Report

Rapid neural computation using single spikes and time-dependent EPSP modulation

BURGI, Pierre-Yves, THORPE, Simon J.

Abstract

Evidence for rapid neural computation in visual cortex are now available: Event related potentials in human visual system have revealed a frontal negativity specific to no-go trials involving complex shape categorization that develops roughly after 150 ms stimulus onset. Furthermore, in VI, in the inferotemporal cortex, and in the superior temporal sulcus, cells can discriminate different complex visual stimuli within 5 to 10 ms. Given the anatomical and physiological constraints imposed by our knowledge of the visual system, such data have major implications for theories of visual computation. The suggestion that firing rate is rarely in excess of one spike in a 5-10 ms period in cortex emphasizes this situation. Cortical maps, which reflect a precise topographic neuronal organization where neighboring sites of maximum relative activity on the cortical plane correspond to nearby points in parameter space, might offer a neural substrate for rapid computation.

Reference

BURGI, Pierre-Yves, THORPE, Simon J. *Rapid neural computation using single spikes and time-dependent EPSP modulation*. 1996, 42 p.

RAPID NEURAL COMPUTATION USING SINGLE

SPIKES AND TIME-DEPENDENT EPSP

MODULATION

Pierre-Yves Burgi[†] and Simon J. Thorpe

Centre de Recherche Cerveau and Cognition,

UMR 5549, Faculté de Médecine de Rangueil,

133 Route de Narbonne, 31062 Toulouse CEDEX, France.

[†] To whom correspondence should be addressed.

Short title: Rapid Neural Computation

April 23, 1996

Abstract

Evidence for rapid neural computation in visual cortex are now available: Event related potentials in human visual system have revealed a frontal negativity specific to no-go trials involving complex shape categorization that develops roughly after 150 ms stimulus onset. Furthermore, in V1, in the inferotemporal cortex, and in the superior temporal sulcus, cells can discriminate different complex visual stimuli within 5 to 10 ms. Given the anatomical and physiological constraints imposed by our knowledge of the visual system, such data have major implications for theories of visual computation. The suggestion that firing rate is rarely in excess of one spike in a 5-10 ms period in cortex emphasizes this situation. Cortical maps, which reflect a precise topographic neuronal organization where neighboring sites of maximum relative activity on the cortical plane correspond to nearby points in parameter space, might offer a neural substrate for rapid computation. We designed a neural architecture performing the dot product operation, an operator involved in cortical mapping, using the timing of single spikes in multiple input sources. The key element in such an architecture is that the amplitude of the excitatory post-synaptic potentials is dependent on the number of spikes that entered the cortical layer. We formulate this dependency mathematically and futher show that cortical mapping as defined in this context can use vectors normalized according to a l^1 norm.

1

1 Introduction

Recent experiments using event related potentials to test processing speed of complex natural images in the human visual system have revealed scalp potentials from frontal sites which allow to differentiate between two classes of stimuli (animal versus non-animal) at 150 ms after stimulus onset (Thorpe *et al* 1996). Given the anatomical and physiological current knowledge of the primate visual system (Felleman and Van Essen 1991), we can infer that transmission of form information involves at least 10 synaptic layers, implicating an average processing time of 15 ms per layer. Recent studies on the dynamics of cells' tuning have provided evidence that fast processing is operating in various topographically organized cortical areas. For instance, within the first 20 ms period of their neuronal response, V1 cells develop selectivity to orientations (Vogels and Orban 1991; Celebrini *et al* 1993) and disparities (J. Masanes and Y. Trotter, personal communication). In the inferotemporal cortex (IT), a period of firing as short as 20 ms and associated with the presentations of complex visual stimuli may be sufficient to support object recognition (Tovée *et al* 1993). Such a performance is attested by experiments on visual masking where briefly flashed visual stimuli followed by the presentation of another pattern (mask) 20 ms later does not significantly impede IT neurons to be selective to shapes (Rolls and Tovée 1994; Kovács *et al* 1995). Furthermore, in the superior temporal sulcus

(STS), another topographically organized cortical area which is involved in face recognition (Harries and Perret 1991), 5 ms has been reported to be sufficient for neurons to discriminate reliably different head views (Oram and Perret 1992). In networks of spiking neurons, the firing rate is generally low and the probability that two spikes occur within 10 ms is small (Abeles *et al* 1990). Thus, periods of time as short as 5 ms severely limit the number of spikes effectively used to process the information and excludes cortical processing based solely on mean firing rates, placing new constraints on neural architectures (Thorpe and Imbert 1989; Thorpe 1990; Burgi and Pun 1994).

One of the most described neural organizations in the brain are cortical maps. They transform the representation of sensory stimuli into local cortical activations that represents the computed values of parameters by sites of maximum relative activity. Their topographical organization tells us that nearby positions of the cortical plane correspond to nearby points in parameter space (for a review, see Knudsen *et al* 1987). The retinotopic and tonotopic maps in the visual and auditory cortices, respectively, exemplify this principle, as does the organization of IT, a cortical area involved in object recognition (Kobatake and Tanaka 1994), where cells selective for a range of similar complex stimuli are grouped together in columns (Fujita *et al* 1992). Such a cell organization allows distance in the cortex to be a surrogate for distance in the representational space, as predicted by Kohonen (1982). In an ever changing environment where the nervous system must

3

react rapidly, cortical maps have been suggested to be efficient in handling large amounts of information by minimizing the neuronal wiring required for local operations (Durbin and Mitchison 1990), and by sorting and processing in parallel components of complex stimuli in a feedforward manner (Knudsen *et al* 1987) (such a preset computation does not exclude recurrent mechanisms operate to sharpen cell's selectivity).

