenced by inhibitory feedback from their target neurons in the next layer. After a time of inhibition, neurons are allowed to respond to new coincidences. The neurons in the input layer continue to fire with probability p_1 at the end of the inhibitory period. This represents a reopening of the transmission of retinal spikes by LGN cells.

5) The jitter in the input layer is modeled as follows: The input neurons are initially set to fire randomly with a probability p_0. However, as soon as their first spike is produced, their firing probability is set to p_1. So, the only role of p_0 is to produce a jitter in the starting time of the neurons in the first layer.

6) Neurons have all the same frequency of sustained firing, determined by p_1.

2.2 Firing onset in the input layer

In order to develop the mathematical expression of the time dependent onset of STM neurons in a layer n we start, in this section, by defining the expressions for the firing rate in the input layer 0. In the next section we then consider the general case of the firing rate in layer $n+1$ as a function of the rates in layer n.

Input neurons in layer 0 have from time 1 on a firing probability p_0. As soon as the first spike is produced, the neuron enters in a state of sustained firing such that, from the next time step on, spikes are produced at random with a constant probability p_1. The probability $P_0(0,k)$ that a neuron in layer 0 has produced no spikes yet up to and including the time step k is:

$$P_0(0,k) = (1-p_0)^k = \overline{p_0}^k. \tag{1}$$

The probability $P_0(f,k)$ that an input neuron has started firing sometimes between $t=0$ and $t=k-1$ included, and therefore has a firing probability p_1 at time k, is given by:

$$P_0(f,k) = 1 - P_0(0,k-1) = 1 - \overline{p_0}^{k-1}. \tag{2}$$

The probability $P_0(1,k)$ that the first spike be produced at time k is given by:

$$P_0(1,k) = p_0 P_0(0,k-1) = p_0 \overline{p_0}^{k-1}. \tag{3}$$

The probability $P_0(k)$ that the neuron produces a spike at time k is:

$$P_0(k) = p_1 P_0(f,k) + P_0(1,k) = p_1 + (p_0 - p_1)\overline{p_0}^{k-1}. \tag{4}$$

The standard deviation σ_k of the starting time in layer 0 is calculated as follows. The average starting time $< k >$ and average square of starting times $< k^2 >$ are given by

$$< k >= \sum_{k=1}^{\infty} k p_0 \overline{p_0}^{k-1} = \frac{1}{p_0} \tag{5}$$
$$< k^2 >= \sum_{k=1}^{\infty} k^2 p_0 \overline{p_0}^{k-1} = \frac{2 - p_0}{p_0^2}.$$

Using the definition of the variance σ_k^2

$$\sigma_k^2 =< k^2 > - < k >^2 = \frac{1 - p_0}{p_0^2}, \tag{6}$$

leads to

$$\sigma_k = \frac{\sqrt{1 - p_0}}{p_0}. \tag{7}$$

2.3 Propagation from layer n to layer $n + 1$

Let us now turn to the activity propagation from layer n to layer n+1. We wish to construct a recursive expression for the firing probability in layer $n + 1$ as a function of the firing probability in layer n. The expression of the firing probability $P_{n+1}(k)$ in layer $n + 1$ is similar to (4): A neuron in layer $n + 1$ produces a spike at time k either because it produces its first spike at that time or because it is in a state of sustained firing at that time. In the latter case, spikes are produced with a probability p_1. Thus

$$P_{n+1}(k) = p_1 P_{n+1}(f, k) + P_{n+1}(1, k), \tag{8}$$

where $P_{n+1}(f, k)$ is the probability for a neuron in layer $n + 1$ to be firing with a sustained rate p_1 at time k, and $P_{n+1}(1, k)$ is the probability for a neuron in layer $n+1$ to produce its first spike at time k.

It is possible to express $P_{n+1}(f, k)$ as a function of all past values of $P_{n+1}(1, i), (i < k)$ using following definition of $P_{n+1}(f, k)$:

$$P_{n+1}(f, k + 1) = P_{n+1}(1, k) + P_{n+1}(f, k). \tag{9}$$

This equation describes the fact that a neuron is in a state of sustained firing at time $k + 1$ either because it was already in that state at time k or because it produced its first spike at time k. Using (9) as a recursive definition of $P_{n+1}(f, k)$ one finds

$$P_{n+1}(f, k) = \sum_{i=1}^{k-1} P_{n+1}(1, i). \tag{10}$$

Therefore $P_{n+1}(k)$ (equation (8)) is only dependent of $P_{n+1}(1, i), (i \leq k)$:

$$P_{n+1}(k) = P_{n+1}(1, k) + p_1 \sum_{i=1}^{k-1} P_{n+1}(1, i). \tag{11}$$

In order to express $P_{n+1}(1, k)$ as a function of the firing probabilities in the previous layer n, we introduce two definitions of $P_{n+1}(c, k)$, the probability that a coincidence

of m input spikes is observed at time k. First, we may note that a coincidence of m input spikes can occur at time k in one of two following exclusive cases: i) The target neuron produces its first spike at time k. ii) The neuron is in a state of sustained firing at time k. This implies that all m inputs are also in a state of sustained firing. In that case, the probability of a coincidence is p_1^m. Thus

$$P_{n+1}(c,k) = p_1^m P_{n+1}(f,k) + P_{n+1}(1,k). \tag{12}$$

The second definition uses only the input firing probabilities $P_n(k)$:

$$P_{n+1}(c,k) = P_n(k)^m. \tag{13}$$

Given a propagation delay τ between two layers, from equations (13), (10) and (12) one finds

$$P_{n+1}(1,k) = P_n(k-\tau)^m - p_1^m \sum_{i=1}^{k-1} P_{n+1}(1,i). \tag{14}$$

Using equation (14) as a recursive definition of $P_{n+1}(1,k)$ one finds

$$P_{n+1}(1,k) = P_n(k-\tau)^m - p_1^m \sum_{i=0}^{k-2} (1-p_1^m)^i P_n(k-\tau-i-1)^m. \tag{15}$$

Incidentally, comparing (14) and (15) one sees that equation (10) can be rewritten as

$$P_{n+1}(f,k) = \sum_{i=0}^{k-2} (1-p_1^m)^i P_n(k-\tau-i-1)^m. \tag{16}$$

Finally, from equations (15) and (16) the expression (8) for the firing probability in layer $n+1$ in dependence of the firing probability in layer n becomes

$$P_{n+1}(k) = P_n(k-\tau)^m + (p_1 - p_1^m) \sum_{i=0}^{k-2} (1-p_1^m)^i P_n(k-\tau-i-1)^m. \tag{17}$$

This is a recursive definition of $P_n(k)$ which could be exploited to define a final expression depending only on n, k, p_1, p_0 and m. We have not attempted to do so as equation (17) can be used to extract approximate dependences on p_0 and p_1, as shown in following sections. Further, equation (17) has a form suitable for a numerical implementation. The firing probabilities predicted by this equation can be compared, in figure 2B, with those produced by simulation of the model. The results of fitting experimental data with expression (17) are shown in section 4.1.

3 Latencies

3.1 Minimum latency (Case $p_1 = 1$)

Simulations show that latencies decrease as p_1 increases and reach a non-zero minimum value as p_1 approaches 1. We will first concentrate on the parameters determining this

minimum latency by evaluating the firing rate (17) in the case $p_1 = 1$. Then we will consider in section 3.2 the more general case of $p_1 < 1$ and find how the latencies approach their minimum as a function of p_1.

In the case $p_1 = 1$, equation (17) reduces to

$$P_{n+1}(k) = P_n(k - \tau)^m. \tag{18}$$

Applying (18) recursively to lower and lower layers one finds

$$P_{n+1}(k) = P_0(k - (n+1)\tau)^{m^{n+1}}, \tag{19}$$

and using (4)

$$P_{n+1}(k) = (1 - \overline{p_0}^{\,k-(n+1)\tau}\,)^{m^{n+1}}. \tag{20}$$

For $p_0 > 0$, the term $\overline{p_0}^{k-(n+1)\tau}$ becomes rapiely very small as k increases. One can therefore approximate

$$P_{n+1}(k) \approx 1 - m^{(n+1)}(1 - p_0)^{k-(n+1)\tau}. \tag{21}$$

Experimentally, latencies are usually defined by the time at which the firing rate rises above the background level. This involves a linear extrapolation of the early rising phase in firing probability [Maunsel and Gibson, 1992]. Theoretically, it is more convenient to define the latency of layer n by the time $k_c(n)$ at which the firing probability reaches 1/2 of the maximum rate p_1

$$P_n(k_c(n)) = \frac{1}{2}p_1, \tag{22}$$

and using (21) we find

$$k_c(n) = \frac{\ln(1/2) - n\ln(m) + n\tau\ln(1 - p_0)}{\ln(1 - p_0)}. \tag{23}$$

From (23) the latency difference between layer n and $n - 1$ is

$$\Delta L_n = k_c(n) - k_c(n-1) = \frac{\tau\ln(1 - p_0) - \ln(m)}{\ln(1 - p_0)}. \tag{24}$$

Using from (23)

$$k_c(0) = \frac{\ln(1/2)}{(1 - p_0)}, \tag{25}$$

one can write (24) as

$$\Delta L_n = k_c(0)\frac{\tau\ln(1 - p_0) + \ln(1/m)}{\ln(1/2)}. \tag{26}$$

Simplified approximate expressions are further developed in sections 3.3 and 5.1.