Consequently, one of the key question addressed in this paper is whether cortical maps or other neural architecture involving competition among neurons for maximal response to a given stimulus can operate on the basis of single spikes in multiple input sources. Such a competitive mechanism is of prime importance for adjusting the response properties of cortical units (Willshaw and von der Malsburg 1976; von der Malsburg and Willshaw 1977; Kohonen 1982). Its realization in terms of single spikes is shown herein to rely on temporal modulation of excitatory post-synaptic potentials (EPSPs) due to the timing of individual spikes. That the spike's order at the level of a single spike is determinant for the amplitude of EPSP has previously been given experimental support (Thomson and Deuchars 1994; Henry Markram, personal communication). Also, in this paper we wish to extend this principle to a whole cell population, with the additional advantage that this EPSP modulation provides a framework for cortical mapping with no need to assume Euclidean normalization of the input and synaptic strength vectors.

Formalization of cortical mapping compatible with fast time access and

4

non-Euclidean normalization is first presented in Section 2. Application of the formal theory to cortical signals is then proposed in Section 3 and followed by a general discussion in Section 4. It must be noted that the formation of cortical maps (Willshaw and von der Malsburg 1976; von der Malsburg and Willshaw 1977; Kohonen 1982) is not addressed here. One possible biological realization has been proposed by Kohonen (1993) and the formation of these maps in the context of spiking neuron will be treated in a future publication.

2   Theory

Cortical mapping can be enonced as follows: A parameter space representative of input stimuli from a space V is to be mapped onto an array of neurons A. Each point in the parameter space is represented as a vector of activities $\vec{x} = (x_1, x_2, \ldots, x_n)^T \in V$, where n is the dimension of the space V. Components of $\vec{x}$ are activation levels, normally thought of as firing rates of action potentials. Each neuron is labelled by its position $r \in A$, and is connected to n input lines via synaptic strengths represented by a vector $\vec{\omega} = (\omega_1, \omega_2, \ldots, \omega_n)^T \in \mathbf{R}^n$.

Upon presentation of an input vector $\vec{x}$, the problem to solve is to find the best match between this vector and the weight vectors. Assuming neurons sum linearly their inputs (Jagadeesh *et al* 1993), their activation can be

expressed by the dot product of input and synaptic strength vectors. In the case these vectors are normalized according to the Euclidean norm (i.e., $x_1^2 + x_2^2 + \ldots + x_n^2 = C^2$, where C is the norm), the amplitude of the neuronal activation is directly proportional to the similitude existing between these two vectors (Figure 1A). Thus, in a first step, the input vector will activate many neurons. Then, and assuming regional neuronal interactions, only the neuron with the largest activity can be maintained according to a winner-take-all (WTA) mechanism. This mechanism implies lateral inhibition and its implementation for networks of spiking neurons is not treated herein, as the topic has been considered elsewhere (Yuille and Grzywacz 1989; Thorpe 1990; Antón *et al* 1992; Coultrip *et al* 1992).

Figure 1: Around Here.

Consider now the problem of cortical selectivity when vectors are normalized according to the following l^1 norm:

$$\sum_{i=1}^{n} x_i = \alpha \quad and \quad \sum_{i=1}^{n} \omega_i = \beta, \tag{1}$$

where $x_i \geq 0$, $\omega_i \geq 0$, $i = 1 \ldots n$, and α and β are two constants. Then, the following Lemma applies:

Lemma 1 *Let the coordinates of $\vec{x}$ and $\vec{\omega}$ in the orthonormal basis $\{\vec{e}_i\}_{i=1}^{n}$ be $(x_i)_{i=1}^{n}$, $(\omega_i)_{i=1}^{n}$, respectively. Given $\vec{x}$ and $\vec{\omega}$ are normalized according to the l^1 norm: $\sum_{i=1}^{n} x_i = \alpha$ and $\sum_{i=1}^{n} \omega_i = \beta$, $x_i \geq 0$, $\omega_i \geq 0$, $i = 1 \ldots n$, then*

6

for a fixed $\vec{x}$ the dot product $\vec{x} \cdot \vec{\omega}$ is maximum for the vector $\vec{\omega} = \beta \vec{e}_k$, where k verifies $x_k = max\{x_i\}_{i=1}^n$.

Proof

Let define vector $\vec{\omega}$ as $\vec{\omega} = (\beta - \mu)\vec{e}_k + \vec{\epsilon}$, where $\sum_i \epsilon_i = \mu$. Then, $\vec{\omega} \cdot \vec{x} = (\beta - \mu)x_k + \vec{\epsilon} \cdot \vec{x}$. Given $\vec{\epsilon} \cdot \vec{x} = x_k(\mu - \sum_{i \neq k} \epsilon_i) + \sum_{i \neq k} \epsilon_i x_i$, the dot product $\vec{\omega} \cdot \vec{x}$ simplifies to $\beta x_k + \sum_{i \neq k} \epsilon_i(x_i - x_k)$. As $x_k = max\{x_i\}_{i=1}^n$, we have $\sum_{i \neq k} \epsilon_i(x_i - x_k) < 0$. Thus, the dot product $\vec{\omega} \cdot \vec{x}$ is maximum when $\vec{\epsilon} = \vec{0}$, that is, for $\vec{\omega} = \beta \vec{e}_k$.

This Lemma thus states that the neuron with, for instance, the weight vector $\vec{\omega} = (1, 0, \ldots, 0)^T$ gets the largest activation within the array A for any input $\vec{x}$ satisfying $x_1 = max\{x_i\}_{i=1}^n$ (Figure 1B). Consequently, only a subset of neurons in array A can be winners (at most n neurons), which is very limiting, and two similar vectors do not necessarily yield the largest dot product.

Let now rewrite the dot product by assuming some modulatory function operates on all synapses, resulting in the attenuation or amplification of the activation level according to some nonlinear function $\Psi(\cdot)$. We get

$$R = \omega_1 \Psi(x_1) + \omega_2 \Psi(x_2) + \ldots + \omega_n \Psi(x_n). \tag{2}$$

Let now assume that the vectors $\vec{x}$ and $\vec{\omega}$ are normalized according to the l^1 norm definition, with $\alpha = n\bar{x}$, and $\beta = n\bar{\omega}$, where $\bar{x}$ and $\bar{\omega}$ denote

respectively the average spike frequency, during stimulation, and average synaptic weight determined over a population of neurons. Then, the dot product defines in Equation 2 becomes

$$R = \sum_{i=1}^{n-1} \omega_i \, \Psi(x_i) + (n\overline{\omega} - \sum_{i=1}^{n-1} \omega_i) \Psi(n\overline{x} - \sum_{i=1}^{n-1} x_i) \qquad (3)$$

In the case of Euclidean normalization, the neuron for which the weight vector is the most colinear with the input vector is the one that yields the largest dot product (Figure 1A). Because of the non-Euclidean norm definition, this measure of similitude between vectors must be abandoned (as stated by Lemma 1). Consequently, the problem of neuronal selectivity is reduced to finding an adequate function $\Psi(\cdot)$ so that the maximum of neural response R (Equation 3) maps uniquely an input vector onto a synaptic weight vector, and defines a bijective application (Figure 1C). As it will be shown in the next section, rapid computation of Equation 3 will be based on the establishment of an equivalence between the modulation of the activation level and the modulation of the effect of individual spikes on somatic potential.

Let express the condition for a vector $\vec{x}$ to be extremum of R:

$$\vec{\nabla}_{\vec{x}} R(\vec{x}, \vec{\omega}) = \vec{0} \qquad (4)$$

Applying the gradient operator to the function R results in a set of nonlinear

8

equations. For a given $\vec{x}$, the set of equations reduces to a linear system of equations describing the ω_k, $k = 1, \ldots, n-1$, whose solution is

$$\omega_k = n\overline{\omega} \; \frac{1}{\Psi'(x_k)} \; \Big(\sum_{i=1}^{n-1} \frac{1}{\Psi'(x_i)} + \frac{1}{\Psi'(n\overline{x} - \sum_{i=1}^{n-1} x_i)} \Big)^{-1}, \qquad (5)$$

where $\Psi'(x_k) = \partial\Psi(x_k)/\partial x_k = \partial\Psi(x)/\partial x$. One characteristic of this solution is that any pair of two components of $\vec{\omega}$ is also an extremum of the function $R_i^* = \omega_i\Psi(x_i) + (\beta^* - \omega_i)\Psi(\alpha^* - x_i)$, where $\beta^* = \omega_i + \omega_j$ and $\alpha^* = x_i + x_j$, where $i, j = 1, \ldots, n$ and $i \neq j$. Then, an extremum of R is a maximum if R_i^*, $i = 1, \ldots n$ verifies

$$\frac{\partial^2 R_i^*}{\partial x_i^2} \bigg|_{\omega_i} < 0, \qquad (6)$$

where ω_i is defined according to Equation 5, substituting $n\overline{\omega}$ and $n\overline{x}$ by β^* and α^*, respectively. Developing Equation 6, one gets

$$\Psi'(\alpha - x_i)\Psi''(x_i) + \Psi'(x_i)\Psi''(\alpha - x_i) < 0. \qquad (7)$$

From this latter equation, it can be deduced that any monotonic concave decreasing or convex increasing functions, such as shown in Figure 2A, guarantee two properties: (1) the extremum of R at $\vec{\omega}$, as defined in Equation 5, is a maximum and; (2) the application $x \rightarrow \omega$ is a bijection. It must be noted that the absolute value of the maximum depends on $\vec{x}$. This fact implies a gain control must regulate the activation level of the neuron so that

9

the highest firing rate occurs only when its optimal input is present.

One remarkable property of Equation 5 is the correspondence it establishes between input signals, defined in terms of firing rates, and unitless synaptic weights (defined over a specific interval). Moreover, such a correspondence is very robust to variations in the composition of the input vector, as it will be seen in the next section. This property is of prime importance for the self-organization of cortical maps.

Figure 2: Around Here.