3.2 Latency in dependence on p_1

Finally we turn to the case $p_1 < 1$, and in particular discuss the three parameter regimes $x > 1$, $x = 1$ and $x < 1$ where $x = (1 - p_1^m)/\overline{p_0}$. Then for large k and $p_1 > p_0$ we show in the appendix A that

i) for $x > 1$

$$\Delta L_n \approx \tau + 1 + \frac{\ln(m)}{p_1^m} \approx A + B/p_1^m. \tag{27}$$

ii) for $x = 1$ and $x < 1$

$$\Delta L_n \approx \tau - \frac{\ln(m) + (m-1)\ln(p_1) + \ln(\frac{p_1 - p_0}{p_1^m - p_0})}{\ln(\overline{p_0})}. \tag{28}$$

Expression (27) shows the $1/p_1^4$ dependence of the latency shift on p_1. However, for p_1 approaching 1, the expression (28) is better suited. In the limit $p_1 = 1$, it reduces to (26). The quality of the approximation (28) is shown in figure 3.

3.3 Comparison with simulations

While the theory focuses on onset-latencies, simulations allow to also explore and visualize more complex dynamic effects, such as those caused by feedback signals.

The simulations are performed with a pyramid of pRAM neurons [Gorse and Taylor, 1990a]. These neurons can act as coincidence detectors, like leaky integrate-and-fire neurons [Bugmann, 1991; 1992a], and are easier to analyze mathematically. The pRAM operates in discrete time-steps Δt with spike trains defined as sequences of ones (a spike) and zeros. We use pRAMs with $m = 4$ inputs set for the coincidence detection mode, i.e. firing only when all 4 input spikes are present during a time-step Δt. As soon as its first spike is produced, the pRAM is set to fire at each subsequent time step with a probability p_1. This produces a random spike train of frequency $f_1 = p_1/\Delta t$ characteristic of the ON-state and simulates the effect of local excitatory feedback. Such a random firing is consistent with the near Poisson distributions of inter-spike intervals observed in biological neurons [Softky and Koch, 1993]. The interlayer propagation time τ is 1 time step.

The pyramidal network is illustrated in figure 1. It has 64 neurons in its input layer (Layer 0), divided in 16 groups of 4 neurons connected to 16 neurons in layer 1. These neurons are divided into 4 groups of 4 neurons connected to 4 neurons in layer 2. These 4 neurons are connected to a single neuron in layer 3. The number of 4 inputs per neuron is small compared to the number of feed-forward inputs in visual neurons (between 40 and 240, see discussion in section 5.1). This has no effect on the principle we attempt to demonstrate and allows short computation times on the PC-486 used for the simulations.

The input neurons are initially set to fire randomly with a firing probability p_0 at each time step. However, as soon as their first spike is produced, their firing probability is set to p_1. So, the only role of p_0 is to produce a jitter in the starting time of the neurons in the first layer. Such a jitter can be seen in the figures 2C and 8B.

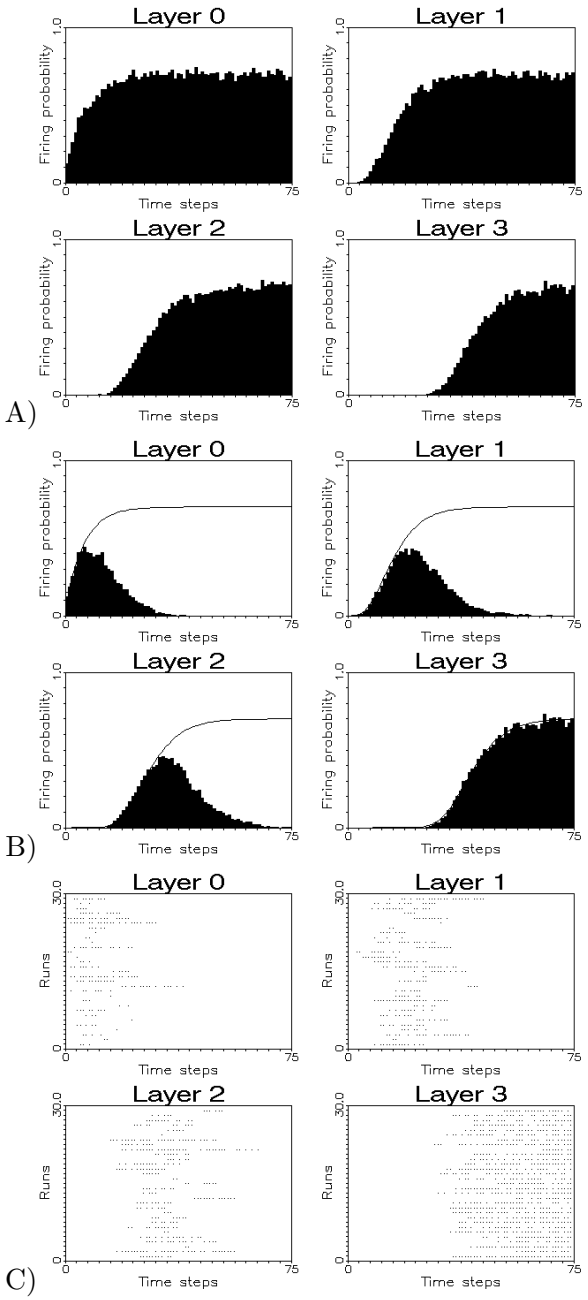


Figure 2: A) Temporal evolution of the firing probability in the case without reset by the inhibitory feedback. B) Same as A) with reset. C) Spike raster of 30 runs in the case with reset. Each dot indicates the time of occurrence of a spike. Simulation parameters: $p_0 = 0.15$, $p_1 = 0.7$. Histograms are based on 500 runs. The continuous line in B) is the theoretical prediction given by equation (17).

Resetting through feedback signals can be realized in two ways: i) When a neuron in layer $n+1$ has fired, the firing probabilities of all its m input neurons in layer n are set to zero. The input neurons are thereby silenced until the end or the run. ii) The input neurons are only temporarily silenced during a inhibitory period lasting a few time steps. After the inhibitory period, these neurons are allowed to respond again to coincidences. For layer 0, neurons start to fire again with probability p_1. As p_1 is usually larger than p_0, there is less jitter in starting time after an inhibitory period. This leads to a better synchronization between the input neurons and consequently shorter durations of sustained firing, as can be seen in figure 8B.

In some simulations (figure 8), the sustained firing in layer 3 is disabled. This prevents a permanent inhibition of layer 2 once the neuron in layer 3 has started firing. In these simulations, the neuron in layer 3 fires only in response to coincidences and inhibits only temporarily layer 2.

The figures 2A and 2B show the observed firing probabilities at each time-step in two cases, respectively without and with feedback inhibition. Inhibition is permanent and the last layer has sustained firing. In the layers 0, 1 and 2, the feedback seems to reduce the firing probability. Actually, neurons are still firing at p_1 when they are in the "ON" state but they do so at different times in different runs. This is exemplified in figure 2C showing a spike raster of 30 runs in the case with feedback. There are no differences in the time course of the firing probability in the last layer which receives no feedback. After the onset time, the firing probability saturates at p_1.

Figure 3 shows the latencies L_0 and ΔL_n for n=1, 2, 3 in the case of a fixed $p_0 = 0.05$ for various values of p_1. The latencies are measured on histograms produced by simulations with no reset, as in figure 2a. In figure 3 the dependence on p_1 of ΔL_n predicted by equation (28) is shown to account satisfactorily for the dependence observed in simulations. The relative latencies converge all to the same minimum value ΔL_{min} for large values of p_1.

With the small values of τ and p_0, and reminding that $L_0 = k_c(0)$, equation (26) reduces to

$$\Delta L_n \approx L_0 \frac{\ln(1/m)}{\ln(1/2)} \approx 2L_0, \tag{29}$$

in the case $m = 4$. The factor 2 is visible for $p_1 = 1$ in figure 3 and in the figure 4. The dependence of the minimum latency on p_0 given by equation (24) reduces to

$$\Delta L_{min} \approx \frac{\ln(1/m)}{\ln(1-p_0)}. \tag{30}$$

The figure 4 shows a good match between the theoretical ΔL_{min} and the one measured by simulations with $p_1 = 1$.

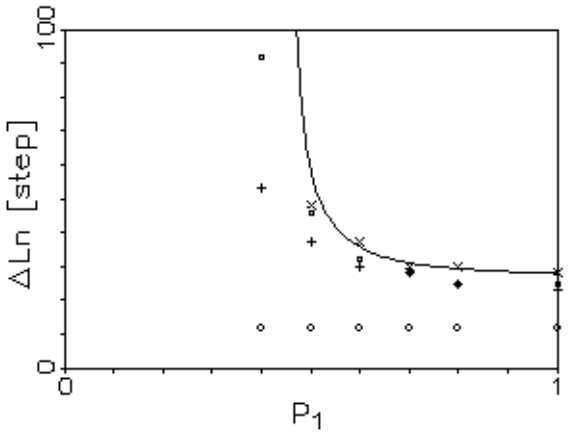


Figure 3: Relative latencies ΔL_n in dependence of the sustained firing probability p_1. $p_0 = 0.05$. Symbols: Data measured on histograms produced by simulation. Circle: L_0, cross: ΔL_1, square : ΔL_2, x: ΔL_3. Full line: Theoretical dependence as in equation (28).

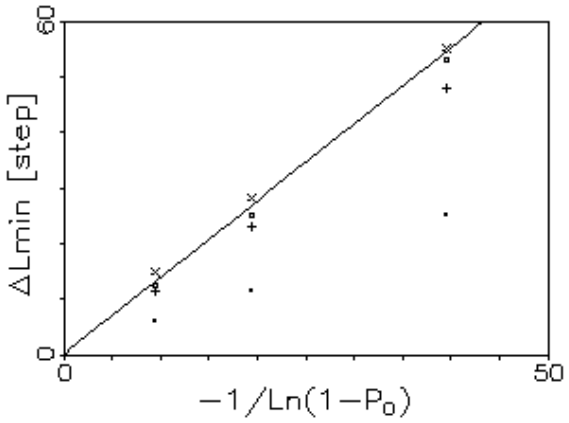


Figure 4: Minimum relative latencies in dependence of p_0, from equation (30). (The jitter in onset time in layer 0 decreases as p_0 increases). Symbols as in Fig. 3 except: dots: L_0.

4 Fitting physiological data

4.1 Fit to latency data in area MST

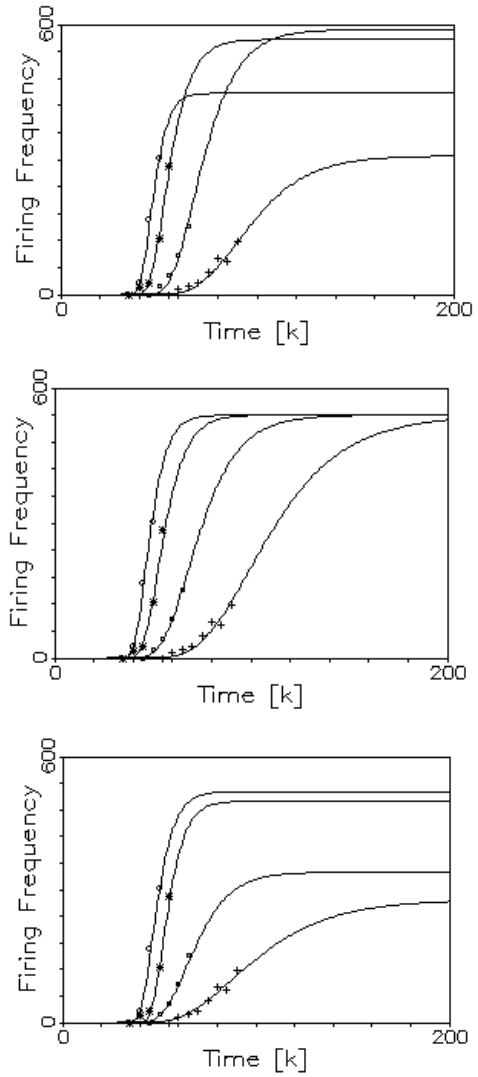
We have attempted to fit the firing probabilities from our theory with experimental data measured during ocular following in area MST of the monkey [fig.8 in Kawano et al., 1994]. It has been proposed that these neurons have a multiplicative AND-function [Saito, 1993] which corresponds well with the hypothesis of our model. In the experiments by [Kawano et al., 1994], an immobile random dot pattern was projected onto the reverse side of a paper screen at time 0, then at time $t = 50ms$, the random dot pattern was accelerated to a maximum angular velocity of 80 deg/sec. The maximum velocity was reached in less than $10ms$. Neurons in MST showed a sharp rise in firing rate approximately $50ms$ after the start of the motion. After an initial peak of the response, the firing rate decreased sharply, indicating inhibition, and then rose again to a lower level where temporal oscillation of the firing rate was observed. Blurring of the stimulus caused an increase in latency and a decrease in maximum firing rate. Blurring was achieved by interposing a plate of ground glass between the projector and the screen, at $1cm$, $3cm$ and $5cm$ from the paper screen.

As for the selection of points to be fitted, we can see in figure 2b) that the feedback starts affecting the shape of the firing probability curve only shortly before the maximum firing rate is reached. Therefore, we have decided to fit only the rising part of the curves observed by [Kawano et al, 1994] in their figure 8, assuming that this phase is only affected by feed-forward processing. We have also subtracted a background firing rate of $7Hz$ from all data. These can be seen in figure 5. Fitting was done with four sets of data referenced by the distance of the ground glass to the screen, $0cm$ corresponding to the sharpest picture and $5cm$ to the most blurred one.

Motion signals which are processed in MST originate in V1 which should therefore correspond to layer 0 in our model. Signals from V1 are processed in MT (layer 1) before reaching MST (layer 2) [Saito, 1993]. Accordingly, we have fitted the expression for $P_2(k)$ to the data (see equation (17)). The three parameters determining the curve are: p_0 (initial jitter), p_1 (level of sustained response), and m (number of inputs of a neuron). A fourth parameter L has been added to account for retinal delays and axonal propagation time up to V1. The interlayer propagation time was assumed to be $\tau = 1$. One time step in the simulation was taken to represent $1ms$ in the data.

The results in [Kawano et al., 1994] are presented in the form of spike density functions that briefly peak at 350Hz for the least blurred case. Such a peak could reflect a higher sustained rate that is capped by the onset of feedback inhibition. It could also be caused by a time-locking of output spikes to the onset of the input, produced by a neuron with a sustained firing of a much lower average frequency. This would concentrate output spikes in a small portion of the post stimulus histogram. For fitting the data with our model assuming Poisson spike trains, we have to adopt the first interpretation. The frequency scale was somewhat arbitrarily normalized so that $600Hz$ in the data correspond to $p_1 = 1$ in the model.

An initial fitting procedure was run with all four parameters free. Best fits were obtained with values of m between $m = 3$ and $m = 6$, depending on the curve. Then we have preset $m = 4$, considering that the degree of blur should not affect the connectivity.



A)

p0	*p1*	*m*	*L*	*ErrorSqAv*	*Blur*
0.214655	0.745379	4	30	0000104	0cm
0.123563	0.944407	4	31	0.000275	1cm
0.077605	0.981172	4	33	0.000305	3cm
0.082796	0.513345	4	31	0.000298	5cm

B)

p0	*p1*	*m*	*L*	*ErrorSqAv*	*Blur*
0.156551	0.9	4	29	0.000282	0cm
0.120577	0.9	4	30	0.000499	1cm
0.068910	0.9	4	28	0.000050	3cm
0.037340	0.9	4	25	0.000378	5cm

C)

p0	*p1*	*m*	*L*	*ErrorSqAv*	*Blur*
0.15	0.869148	4	28	0.000214	0cm
0.15	0.832389	4	33	0.000234	1cm
0.15	0.565189	4	30	0.000143	3cm
0.15	0.461537	4	29	0.000168	5cm

Figure 5: Experimental data of [Kawano et al., 1994] and the fitted firing probability curves (continuous lines) produced by the model. The fitted parameters are given on the adjacent tables. The experimental data points correspond to blurring distances of 0 (o symbol), 1 ($*$ symbol), 3 (box symbol) and 5cm (+ symbol). For fitting, Layer 0 corresponds to V1, layer 1 to MT and layer 2 to MST. The parameters of best fit are shown below the figures. m is the number of synchronous input spikes needed by a neuron to fire, p_0 sets the jitter in V1, p_1 is the level of sustained responses in all layers, L is the retinal reaction time added to the propagation time between retina and V1, $ErrorSqAv$ is the average square error per data point. **A)** Fitting done with fixed parameter: $m = 4$, free parameters p_0, p_1, L **B)** Fixed: $m = 4$, $p_1 = 0.9$. Free: p_0 and L. **C)** Fixed: $m = 4$, $p_0 = 0.15$. Free: p_1 and L.

A second fitting procedure was run with $m = 4$ and all other parameters free. The results are shown in figure 5a. A third fitting procedure was done with $m = 4$ and $p_1 = 0.9$ and the other two parameters free (figure 5b). A fourth fitting was done with $m = 4$ and $p_0 = 0.15$, the remaining two parameters being free (figure 5c).

With fixed input jitter (p_0 preset) or fixed sustained rate (p_1 preset), good fits are obtained. Although there is a slightly better fit when p_1 is variable, the fitting procedure gives no clear cut indication on how blur affects the neuronal responses: increased jitter or reduced firing rates ? Only examination of the spike rasters corresponding to the different conditions of blur may tell if high blur leads to increased jitter or to a reduced instantaneous firing rate. This is because peaks in histograms can be affected by synchronization or de-synchronization effects, e.g. dues to feedback inhibition, as can be seen in figure 2b)). Methods for estimating the instantaneous firing rate are described in [Paulin, 1992].

We may note that the propagation time is consistently found to be $L \approx 30ms$. In the experiments by [Kawano et al, 1994] is was observed that a stationary random dot pattern presented suddenly at time $t = 50ms$ elicited a brief response in the same MST neurons after a delay of $30ms$. Our fitting procedure may indicate that this delay is purely the propagation time of spikes. This would mean that neurons in V1, MT and MST respond within $1ms$ to arriving spikes. In our model, this is only possible if all arriving spikes are perfectly synchronous, excluding the possibility for any jitter. Future physiological data may tell if a sudden presentation of a bright stimulus to dark-adapted subjects can trigger a fast wave of synchronous spikes through the visual system.

Finally, [Kawano et al, 1992] observe that, whatever the blur conditions, after the end of the stimulus, neurons stop firing with the same offset latency, independent of their onset latencies. This onset-offset asymmetry is also observed in an earlier study [Bugmann, 1992b], and in this work (Fig. 8).

4.2 Fit to psychometric backward masking curves

In the figures 2b and 2c it is clearly visible that all neurons in a layer do not start firing a the same time nor stop at the same time. There are cases where some neurons in layer 0 are still firing while some neurons in layer 2 have already started firing. This indicates that part of the visual information has already reached layer 2 while other parts are still localized in layer 0. Therefore, although the network operates in a pure feed-forward way, pieces of information propagate at different speed and do this in a stochastic manner. In different runs, different neurons may be part of a slow or fast path.

This de-localization of the information led us to consider the application of this model to the explanation of the backward masking phenomena. In a typical backwards masking experiment, a subject has to detect, in a first image presented briefly, the presence or the absence of some visual pattern among an array of distractors, e.g. a horizontal bar among vertical bars. The first visual stimulus is shown at time 0 during typically 33ms, a blank screen follows and, at time SOA (Stimulus Onset Asynchrony), a masking stimulus is displayed containing, for instance, an array of crosses (the mask) [Zohary et al., 1990]. As the delay SOA becomes smaller, the performance of the

subject decreases and reaches the chance level where he or she produces 50% of correct responses. The intriguing feature of backwards masking is that a later stimulus can affect the processing of earlier input, and it does so in a probabilistic way. Examples of psychometric response curves are shown on figure 7, where the range 50% to 100% is rescaled between 0 and 1.