3   Cortical Mapping for Single Spikes

In cortex, activation levels are encoded by mean firing rates of individual spikes. When such rates are low, the time period for estimating the activation level by integrating individual spikes is expected to be longer than the 15 ms time period during which the dot product calculation should be completed. Two neurophysiological facts give a partial answer to this apparent contradiction between low temporal spiking rates and fast computation time. First, it has been shown that the mean firing rate of a neuron measured in a 20 ms period is well correlated with the rate measured in a period of several hundreds milliseconds (Tovée *et al* 1993). This means that the interspike interval between the first and second spike should be representative of the mean firing rate. Second, the latency of the first spike for the initial

10

response of cortical neurons has been observed to be (negatively) correlated to the mean firing rate (Vogels and Orban 1991; Celebrini *et al* 1993), that is higher firing rates yield shorter latencies. We now demonstrate that either the latency of the first spike or the interspike interval is compatible with cortical mapping as described in the previous section.

We start by assuming that within a cortical layer, the first spikes related to the stimulus presentation are those that yield the largest EPSPs. Then, because of inhibitory interactions within the cortical layer, and because of channel desensitization, the amplitude of the response to the subsequent spikes will be more depressed. Neurophysiological evidence for such a depressive mechanism has been described at the level of single cells (Thomson *et al* 1993; Henry Markram, personal communication). Let assume that the strength of such a depression can be described by a monotonic increasing function. For the sake of the analysis we choose the exponential function $1 - \exp(-t/\tau)$, where τ is a time constant characterizing the time course of the depression, and $t = 0$ would correspond to the latency of the very first spike. The effect of this depression on the amplitude of the EPSP can be set to be proportional to $A \exp(-t/\tau)$, where A is spike amplitude (assumed to be identical for all spikes).

To use the information contained in the latency of the first spike, we propose as a first estimation that firing rate, x, and latency, l, are linked together by the relation $x = c/l$, where c is a constant scale factor (set to

11

0.1 in our experiments). Given this mean-firing-rate-to-latency relationship, the modulatory function for the activation level x becomes

$$\Psi(x) = A \exp(-\frac{c}{\tau x}). \tag{8}$$

According to this formula, the first spikes contained in high firing-rate-input signals yield larger EPSPs than those contained in lower firing-rate-input signals, and this EPSP modulation is controlled by the time constant τ (which was fixed to 10 ms). Because the above function $\Psi(\cdot)$ is not a pure monotonic decreasing function (Figure 2B), the domain for which the inequality in Equation 7 is verified must be determined. This domain is delimited by the mean firing rates in the range $x \in [\alpha(1 - \sqrt{\Delta})/2, \alpha(1 + \sqrt{\Delta})/2]$, where $\Delta = (\alpha - c/\tau)/(\alpha + c/\tau)$. For $n = 10$, $\bar{x} = 25\,Hz$, and $c/\tau = 10$, firing rates must be comprised between 4.9 and 245 Hz.

Substituting $\Psi(x)$ in Equation 5 with its new definition (Equation 8) and performing the calculation for a large set of input patterns (several thousands) normalized according to the l^1 norm, we obtained the set of points shown in Figures 3A and B, for 10 and 50 inputs, respectively ($\tau = 10$ ms, $\bar{x} = 25$ Hz, $\bar{\omega} = 0.2$). Apparent in these figures is the decrease in the variance of the distribution of the points resulting from an increase in the number of synaptic inputs. To show the relationship between synaptic weight and latency of the first spike, the points plotted in Figure 3A have

Figure 3: Around Here.

been replotted in Figure 3C by converting firing rates into latencies.

Doing the dot product calculation uniquely on the base of the first spike's latency represents an extreme case, which might be the only possibility to account for access time to cortical maps as short as 5 ms. While from a mathematical point of view, the time it takes to perform the calculation should depend on the largest possible latency (which can be several tens of milliseconds), it might be that the most active neuron is determined well before. Also, looking at Figure 3C, one can conclude (for the parameters used) that most input signals have small latencies (< 10 ms). [After looking at cortical physiological recordings (Maunsell and Gibson 1992; Celebrini et al 1993; Nowak et al 1995), we approximated the density of response latencies by a Rayleigh's distribution, that is, $p(l) = l/\sigma^2 \exp(-0.5l^2/\sigma^2)$, with $\sigma = 4\,ms$, which corresponds to a mean firing rate of 25 Hz (with $c = 0.1$ and maximum frequency limited to 100 Hz).] Consequently, a more pragmatic approach consists in establishing, for a given population of neurons, the error rate one commits in determining which neuron is the most active for a given input when the arrival of input spikes is limited by a time window of variable width. Indeed, when all input information is available, there is a one-to-one relation where each neuron $i = 1 \ldots n$, responds with the highest rate to one (among n) specific input x_i, which might not be

13

the case under application of a time window where neuron i might respond with the highest rate to an input x_j, with $j \neq i$. Such error rates were determined by computer simulations for various time windows, and for a network comprising 1000 neurons with 3, 10 and 30 synapses per neuron. The results, shown in Figure 4, reveal that cortical mapping is reasonably accurate (error rate less than 5%) within 4 to 8 ms, depending on the ratio between the number of neurons and the number of synapses. Because the neuron with the highest activation level is also the most likely to fire first, it can veto the activation of neighboring neurons within a few milliseconds (such an efficient implementation of the WTA function has previously been proposed – Thorpe 1990; Coultrip et al 1992; Barnden and Srinivas 1993). Thus, the whole competitive mechanism, that is, the dot product and WTA calculations can be achieved easily within a time period of 15 ms using the information encoded in the latency of the first spikes.

Figure 4: Around Here.