In our model, information remains localized in layer 0 during a certain time after the onset of the stimulus, until the target neuron in layer 1 has seen a coincidence and starts firing. While information is held in the first layer, it is possible for activity generated by a new stimulus to interfere with activity representing the previous stimulus. For instance, based on electro-physiological recording, it has been suggested that the effect of the mask is to reduce the duration of the spike trains carrying sensory information [Felsten and Wasserman, 1980]. We have implemented this idea by limiting the duration of the sustained firing in layer 0 to the time SOA. Figure 6 shows that the ending of the firing in layer 0 at a predetermined time is sometimes premature. It results in a number of runs where neurons in layer 1 fail to start firing. In these cases, the neuron in layer 3 cannot starting firing and cannot inhibit its other inputs in layer 2. This suggests that prolonged sustained firing may be the electro-physiological correlate of masking, and has actually been observed by [Roll and Tovee, 1994]

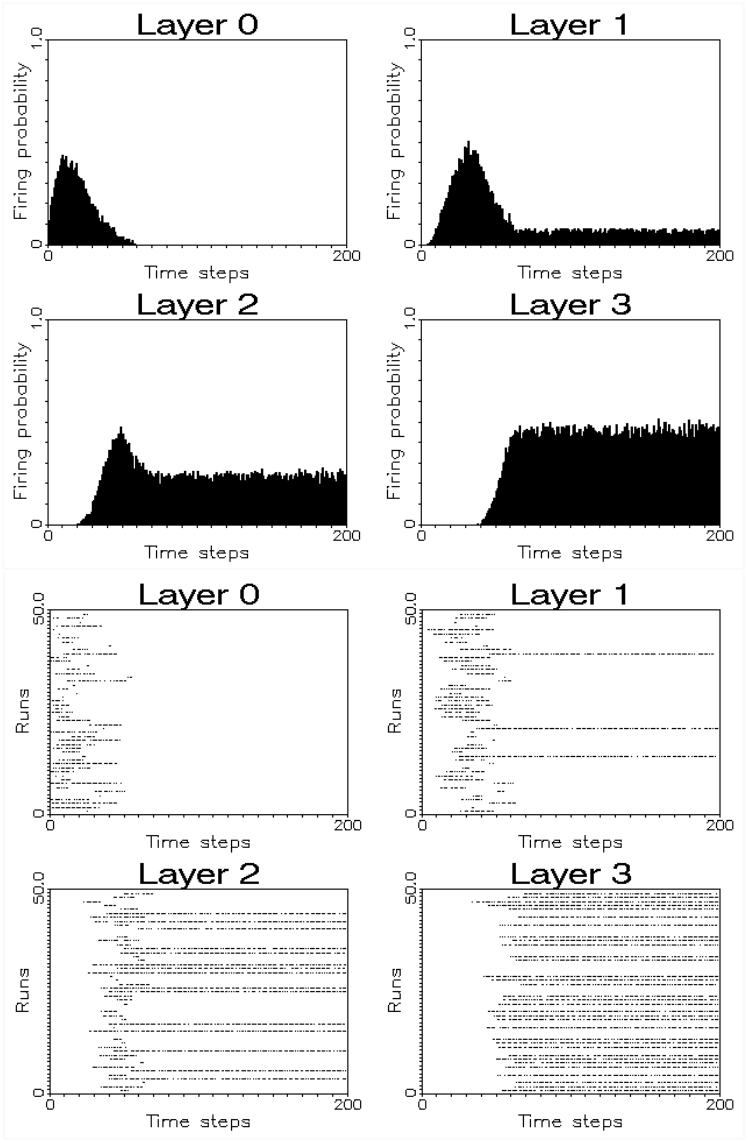


Figure 6: Post stimulus histograms and spike rasters observed when the sustained firing in layer 0 is interrupted at the time step 60. We used $p_0 = 0.08$ and $p_1 = 0.8$.

To determine the psychometric curve predicted by our model, we have assumed that the subject gives a correct response if information reaches layer 3. However, if the neuron in layer 3 fails to enter the sustained mode, the subject has no information on the stimulus and must reply at random. In our model, the mask interrupts firing in layer 0. If this prevents any neuron in layer 1 from starting to fire, then information cannot propagate further. So, for our purpose, the probability of a correct response is dependent on the probability $P_{on}(k_s)$ of all neurons in layer 1 to be ON at the time k_s when the sustained firing in layer 0 is stopped by the mask. We assume that each neuron in layer 1 has m inputs, and N active neurons in layer 1 are needed by subsequent layers. The probability $P_c(k_s)$ to give a correct response if the firing in layer 0 is stopped at time k_s is calculated as follows:

$$P_c(k_s) = P_{on}(k_s) + 0.5(1 - P_{on}(k_s)). \tag{31}$$

So, when $P_{on}(k_s) = 0$ we have $P_c(k_s) = 0.5$ corresponding to a correct response given by chance.

The probability $P_{on}(k)$ that all N neurons in layer 1 are ON at time k is:

$$P_{on}(k) = P_{1,on}(k)^N, \tag{32}$$

where $P_{1,on}(k)$ is the probability of one neuron in layer 1 to be ON. When a neuron in layer 1 is firing with a probability p_1, i.e. it is in a state of sustained firing, it is considered to be on:

$$P_{1,on}(k) = \frac{p_1(k)}{p_1}, \tag{33}$$

where $p_1(k)$ is given by

$$P_1(k) = P_0(k-\tau)^m + (p_1 - p_1^m)\sum_{i=0}^{k-2}(1-p_1^m)^i P_0(k-\tau-i-1)^m, \tag{34}$$

from 17. In order to evaluate this expression for small values of k, we assume that $P_n(k') = 0$ for values of $k' < 1$.

Finally, the probability for a correct response $P_c(k_s)$ when the input spike trains are stopped at time k_s can be found by applying (34), (33) and (32) to (31). This probability depends on 4 parameters: p_0 which determines the jitter in onset time in the first layer, p_1 the level of sustained firing in all layers, m the fan-in of a neuron in layer 1 and N the number of neurons in layer 1 which must be ON to allow information propagation to subsequent layers.

Qualitatively we see that the probability of a correct response is given by the probability of a neuron in layer 1 to be ON to the power N. This power delays the onset-time of the curve and contributes to the threshold of the psychometric curve. The larger the number N of neurons needed in layer 1 to construct the response, the larger the threshold (minimum SOA value for a non-random response). Therefore, a prediction of our model is that, the more complex the information to be stored is, i.e. coded by a larger number of input neurons, the larger the threshold of the psychometric curve.

Note that the time-course of the probability of correct response is a sigmoid function for which a "response onset threshold" could be defined in a somewhat arbitrary way.

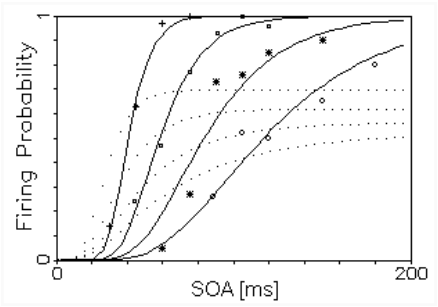


Figure 7: Psychometric curves produced by our model of backward masking. Symbols: Data from fig. 4 in [Bergen and Julesz, 1983]. Full lines: probability for the neuron in layer 3 to start sustained firing in dependence on the stop time SOA in layer 0 (assuming 1 time step = 1 ms). Dotted lines: Time dependent firing probabilities of neurons in layer 1 which produce the best fit to the experimental curves. Other parameters: $p_0 = 0.084$, $m = 4$, $N = 6$.

We have not attempted to do so, but this is clearly possible, e.g. based on a slope intercept with the time axis, leading to the definition of a threshold depending on parameters of the model. It is notable that this model generates a threshold, in contrast to others where the threshold itself is a free parameter, e.g. [Zohary et al. 1990].

To test the model we have used a simple fitting routine to find which parameters allow the reproduction of experimental probabilities of correct responses. We have used data from Bergen and Julez as shown in figure 6 of the paper of [Zohary et al., 1990]. In the experiments of Bergen and Julez the target was always a vertical bar while the distractors differed from the target by various angles. We have taken the data for $20, 30, 60$ and $90 degrees$ (shown in figure 7).

The fitting was done in two steps. First we have allowed all parameters to be adjusted by the fitting routine. This gave us a range of values for the parameters of each of the four curves. Various combinations of parameters allowed similar quality of fitting as determined by the mean square error. In the second phase, we have averaged the parameters over all runs and over all curves and have fixed some parameters to their average values and run the fitting procedure again on the remaining free parameters. It turned out that, with adequate values of p_0, m and N, the only essential free parameter needed to fit all four curves was p_1. By doing so, the average error per point of the fit was not substantially increased compared to the best fit with all parameters free (from 5% to 6% for the $20 deg$ data, from 7% to 10% for the $30 deg$ data, from 3.5% to 3.6% for the $60 deg$ data and from 2.7% to 2.9% for the $90 deg$ data). Using p_0 and p_1 as free parameters gives no better results, although, if layer 0 actually represents some higher area in the visual system, it might be more consistent to have its jitter dependent of its level of sustained firing.

5 Discussion

5.1 Role of the retinal jitter

The model shows two components of the latencies: 1) The initial jitter or time necessary for all m input neurons to be in the ON state. This depends on p_0 in the model. 2) The time for a coincidence to occur which depends on p_1.

The initial jitter, corresponding physiologically to the fluctuations in onset times in retinal ganglion cells [Levick, 1973], determines the absolute minimum in per-layer-computation time ΔL_{min}. If there is no jitter, as during electrical stimulations [Ferster and Lindstrom, 1983], the minimum latency, as defined by equation (24), reduces to the interlayer propagation delay τ. This can be as small as 1 ms in the visual system [Ferster and Lindstrom, 1983] and is 1 time step in our simulations. The value of p_1 is then irrelevant because the first spike generated in the input layer propagates through all layers.

If there is a jitter in the first layer, and p_0 is small, then the standard deviation of the starting times (7) becomes $\sigma_k \approx 1/p_0$ and the minimum latency becomes from equation (24):

$$\Delta L_{min} \approx \tau + \sigma_k \ln(m). \tag{35}$$

This corresponds to the asymptotic behavior near $p_1 = 1$ in figure 3.

As ΔL_{min} increases with the number of inputs, the spread of latencies of neurons in a same layer [Best et al., 1989; Maunsell and Gibson, 1992] could reflect differences in fan-in.

Using physiological values of $\sigma_k = 8ms$ measured by [Levick, 1973] in the cat, one can see in (35) that almost all of the interlayer latency of $\Delta L_n \approx 10ms$ could be accounted for by the retinal jitter. This would imply that the physiological equivalent to p_1 is close to 1 and that m is very small, around 3-4, as used in our simulations.

The number of inputs suggested by our results is much smaller than even the smallest current estimations of approximately 240 inputs in V1 based on count of synapses [Ahmed et al., 1994] or over 40 inputs based on the amplitude of EPSPs [Softky and Koch, 1993]. Note, however, that [Tollhurst et al., 1983] estimated that perceptual decisions are made on the basis of the firing of 2-8 neurons. This is an aspect of the model which deserves further attention.

The probability of a spike-inducing event is taken to be p_1^m in our model, assuming coincidences of one spike from each input within one time step. The value of p_1^m decreases rapidly as the number m of input increases. In biological systems the spiking conditions are likely to be less stringent. For instance, as mentioned in the introduction, the physiological time window for coincidence is closer to 5ms than to 1ms. Therefore, a larger number of input neurons could achieve similar coincidence rates than the small number suggested above, or by the results of fitting the latency curves of [Kawano et al., 1994] (section 4.1). Another point to consider is the low selectivity of neurons in the visual system. Their firing can be induced in some cases by less than 50% of active inputs [Bugmann, 2007]. This can be modeled as input spike configurations with $m-1, m-2, ..., m-\frac{m}{2}$ coincident input spikes being able to generate output spikes. Thus a large number of inputs would find it easier to meet firing conditions. These

same effects could also be used to make lower input rates more effective. Therefore, the same general dynamic behavior of the model could be obtained with smaller values of p_1 and larger values of m.

Further experimental and theoretical work could help determine if the visual system actually operates in a regime of minimal latency where the retinal jitter is the main limiting factor. In this case, an experimental manipulation of the retinal jitter may have a strong observable effect on the measured latencies and even on behavioral reaction times.

5.2 Short-term memory

The need for sustained firing in the model arose from the observation that coincidence detecting neuron models always produce output spike trains with a lower rate than input spike trains [Bugmann, 1991, 1992a, 1992b].

A local recurrent excitatory network that increases the gain of neurons appears therefore to play a crucial role in information propagation. Such a network also gives new memory properties to neurons, allowing the local storage of information. This is exploited in a model for serial and parallel pattern recognition [Bugmann and Taylor, 1993a]. Local memory may also allow the design of asynchronous digital computers. For instance, the network in figure 1 is a pyramid of AND-gates which operates with random inputs without requiring a clock. Indeed, this reduces the computation speed and it is probably not very useful. However, if a computer had to be designed for processing unsynchronized inputs, e.g. using single-electron currents, local memory may become a necessary feature.

In our model, prolonged sustained firing is not observed during normal operation. This is due to the feedback inhibition by an activated target neuron. However, if the target neuron failed to fire, due to an insufficient number of active inputs, or simply because it is temporarily inhibited, then prolonged sustained firing becomes observable in its input layer. This is illustrated in the backward masking simulation (section 4.2) and is experimentally observed by [Tovee and Rolls, 1994] for short SOAs of the mask, a condition where the mask is most effective in interfering with the perception of the stimulus. Thereby, prolonged sustained firing could be an indicator of failure of information propagation. It may be possible to interpret experimental data on short-term memory neurons from this point of view. For instance, if an output neuron is inhibited during the cue - action delay during delayed matching experiments [Funahashi et al.,1992], then the neurons in the previous layer would exhibit sustained firing.

5.3 Is sustained firing necessary?

In our model, the property of sustained firing is introduced to prevent the loss of information in multilayer systems.

There is a question of whether this line of reasoning still hold if one uses more biologically plausible models of neurons, with longer coincidence time windows and lesser requirements on the fraction of input that need to contribute a spike to a coincidence. Could such neurons produce output firing rates of the same level as the input firing rates?

To answer this, we simulated a leaky integrate and fire neurons model with depressing synapses using parameters corresponding to experimental data in [Boudreau and Ferster, 2005] (see more details in Appendix B). The results show that even a neuron with a very low selectivity, able to respond if no more than 44% of its inputs are active, still shows a more than 50% drop in average output frequency, compared to the input frequency, when 100% of its inputs are active. It should be noted that the low average value hides a short onset peak of a very high frequency ($\approx 500Hz$) lasting $\approx 30ms$. Initial peaks are commonly observed in experimental PSTHs. While they can accelerate the production of first spikes, and produce similar initial frequency peaks in their targets, they cannot affect the firing rate during the response plateau that follows. Plateau levels are affected by the plateau level in the input neurons, and these are subject to a layer-by-later drop.

Therefore, the problem of supporting biological plateau levels in more realistic neurons models still points to the need for a frequency-boosting mechanism. Given the numbers above, a gain of 2-3 is necessary to ensure information propagation (each coincidence generate 2-3 output spikes). If this is achieved by a feedback mechanism initiated by an output spike, such a gain will generate self-sustained firing.

In the literature, there are data in support of self-sustained firing, e.g. based on the observation of responses of much longer duration than the stimulus [Rolls and Tovee, 1994].

However, there are also data showing responses that end with a fixed delay after the end of the stimulus [Kawano et al., 1992; Bair et al., 2002]. These are evidence that the presence of the stimulus is needed to maintain the response. However, this is not inconsistent with our model where sustained firing is interrupted by inhibitory feedback from target layers. Here, the presence of the input stimulus is needed to re-start the activity in the network, leading to PSTH similar to those observed in the brain (Fig. 8).

The prolonged response of retinal cells could also cause a response of a longer duration than the stimulus, but, as onsets propagate more slowly than offsets [Kawano et al., 1992; Bugmann, 1992b], a retina-driven response would become shorter and shorter as it progresses through cortical layers and visual areas. Therefore, there is value in a local duration-extending mechanism.

On the other hand, fitting the model to latency and masking data raises the possibility that he level of p_1 may be stimulus dependent. Some experiments also show that the plateau level of the response encodes properties of the stimulus. For instance, in a motion sensitive cell in area DLPN [Kawano, 1992], the plateau level reflects the speed of the stimulus, while the initial transient and latency is mainly affected by contrast. Similar observations are made in [Oram and Perret (1992); Oram, 2010] on face detection cells. In our model, the plateau level reflects several processes, such as self-sustaining mechanisms, inhibitory feedback and re-starting processes. It remains to be examined carefully how these relate to stimulus properties. Despite a recent study on the effect of contrast on response times [van Rossum et al., 2008], much work in this areas remains to be done.

It will also be worth investigating how a this network can actually be implemented with neural hardware. It may have to take the form of a partially inhibited high-gain excitatory feedback loop similar to the microcircuit described by [Douglas and

Martin, 1991]. A dynamic systems approach is examined in [Brody et al., 2003]. It may be worth investigating the connectionist model by [Zipser, 1991] where the level of persistent sustained firing is stimulus dependent.

5.4 Interlayer feedback and its temporal effects

Interlayer feedback is present in our model for consistency reasons, allowing the system to process new information. Due to the pyramidal architecture of the network, the feedback connections are very simple, from target neuron to its source neurons. In a more complex architecture however, if several target neurons share an input neurons in the previous layer, then some logic must be introduced in the feedback circuit. For instance, an input neurons should not be reset by one target neuron if another target neuron has not processed the information yet. Such considerations can rapidly lead to very complex feedback circuits, requiring several layers of neurons. In the visual cortex, layers V and VI are believed to be involved in the management of feedback signals.

Feedback neither affects the velocity of the activity propagation nor the backward masking and is not needed in the theory of these phenomena. Feedback can however generate the complex temporal features of postsynaptic histograms (see e.g. figure 8) and causes a synchronization of input neurons. The synchronization mechanism is relatively simple: Once a neuron in layer $n+1$ has fired, it resets all its m inputs in layer n at the same time. Consequently, these dis-inhibit their inputs in layer $n-1$ and start receiving spike trains. At the end of the inhibitory period, it is likely that all m neurons in layer n start firing at nearly the same time. The required duration of the sustained firing becomes then much shorter, as can be seen in figure 8. More detailed results are presented in [Bugmann and Taylor, 1994b]. A consequence of such a top-down mechanism for synchronization is that synchronization is the by-product of recognition (a neuron in layer $n+1$ firing) rather than its cause [Bugmann, 1997].