It is certainly optimistic to assume that the latency of the first spike is accurate and reproductible from one trial to the next. Rather, it has been suggested that, although accurate (a point of view not shared by all – Shadlen and Newsome 1994), the repeating firing patterns in cortex might deviate by a few milliseconds (Abeles et al 1993; Mainen and Sejnowski 1995). Thus, we tested the effect of adding a normal-distributed jitter to

14

the firing time of the first spike, with standard deviation in the range 1 to 6 ms. The results are shown in Figure 5A for a network comprising 100 cells, each cell having 30 synapses. From this figure it is apparent that while a standard deviation of 1 ms does not affect the accuracy of cortical mapping, the error rate rises abruptly for larger standard deviations.

Although the latency of the first spike might well permit to find out which neuron responds best to a specific input pattern, the interspike interval (between the first and second spike, second and third spike, and so on) should improve the tolerance to jitters. Neglecting any leakage current at the soma, we can express the equivalent EPSP due to the combining effects of the first and n successive spikes separated by a mean interspike interval ξ as follows

$$\Psi_n(x) = A \exp(-\frac{c}{\tau x}) \frac{(1 - r^{n+1})}{1 - r}, \tag{9}$$

where $r = exp(-\frac{1}{\xi\tau})$. Functions $\Psi_n(x)$ for two ($n = 1$) and three ($n = 2$) spikes are shown in Figure 2B. Because the shape of these functions ressembles that for $n = 0$, the same theory as before applies. It can be verified that the main effect of using several spikes for the dot product calculation is to increase the tolerance to jitters in the spikes' arrival time. Figure 5A shows the low error rate one gets to jitters up to 3 ms standard deviation when two or three spikes are used. Also visible in this figure is the relatively small gain obtained in using the third spike. Indeed, at the time

15

the third spike comes the depression of the EPSP is so large that this spike does not contribute much to the somatic potential. By changing the time constant of the exponential decreasing function, the network better benefits from the second and successive spikes. For instance, when using the second spike, increasing the time constant from 10 to 30 ms results in improving the tolerance to jitters from 3 to 6 ms standard deviation (on the basis of a 5% error rate), as shown in Figure 5B. This figure also shows that an increase in the time constant does not improve the error rate when only the first spike is used.

Figure 5: Around Here.

4   Discussion

Based on current knowledge of the primate visual system (Felleman and Van Essen 1991) and on recent data concerned about processing speed of a categorization task involving complex natural images (Thorpe *et al* 1996), on can infer that no more than 15 ms is left per cortical layer to process visual information. Also, this study aimed at assessing whether neural computations performed within this time period, or shorter than the average interspike interval, can be used efficaciously by the brain. Particularly, we considered the general computation consisting in mapping inputs representative of stimulus features onto sites of maximum relative activity in cortical

16

areas. For instance, cortical visual maps topographically organized (Hubel and Wiesel 1968; Harries and Perret 1991; Fujita *et al* 1992), selectively respond within 5 to 20 ms to orientations (Celebrini *et al* 1993), disparities (J. Masanes and Y. Trotter, personal communication), shapes and faces (Tovée *et al* 1993; Oram and Perret 1992), in spite of the low mean firing rates (less than 100 Hz) characterizing cortical neuronal signals (Abeles *et al* 1990; Rolls and Tovée 1995). Because cortical mapping relies entirely on patterns of connectivity (Knudsen *et al* 1987), its underlying computation is interpretable in terms of competition between input vectors for a given synaptic weight vectors. Resolution of this problem involves the dot product operator, whose biophysical implementation, respectful of the time constraints has been the focus of this paper.

In an Euclidean space, the dot product computation is ruled by the principle that two colinear vectors yield the largest product (Figure 1A). Application of this principle to cortical maps has previously been demonstrated (Kohonen 1982). But to account for a biophysically realistic implementation of cortical maps, Euclidean normalization is not the most desirable as it involves the square operator. We thus used a simpler vector's normalization defined according to the l^1 norm. Consequently, we had to abandon the notion of vector colinearity, and cortical mapping was defined on the basis of an one-to-one bijective application, where one neuron responds with the highest rate to one specific input vector (Figure 1C). In our model, such

17

a bijection has been shown to result from a nonlinear transformation of the input vectors according to any concave monotonic decreasing or convex monotonic increasing functions.

Is there supportive physiological evidence for such a nonlinear transformation of the activation levels? In cortex, activation levels are coded by firing rates. Also, the latency of the first spike has been observed to be correlated to the mean firing rate, but only during the initial phase of neuronal response (Vogels and Orban 1991; Celebrini *et al* 1993). Furthermore, at the level of single cells, the amplitude of the EPSP has been shown to be dependent on the order of the spike's arrival, and to be larger for the first than for the second spike (Thomson *et al* 1993; Henry Markram, personal communication). This has been interpreted in terms of desensitization of receptor channels (Thomson and Deuchars 1994). In this paper, we extended this concept of spike order to a whole cell population, where the precedence of an EPSP at the level of one cell affects the amplitude of the EPSPs at the other cells. Although so far there is no experimental evidence supporting this view, one could think of an inhibitory component that becomes more prominent as the number of EPSP within the cell population increases. Such a scheme would give more importance to signals with short latencies (Burgi and Pun 1994). Furthermore, to account for the repetitive firing of the winner neuron, a periodical resetting of the inhibition must be assumed. Such a resetting could operate every 10 to 20 ms to guarantee computation of

cortical mapping is over (this periods would yield oscillatory potentials in the range 50-100 Hz, which is not incompatible with periodical neuronal activities that have been observed in striate cortex – Maunsell and Gibson 1992; Nowak *et al* 1995).