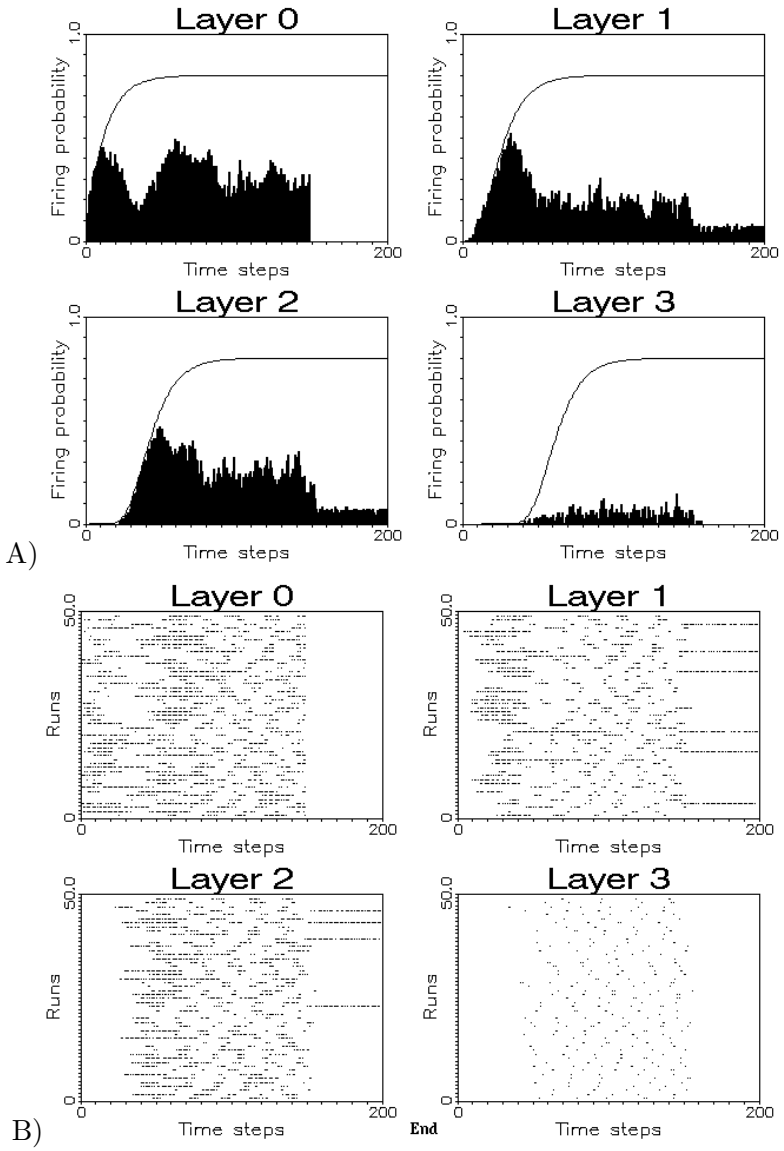


Figure 8: A) Poststimulus histograms produced by the simulation of the model in figure 1 and theoretical firing probabilities (full line). The parameters are $m = 4, p_1 = 0.8, p_0 = 0.08, \tau = 1$. The reset function is used and the neurons are allowed to re-enter a state of sustained firing as soon as a new coincidence occurs. Inhibition lasts for 10 time steps. The neuron in the last layer has no sustained firing (because it would inhibit permanently the neurons in layer 2) and responds only to coincidences. The input (layer 0) is switched off at t=150ms. One can see that firing stops in subsequent layers despite sustained firing. Histograms are based on 100 runs. B) Spike rasters produced by simulations with the same parameters as figure A).

6 Conclusion

The work presented here is a formal analytical description of activity propagation in a simple multilayer network of coincidence-detecting neuron models.

The analyzed multilayer neural network exhibits stochastic propagation of neural activity. Such propagation has interesting features, such as information de-localization, that could explain backward masking. Stochastic propagation is normally observed in simulations of networks of spiking neurons. One of the contributions of this paper is to offer a method for formalizing and quantifying such effects, albeit in a simplified system.

The mathematical model describes how firing frequency levels and onset-time jitter might contribute to latencies in the visual system. It is suggested that retinal jitter plays a major role, which may be confirmed experimentally.

The latency differences between individual neurons could be explained here by differences in the number of their inputs.

The model assumes that all neurons exhibit self-sustained firing triggered by the recognition of features of the stimulus. If not inhibited, such firing can be interpreted as a form of short-term memory. Fitting the model to experimental data produced a good fit, and raises questions of whether levels of sustained firing depend on the quality of the stimulus. It is expected that experimentally measured firing frequencies reflect a combination of multiple factors, such as intrinsic level of sustained firing, duration of the effect of feedback inhibition and time to restart.

An interesting aspect of the model is its potential to describe within a single framework a number of apparently unrelated characteristics of visual information processing, such as latencies, backward masking, synchronization and temporal pattern of post-stimulus histograms. Due to its simplicity, the model can easily be understood, refined and extended.

Acknowledgments
This work was supported in part by SERC/EPSRC under grant GR/H22495.

7 Appendix A

In this appendix, we attempt to find a self-consistent solution of equation (17) valid for large k in the form:

$$P_n(k) \approx p_1 + \varepsilon_n(k). \tag{36}$$

Then, under the assumption that only terms linear in the ε_ns need to be kept in (17), that reduces to

$$\varepsilon_{n+1}(k) = mp_1^{m-1}\varepsilon_n(k-\tau) + (p_1 - p_1^m)mp_1^{m-1}\sum_{i=0}^{k-2}(1-p_1^m)^i\varepsilon_n(k-i-1-\tau). \tag{37}$$

We attempt to justify

$$\varepsilon_n(k) = A_n X_1^k + B_n(k) Y_1^k, \tag{38}$$

where $X_1 = \overline{p_0}$; $Y_1 = 1 - p_1^m = xX_1$ and $x = \frac{1-p_1^m}{\overline{p_0}}$.

From (4) we know that $P_0(k) = p_1 P_0(f,k) + P_0(1,k) = p_1 + (p_0 - p_1)\overline{p_0}^{k-1}$ so that

$$\varepsilon_0(k) = (p_0 - p_1)\overline{p_0}^{k-1}. \tag{39}$$

Thereby $A_0 = \frac{p_0 - p_1}{\overline{p_0}}$ and $B_0(k) = 0$.

Substitution of (39) into (37) leads to

$$\begin{aligned}
\varepsilon_1(k) = & \quad mp_1^{m-1}[\tfrac{p_0-p_1}{\overline{p_0}}\overline{p_0}^{k-\tau} + (p_1 - p_1^m)\textstyle\sum_{i=0}^{k-2}(1-p_1^m)^i \tfrac{p_0-p_1}{\overline{p_0}}\overline{p_0}^{k-i-1-\tau}] \qquad (40)\\
= & \quad mp_1^{m-1}[\tfrac{p_0-p_1}{\overline{p_0}}\overline{p_0}^{k-\tau} + (p_1 - p_1^m)\tfrac{p_0-p_1}{\overline{p_0}}\overline{p_0}^{k-1-\tau}\textstyle\sum_{i=0}^{k-2} x^i]\\
= & \quad mp_1^{m-1}[A_0 X_1^k \overline{p_0}^{-\tau} + A_0(p_1 - p_1^m)X_1^k \overline{p_0}^{-1-\tau}(\tfrac{1-x^{k-1}}{1-x})]\\
= & \quad mp_1^{m-1}(A_0\overline{p_0}^{-\tau} + A_0\tfrac{p_1-p_1^m}{1-x}\overline{p_0}^{-1-\tau})X_1^k - mp_1^{m-1}(A_0(\tfrac{p_1-p_1^m}{1-x})\overline{p_0}^{-1-\tau}\tfrac{1}{x})Y_1^k\\
= & \quad A_1 X_1^k + B_1 Y_1^k,
\end{aligned}$$

with

$$A_1 = mp_1^{m-1}A_0(\overline{p_0}^{-\tau} + \frac{p_1 - p_1^m}{1-x}\overline{p_0}^{-1-\tau}), \tag{41}$$

and

$$B_1 = -mp_1^{m-1}A_0(\frac{p_1 - p_1^m}{1-x})\frac{\overline{{p_0}^{-1-\tau}}}{x}. \tag{42}$$

This process can be continued iteratively, the dependence on k of $B_n(k)$ appearing only for $k \leq 2$. We find

$$\varepsilon_2(k) = A_2 X_1^k + B_2(k) Y_1^k, \tag{43}$$

with

$$A_2 = mp_1^{m-1}[x_1^{-\tau} + \frac{p_1 - p_1^m}{1-x}x_1^{-1-\tau}]A_1, \tag{44}$$

and

$$B_2(k) = mp_1^{m-1}[B_1 Y_1^{-\tau} - \frac{p_1 - p_1^m}{1-x}\frac{A_1}{x}x_1^{-1-\tau} + (p_1 - p_1^m)B_1 Y_1^{-1-\tau}(k-2)]. \tag{45}$$

In general, (38) is verified with

$$A_{n+1} = mp_1^{m-1}x_1^{-\tau}[1 + \frac{p_1 - p_1^m}{1-x}x_1^{-1}]A_n, \tag{46}$$

and

$$B_{n+1}(k) = mp_1^{m-1}[B_n(k)Y_1^{-\tau} - \frac{p_1 - p_1^m}{1-x}\frac{A_n}{x}x_1^{-1-\tau} + (p_1 - p_1^m)Y_1^{-1-\tau}\sum_{i=0}^{k-2} B_n(k-1-i-\tau)]. \tag{47}$$