Modulation of the EPSP's amplitude, coupled with the existence of a correlation between the latency of the first spike and firing rate allowed us to consider the extreme case where stimulus decoding is based on single spikes in multiple sources. Consequently, determination of the most active neuron upon presentation of a stimulus can be very fast (Figure 4). Compared to alternative schemes where the timing of spikes is decoded by well adjusted transmission delays through the axons and dendritic trees (Tank and Hopfield 1987; Hopfield 1995) or using coincidence detectors (Abeles *et al* 1982; Bugmann 1991; Koch and Poggio 1992), our model is less demanding in spike's synchronization. Nevertheless, it has to assume like the other schemes that spike timing is reasonably precise (Abeles *et al* 1993; Mainen and Sejnowski 1995). To test the robustess of our model to variations in the timing of individual spikes, normal distributed jitters with zero mean and standard deviations varying between 1 and 6 ms were added to the spikes' arrival time. The results, shown in Figure 5A, demonstrate that the network works with a low error rate for up to 2 ms standard deviation and more if the ratio number of synapses to number of cells is larger (data not shown). In this paper we also analyzed the case where, besides using the latency of

19

the first spike, the interval interspike between the first and second spike (and second and third spike) is considered for the calculation. The advantage of such an approach is to improve the tolerance to jitters (Figure 5A), up to 6 ms standard deviation, depending on the time constant of the modulatory function of the EPSPs (Figure 5B). The fact that more than one spike is needed to provide a precise cortical mapping when large variations in the timing of single spikes are present is compatible with a recent quantitative study on temporal coding in primates which stipulates that a code based on the time of the first spike can convey at most half of the total information (Heller *et al* 1995). Indeed, the first spike could allow for a coarse decoding, which would eventually refine as more spikes enter the layer and bring complementary information.

It has recently been suggested that if the time advance of a spike is a function of the logarithm of the intensity, then recognition can be done in a scale-invariant fashion (Hopfield 1995). In our model, given variations in stimulus intensity are encoded in firing rate, then an increase (respectively, decrease) in the strength of a stimulus would result in advancing (delaying) each action potential by a certain factor. Consequently, the overall spike rate at a given cortical layer would increase (decrease), and with it the time course of the EPSP's depression. Thus, besides offering fast computation, our scheme could also work in a scale-invariant fashion. It remains to be seen whether the mechanism advocated in this model and by which the

20

amplitude of the EPSPs depends crucially on the order of the spikes within a cell population can be given experimental evidence.

Acknowledgments

References

Abeles M 1982 *Local Cortical Circuits. An Electrophysiological Study* (Berlin: Springer)

Abeles M, Vaadia E and Bergman H 1990 Firing patterns of single units in the prefrontal cortex and neural network models *Network* **1** 13-25

Abeles M, Bergman H, Margalit E and Vaadia E 1993 Spatiotemporal firing patterns in the frontal cortex of behaving monkeys *J. Neurophysiol.* **70** 1629-1638

Antón P S, Granger R and Lynch G 1992 Temporal information processing in synapses, cells, and circuits *Single Neuron Computation* ed T McKenna, J Davis and S F Zornetzer (Boston: Academic Press) pp 291-313

Barnden J A and Srinivas K 1993 Temporal winner-take-all networks: A time-based mechanism for fast selection in neural networks *IEEE Trans. Neural Networks* **4** 844-853

Bugmann G 1991 Summation and multiplication: two distinct operation domains of leaky integrate-and-fire neurons *Network* **2** 489-509

Burgi P-Y and Pun T 1994 Asynchrony in image analysis: using the luminance-to-response-latency relationship to improve segmentation

J. Opt. Soc. Am. A **11** 1720-1726

Celebrini S, Thorpe S J, Trotte, Y and Imbert M 1993 Dynamics of orientation coding in area V1 of the awake primate *Visual Neurosci.* **10** 811-825

Coultrip R, Granger R and Lynch G 1992 A cortical model of the winner-take-all competition via lateral inhibition *Neural Networks* **5** 47-54

Durbin R and Mitchison G 1990 A dimension reduction framework for understanding cortical maps *Nature* **343** 644-647

Felleman D J and Van Essen D C 1991 Distributed hierarchical processing in the primate cerebral cortex *Cerebral Cortex* **1** 1-47

Fujita I, Tanaka K, Ito M and Cheng K 1992 Columns for visual features of objects in monkey inferotemporal cortex *Nature* **360**, 343-346

Harries M H and Perret D I 1991 Visual processing of faces in temporal cortex: physiological evidence for a modular organization and possible anatomical correlates *J. Cognitive Neurosci.* **3** 9-24

Heller J, Hertz J A, Kjaer T W and Richmond B J 1995 Information flow and temporal coding in primate pattern vision *J. Comput. Neurosci.* **2** 175-193

Hopfield J J 1995 Pattern recognition computation using action potential timing for stimulus representation *Nature* **376** 33-36

Hubel D H and Wiesel T N 1968 Receptive fields and functional architecture of monkey striate cortex *J. Physiol.* **195** 215-243

Jagadeesh B, Wheat H S and Ferster D 1993 Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex *Science* **262** 1901-1904

Knudsen E I, du Lac S and Esterly S D 1987 Computational maps in the brain *Ann. Rev. Neurosci.* **10** 41-65