Inspecting $B_0, B_1, B_2(k), ...$ one finds that for $n > 0$

$$B_n(k) = \sum_{i=0}^{n-1} \gamma_{i,n} k^i. \tag{48}$$

Keeping only the term with the highest power of k:

$$B_n(k) \approx \gamma_{n-1,n} k^{n-1}, \tag{49}$$

where

$$\gamma_{n-1,n} = B_1 m p_1^{m-1}(p_1 - p_1^m)Y_1^{-1-\tau}. \tag{50}$$

So, for large values of k and for $n > 0$, we obtain the approximation

$$\begin{aligned} B_n(k) &\approx B_1[mp_1^{m-1}(p_1 - p_1^m)Y_1^{-1-\tau}]^{n-1}k^{n-1} &(51)\\ &\approx B_1\beta_1^{n-1}k^{n-1}. &(52)\end{aligned}$$

There is a similar but exact expression for A_n for $n \leq 0$

$$\begin{aligned} A_n = &\quad A_0[mp_1^{m-1}(1 + \tfrac{p_1 - p_1^m}{1-x}X_1^- 1), X_1^{-\tau}]^n &(53)\\ = &\quad A_0\alpha_1(x)^n &(54)\end{aligned}$$

so that

$$\varepsilon_n(k) \approx A_0\alpha(x)^n\overline{p_0}^k + B_1(x)\beta_1^{n-1}x^k\overline{p_0}^k. \tag{55}$$

Latencies k_c are defined by

$$P_n(k_c) = \frac{p_1}{2}, \tag{56}$$

or, from (36)

$$\varepsilon_n(k_c) = -\frac{p_1}{2}. \tag{57}$$

To solve equation (55) for k_c, it is convenient to distinguish 3 domains for the values of x:

$$\begin{array}{lll} I & x > 1 & (1 - p_1^m > \overline{p_0}). \\ II & x < 1 & (1 - p_1^m < \overline{p_0}). \\ III & x = 1 & (1 - p_1^m = \overline{p_0}). \end{array} \tag{58}$$

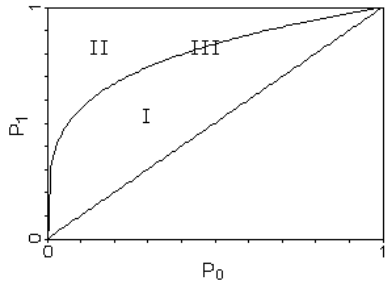


Figure 9: Domains for the parameter x.

Reminding that $p_1 > p_0$ is required for $\varepsilon_n(k)$ to be < 0, the three domains correspond to the areas marked in figure 9. For $p_1 < p_0$ we would observe a decrease in firing rate in post-stimulus histograms.

Domain I

For $x > 1$ and large values of k_c, the second term in the right hand side of (55) is dominant:

$$\varepsilon_n(k) \approx -B_1(x)\beta_1^{n-1}x^{k_c}\overline{p_0}^{k_c}. \tag{59}$$

Neglecting the term in $\ln(k_c)$ one finds

$$k_c \approx \frac{\ln(\frac{p_1}{2}) - \ln(-B_1(x)\beta_1^{n-1})}{\ln(1-p_1^m)}, \tag{60}$$

and the interlayer latency difference becomes

$$\begin{aligned} \Delta L_n &= k_c(n) - k_c(n-1) && (61)\\ &\approx -\frac{\ln(\beta_1)}{\ln(1-p_1^m)} && (62)\\ &\approx \tau + 1 - \frac{\ln(m) + (m-1)\ln(p_1) + ln(p_1 - p_1^m)}{\ln(1-p_1^m)}. && (63) \end{aligned}$$

Assuming further $p_1^m << 1$ and $\ln(p_1) \approx 0$ one gets

$$\Delta L_n \approx \tau + 1 + \frac{\ln(m)}{p_1^m)}. \tag{64}$$

Domain II

For $x < 1$, the second term on the right hand side of (55) behaves as $k_c^{n-1}x_{k_c}$. It goes to zero for $k_c \to \infty$ and has a maximum at $k_m = (n-1)/\ln(1/m)$. Then for values of $k_c >> k_m$ we have

$$\varepsilon_n(k) \approx A_0\alpha(x)^n\overline{p_0}^{k_c}. \tag{65}$$

Then

$$k_c \approx \frac{\ln(\frac{p_1}{2}) - \ln(-A_0) - n\ln(\alpha(x))}{\ln(\overline{p_0})}, \tag{66}$$

$$\Delta L_n \approx \tau - \frac{\ln(m) + (m-1)\ln(p_1) + \ln(\frac{p_1 - p_0}{p_1^m - p_0})}{\ln(\overline{p_0})}. \tag{67}$$

In the limit $p_1 \to 1$, (67) reduces to (24).

Domain III

For $x = 1$ and large k_c, the first term in the right-hand side of (55) is still dominant and we en up with same expression as in domain II.

8 Appendix B

This appendix describes the simulation of the firing rate of a more realistic neuron model (than a simple coincidence detector).

Using a Leaky Integrate and Fire (LIF) neuron as in [Bugmann, 2012], partial reset to 91% of the firing threshold. Depressing synapses were used with parameters inspired by [Boudreau and Ferster, 2005]. The fraction of weight depressed by each input spike was 75%. The recovery time constant was $80ms$. The somatic potential decay time constant was $\tau_{RC} = 20ms$. These values produced a reasonable fit to their figure 9F.

8.1 Initial conditions

A notable observation in [Boudreau and Ferster, 2005] is that V1 synapses are already depressed to some degree at the start of the stimulus, due to the background activity of LGN cells. To determine what value to use, we conducted simulations using inputs firing at their average background figure of $10Hz$. The result was an asymptotic weight value of 23% of its non-depressed value (77% depression).

8.2 Weights for maximal selectivity

We then set the initial weights of a neuron with $m = 25$ inputs to 23% of a value w and started the simulation with each input producing Poisson spikes trains at $200Hz$. We increased the value of w to find the value w_1 for which the neuron started firing. We then reduced progressively the number of active inputs to find the lowest value m' for which the neuron still responds. The selectivity S of the neuron was calculated as:

$$S = 1 - \frac{m - m'}{m}. \tag{68}$$

We then repeated the procedure for input firing rates of $200, 150, 100, 50, 25Hz$.

We found $S = 0.96$ for $w = w_1$ and the neuron did only respond to inputs rates of $200Hz$. Each run simulated 1 second of real-time to determine the average output firing rate.

8.3 Low selectivity regime

We then set $w = 2.2w_1$ and repeated the above measurements. Table 1 summarize the results.

Table 1: Results of simulating a low selectivity neuron

Input Frequency [Hz]	Av. Output Frequency [Hz]	Selectivity S (Equ. 68)
200	65	0.4
150	57	0.44
100	44	0.48
50	20	0.6
25	3	0.77

The results show that the weight value selected coresponds to a neuron with very low selectivity, able to start firing when only 44% of inputs are active. This is plausible for visual neurons [Bugmann, 2007]. Despite such a low selectivity, the output firing rate of the neuron is always less than half of the input firing rate.

The low average hides a strong frequency peak lasting approximately 30ms for the 200Hz case, where the instantaneous rate is above 500Hz. This peak is due to the larger initial EPSP sizes, despite initial depression, the low selectivity and partial reset that makes it easier for a spike to be produced immediately after a previous one.

9 Nomenclature

k : Time (in discrete steps).

$k_c(n)$: Time at which the firing probability of a neuron in layer n reaches 1/2 of the sustained rate p_1.

ΔL_n : Interlayer latency $= k_c(n) - k_c(n-1)$

ΔL_0 : Time for the firing probability of an input neuron to reach 1/2 of the sustained rate p_1 (formaly equal to $k_c(0)$).

m : Number of neurons in layer $n-1$ from which a neuron in layer n receives inputs.

n : Layer identification ($n = 0$: Input layer)

p_0 : Initial firing probability of input neurons.

p_1 : Firing probability during sustained firing.

$P_n(k)$: Probability that a neuron in layer n produces a spike at time k.

$P_n(0,k)$: Probability that a neuron in layer n has produced no spike yet up to time k included.

$P_n(1,k)$: Probability that a neuron in layer n produces its first spike at time k.

$P_n(c,k)$: Probability that a neuron in layer n experiences a coincidence of m input spikes at time k and thus produces and output spike at time k.

σ_k : Standard deviation of the starting time of the sustained firing in layer 0.

τ : Propagation time delay between 2 layers, in time steps.

10 References

Ahmed B., Douglas R.J., Martin K.A. and Nelson C. (1994) "Polyneural innervation of spiny stellate neurons in cat visual cortex", J. Comp. Neurol., 341, 39-49.

Bair W., Cavanaugh J.R., Smith M.A. and Movshon J.A. (2002) "The timing of response onset and offset in macaque visual neurons". J. Neuroscience, 22, 3189-3205.

Bergen J.R.and Julesz B. (1983) "Rapid discrimination of visual patterns." IEEE Trans. SMC-13, 857-863, 1983.

Best J., Reuss S. and Dinse H.R.O. (1986) "Lamina-specific differences of visual latencies following photic stimulations in the cat striate cortex", Brain Research, 385, 356-360

Brody C.D., , Romo R., Kepecs A. (2003) "Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations" Current Opinion in Neurobiology, 13 (2) 204211.

Boudreau C.E. and Ferster D.E. (2005) "Short-term depression in thalamocortical synapses of cat primary visual cortex" J. Neurosci. 25(31):7179-7190.

Bugmann G. (1991) "Summation and multiplication: Two distinct operation domains of leaky integrate-and-fire neurons", Network, 2, 489-509.

Bugmann G. (1992a) "Multiplying with neurons: Compensation of irregular input spike trains by using time-dependent synaptic efficiencies", Biol. Cybern., 68, 87-92.

Bugmann G. (1992b) "The neuronal computation time" in Aleksander I. and Taylor J.G. (eds) "Artificial neural networks II", Elsevier, 861-864.