Kobatake E and Tanaka K 1994 Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex *J. Neurophysiol.* **71** 856-867

Koch C and Poggio T 1992 Multiplying with synapses and neurons *Single Neuron Computation* ed T McKenna, J Davis and S F Zornetzer (Boston: Academic Press) pp 315-345

Kohonen T 1982 Self-organized formation of topologically correct feature maps *Biol. Cybern.* **43** 59-69

Kohonen T 1993 Physiological interpretation of the self-organizing map algorithm *Neural Networks* **6** 895-905

Kovács G, Vogels R and Orban G A 1995 Pattern backward masking results from response interruption: direct physiological evidence *Perception* **24 suppl.** 44

Mainen Z F and Sejnowski T J 1995 Reliability of spike timing in neocortical neurons *Science* **268** 1503-1506

Malsburg C von der and Willshaw D J 1977 How to label nerve cells so that they can interconnect in an ordered fashion *Proc. Natl. Acad. Sci. USA* **74** 5176-5178

Maunsell J H R and Gibson J R 1992 Visual response latencies in striate cortex of the macaque monkey *J. Neurophysiol.* **68** 1332-1344

Nowak L G, Munk M H J, Girard P and Bullier J 1995 Visual latencies in areas V1 and V2 of the macaque monkey *Visual Neurosci.* **12** 371-384

Oram M W and Perret D I 1992 Time course of neural responses discriminating different views of the face and head *J. Neurophysiol.* **68** 70-84

Rolls E T and Tovée M J 1994 Processing speed in the cerebral cortex and the neurophysiology of visual masking *Proc. R. Soc. Lond.* B **257** 9-15.

Rolls E T and Tovée M J 1995 Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex *J. Neurophysiol.* **73** 713-726

Sejnowski T J 1995 Time for a new neural code? *Nature* **376** 21-22

Shadlen M N and Newsome W T 199 Noise, neural codes and cortical organization *Curr. Opinion Neurobiol.* 4 569-579

Tank D W and Hopfield J J 1987 Neural computation by concentrating information in time *Proc. Natl. Acad. Sci. USA* **84** 1896-1900

Thomson A M Deuchars J and West D C 1993 Large, deep layer pyramid-pyramid single axon EPSPs in slices of rat motor cortex display paired pulse and frequency-dependent depression, mediated presynaptically and self-facilitation, mediated postsynaptically *J. Neurophysiol.* **70** 2354-2369

Thomson A M and Deuchars J 1994 Temporal and spatial properties of local circuits in neocortex *TINS* **17** 119-126

Thorpe S J and Imbert M 1989 Biological constraints on connectionist models *Connectionism in Perspective* ed R Pfeifer, Z Schreter and F Fogelman-Soulié (Amsterdam: Elsevier) pp 63-92

Thorpe S J 1990 Spike arrival times: a highly efficient coding scheme for neural networks *Parallel Processing in Neural Systems and Computers* ed R Eckmiller, G Hartmann and G Hauske (Amsterdam: Elsevier) pp 91-94

Thorpe S J, Fize D and Marlot C 1996 The speed of processing in the human visual system *Nature* in press

Tovée M J, Rolls E T, Treves A and Bellis R P 1993 Information encoding and the responses of single neurons in the primate temporal visual cortex *J. Neurophysiol.* **70** 640-654

Vogels R and Orban G A 1991 Quantitative study of striate single-unit responses in monkeys performing an orientation discrimination task *Exp. Brain. Res.* **84** 1-11

Willshaw D J and Malsburg C von der 1976 How patterned neural connections can be set up by self-organization *Proc. R. Soc. London B* **194** 431-445

Yuille A L and Grzywacz N M 1989 A winner-take-all mechanism based on presynaptic inhibition feedback *Neural Comput.* **1** 334-347

Figure Legends

Figure 1: Vectors Maximizing the Dot Product. The vector configuration for which the dot product is maximum depends on the norm used. These three figures illustrate two examples of vector configuration $\{\vec{x}_i, \vec{\omega}_i\}$, $i = 1, 2$, for which, given $\vec{x}$, vector $\vec{\omega}$ yields the largest dot product. (A) Euclidean norm: $x_{i1}^2 + x_{i2}^2 = \omega_{i1}^2 + \omega_{i2}^2 = 1$. The maximum occurs for colinear vectors $\vec{\omega}$. (B) l^1 norm: $x_{i1} + x_{i2} = \omega_{i1} + \omega_{i2} = 1$. Lemma 1 applies. (C) Bijective application $\vec{x} \rightarrow \vec{\omega}$; norm as in (B).

Figure 2: Modulatory Functions. (A) Monotonic concave decreasing (solid line) and convex increasing (dotted line) functions. (B) Equivalent EPSP as a function of firing rate for an exponential (convex) increasing time-dependent depression, resulting from one (solid line), two (dashed line), and three (dotted dashed line) spikes.

Figure 3: 2-D Projection of Optimum Vectors. Components of the vectors that verify Equation 4 for the exponential function ($\tau = 10$ ms) have been projected onto a two-dimensional representation. The dimension of the vectors is 10 in (A) and 50 in (B). Abscissa and ordinate represent mean firing rate and synaptic weight, respectively. Visible in these two figures is the smaller dispersion of the points when the vector's dimension is larger. In (C), the vectors projected in (A) have been replotted after conversion of the mean firing rates into latencies.