Bugmann G. (1997) "Binding by Synchronisation: A Task Dependence Hypothesis" Brain and Behaviour Sciences, 20, pp. 685-686.

Bugmann G. (2012) "Modelling fast stimulus-response association learning along the occipito-parieto-frontal pathway following rule instructions." Brain Research, vol.1434, pp. 73-89.

Bugmann G. and Taylor J.G. (1993a) "A stochastic short-term memory using a pRAM neuron and its potential applications" in Beale R. and Plumbley M.D. (eds) "Recent advances in Neural Networks", Prentice Hall, In press. Also avialable as Research Report NRG-93-01, School of Computing, University of Plymouth, Plymouth PL4 8AA, UK.

Bugmann G. and Taylor J.G. (1993b) "A model for latencies in the visual system" in Gielen S. and Kappen B. (eds), Proc. of the Int. Conf. on Artificial Neural Networks, Amsterdam (ICANN'93), Springer, p.165-168.

Bugmann G. and Taylor J.G. (1994a) "Role of Short-term memory in neural information propagation", Extended Abstract Book of the Symposium on Dynamics of Neural Processing, Washington DC, June 6-8, p.132-136.

Bugmann G. and Taylor J.G. (1994b) "A top-down model for Neuronal Synchronization", Research Report NRG-94-02, School of Computing, University of Plymouth, Plymouth PL4 8AA, UK.

Bugmann G., and Taylor J.G. (2005) "A Model of Visual Backward Masking" Biosystems, 79, 151-158.

Bugmann G. (2007) "Determination of the fraction of active inputs required by a neuron to fire." Biosystems, 89, Issues 1-3, 154-159.

Bugmann G. (2009) "A Neural Architecture for Fast Learning of Stimulus-Response Associations." Proceeding of IICAI'09, Tumkur (Bangalore), pp 828-841

Bugmann G., Christodoulou C. and Taylor J.G. (1997) "Role of Temporal Integration and Fluctuation Detection in the highly irregular firing of a Leaky Integrator Neuron Model with Partial Reset" Neural Computation, 9, pp. 985-1000.

Burgess N. and Hitch G.J. (1992) "Towards a network model of the articulatory loop", J. of Memory and Language, 31, 429-460.

Burgi P.-Y. and Pun T. (1991) "Temporal analysis of contrast and geometrical selectivity in the early visual system", in: Blum. P (ed) "Channels in the visual nervous system: Neurophysiology, Psychophisics and models", Freund, London, p.273-288.

Celebrini S., Thorpe S., Trotter Y. and Imbert M. (1993) "Dynamics of orientation coding in area V1 of the awake primate". *Visual Neuroscience*, 10 , pp 811-825.

Douglas R.J. and Martin K.A.C. (1991) "A functional microcircuit for cat visual cortex", J. Physiol., 440, 735-769.

Felsten G. and Wasserman G.S. (1980) "Visual masking: Mechanisms and theory." Psychological Bulletin, 88, 329-353.

Ferster D. and Lindstrom S. (1983) "An intracellular analysis of geniculo-cortical connectivity in area 17 of the cat", J. Physiol., 342, 181-215.

Funahashi S., Bruce C.J. and Goldman-Rakic P. (1989) "Mnemonic Coding of Visual Space in the Monkey's Dorsolateral Prefrontal Cortex" J. of Neurophysiology, 61, 331-349.

Fuster J.M., Bauer R.H. and Jervey J.P. (1981) "Effects of cooling inferotemporal cortex on performance of visual memory tasks", Experimental Neurology, 71, pp. 398-409.

Garey L.F., Dreher B. and Robinson S.R.(1991) "The organization of the visual thalamus", Chap. 3 in: Dreher B. and Robinson S.R. (eds) "Neuroanatomy of the visual pathways and their development", Vol. 3 of the series "Vision and Visual Dysfunction", MacMillan, 176-234.

Gorse D. and Taylor J.G. (1990a) "A general model of stochastic neural processing", Biol. Cybern., 63, 299-306.

Gorse D. and Taylor J.G. (1990b) "Hardware realisable training algorithms. Proc. Int. Conf. on Neural Networks, Paris, France, 1990, pp. 821824.

Granger R., Ambros-Ingerson J., Staubli U. and Lynch G. (1990) "Memorial operation of multiple, interacting simulated brain structures", in Gluck M.A. and Rumelhart D. (eds) "Neuroscience and Connectionist Theory", Lawrence Erlenbaum Associates, London, 95-129.

Grossberg S. (1982) "Contour Enhancement, Short-Term Memory, and Constancy in Reverberating Neural Networks" reprinted in Grossberg S., "Studies of Mind and Brain", D.Reidel Publishing Company, Boston, 334-378.

Houghton G. (1990) "The problem of serial order: A neural network model for sequence learning and recall", in Dale R., Mellish C. and Zock M. (eds.) "Current research in natural language generation", Academic Press, London, 287-319.

Humphrey G.W. and Muller H.J. (1993) "SEarch via Recursive Rejection (SERR): A connectionist model of visual search", Cognitive Psychology, 25, 43-110.

Kawano K., Shidara M. and Yamane S. (1992) "Neural activity in Dorsolateral

Pontine Nucleus of Alert Monkey During Ocular Following Response", J. of Neurophysiol., 67, 680-703.
Kawano K., Shidara M., Watanabe Y. and Yamane S. (1994) "Neural Activity in Cortical Area MST of Alert Monkey During Ocular Following Responses", J. of Neurophysiology, 71, 2305-2324.
Lesica N.A. and Stanley G.B. (2004) "Encoding of Natural Scene Movies by Tonic and Burst Spikes in the Lateral Geniculate Nucleus". The Journal of Neuroscience, 24(47):1073110740.
Levick W.R. (1973) "Variation in the response latency of cat retinal ganglion cells", Vision Res., 13, 837-853.
Logothetis N.K., Pauls J. and Poggio T. (1995) "Shape representation in the inferior temporal cortex of monkeys" Current Biology, 5:552-563
Maunsell J.H.R. and Gibson J. (1992) "Visual response latencies in striate cortex of the macaque monkey", J. Neurophysiol., 68, 1332-1344.
McCormick D.A., Shu Y., Hasenstaub A., Sanchez-Vives M., Badoual M. and Bal T. (2003) "Persistent Cortical Activity: Mechanisms of Generation and Effects on Neuronal Excitability". Cerebral Cortex, 13 (11) 1219-1231.
Muller J.R., Metha A.B., Krauskopf J., and Lennie P. (2001) "Information conveyed by onset transients in responses of striate cortical neurons" J. of Neuroscience, 21(17), 6878-6990.
Oram M.W. (2010) "Contrast induced changes in response latency depend on stimulus specificity." Journal of Physiology - Paris, 104, 167175
Oram M.W and Perrett D.I. (1992) "Time course of neural responses discriminating different views of face and head", J. of Neurophysiology, 68, 70-84.
Paulin M.G. (1992) "Digital filter for firing rate estimation", Biol. Cybernetics, 66, 525-531.
Reich D.S. , Mechler F. and Victor J.D. (2001) "Temporal Coding of Contrast in Primary Visual Cortex: When, What, and Why" J Neurophysiol 85:1039-1050.
Rolls E.T. and Tovee M.J. (1994) "Processing speed in the cerebral cortex and the neurophysiology of visual masking" Proc. R. Soc. Lond. B vol. 257 no. 1348, 9-15.
van Rossum M.C.W., van der Meer M.A.A., Xiao D. and Oram M.W. (2008) "Adaptive integration in the visual cortex by depressing recurrent cortical circuit" Neural Computation 20:1847-1872.
Roy S.A. and Alloway K.D. (2001 "Coincidence Detection or Temporal Integration? What the Neurons in Somatosensory Cortex Are Doing" The Journal of Neuroscience, 21(7):24622473.
Saito H.-A.(1993) "Hierarchical Neural Analysis of Optical Flow in the Macaque Visual Pathway", in: Ono. T, Squire L.R., Raichle M.E., Perret D.I. and Fukuda M. (eds) "Brain Mechanisms of Perception and Memory", Oxford Univ. Press, 121-140.
Sherman S.M. and Koch C. (1986) "The control of retinogeniculate transmission in the mamalian lateral geniculate nucleus", Exp. Brain Res., 63, 1-20.
Softky W.R. and Koch C. (1993) "The highly irregular firing of cortical-cells is inconsistent with temporal integration of random EPSPs", J. Neurosci., 13, 334-350.
Thomson AM, Deuchars J. (1994) "Temporal and spatial properties of local circuits in neocortex." Trends Neurosci. 17: 119-126

Thorpe S.J. and Imbert M. (1989) "Biological constraints on connectionist modelling", in Pfeifer R. et al. (eds.) "Connectionnism in perspective", Elsevier, 63-92.
Tolhurst D. J., Movshon J. A. and Dean A. F. (1983) "The statistical reliability of signals in single neurons in cat and monkey striate cortex." Vision Research 23, 775-785.
Troyer T.W. and Miller K.D.(1997) "Physiological gain leads to high ISI variability in a simple model of a cortical regular spiking cell". Neural Computation, 9(5):971-983.
Van der Loos, H. and Glaser, E.M. (1972) "Autapses in neocortex cerebri: Synapses between a pyramidal cell's axon and its own dendrites", Brain Res., 48, 355-360.
Verri A., Straforini M. and Torre V. (1992) "Computational aspects of motion perception in natural and artificial neural networks", in Orban G.A. and Nagel H.H. (eds.) "Artficial and biological vision systems", Springer, 71-92.
Zipser D. (1991) "Recurrent network model of the Neural mechanism of short-term active memory", Neural Computation, 3, 179-193.
Zohary E., Hillman P. and Hochstein S. (1990) "Time course of perceptual discrimination and single neuron reliability", Biological Cybernetics, 62, 475-486.