29

Figure 4: Speed of Cortical Mapping. To assess the time required to perform cortical mapping with confidence, a time window of variable length (shown in abscissa) limited the number of spikes effectively used. Error rates are given for a network comprising 1000 neurons with 3, 10, and 30 synaptic inputs. Increasing the number of synaptic inputs when the number of neurons is fixed dramatically reduces the error rate or, equivalently, speeds up processing.

Figure 5: Robustess of Cortical Mapping to Jitters in Spikes' Arrival Time. Jitters taken from a normal distribution (with zero mean) were added to latency of the first, second and third spike. (A) Dependence of the error rate as a function of standard deviation ($\tau = 10$ ms). Cortical mapping is more robust to jitters when the second (dashed line) and third (dotted-dashed line) spike's arrival time complement the latency of the first spike (solid line). (B) Effect of the time constant of the EPSP modulation on the error rate for a 6 ms standard deviation. A slower establishment of the EPSP's depression leads to a larger tolerance to jitters in the case the interspike interval is used. Error rates are given for a network comprising 100 cells with 30 synapses per cell.
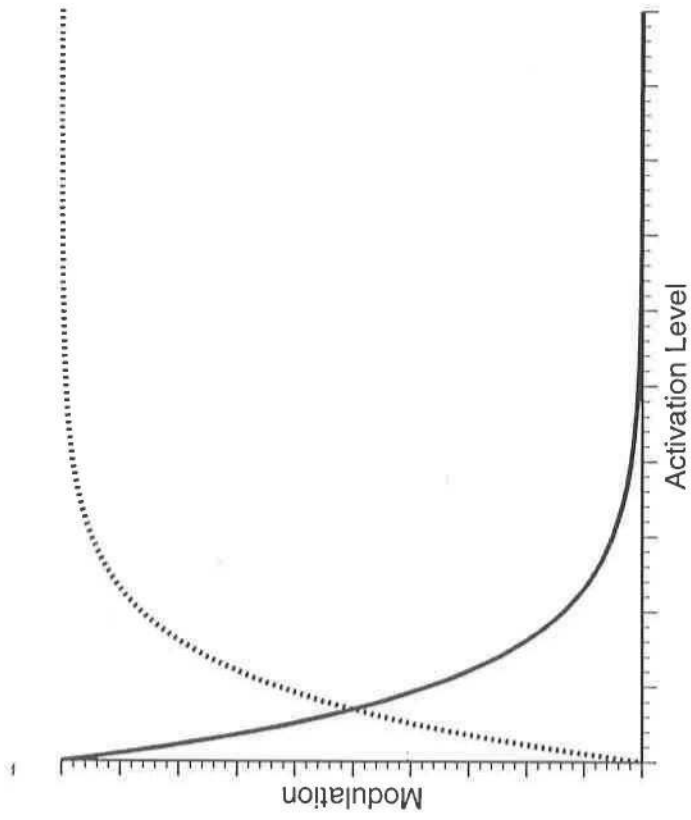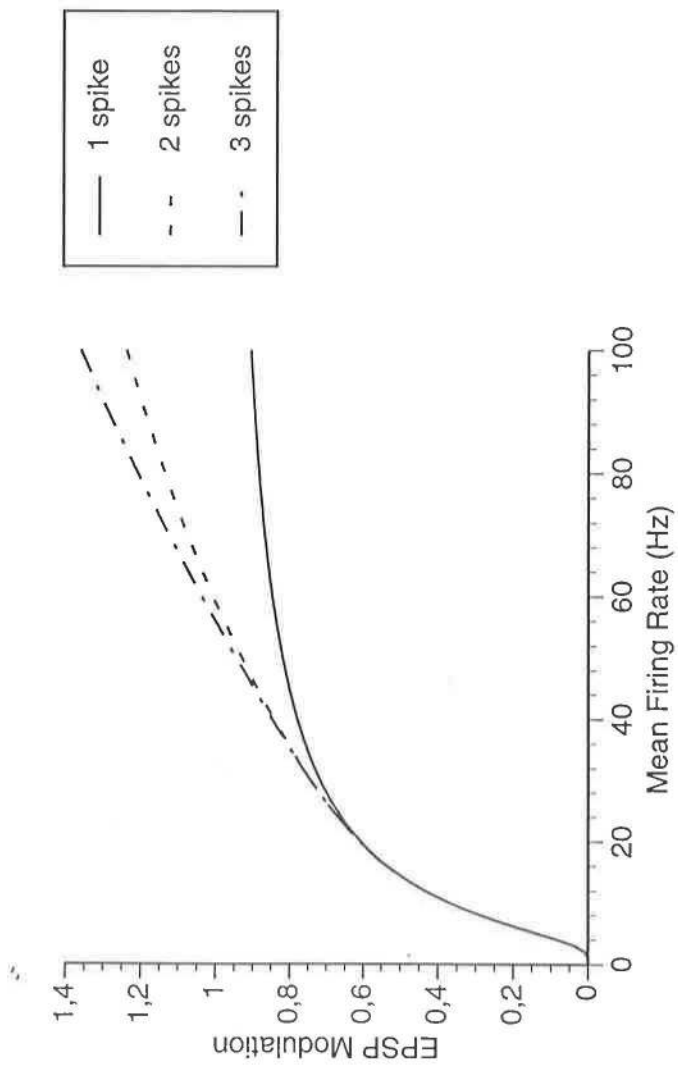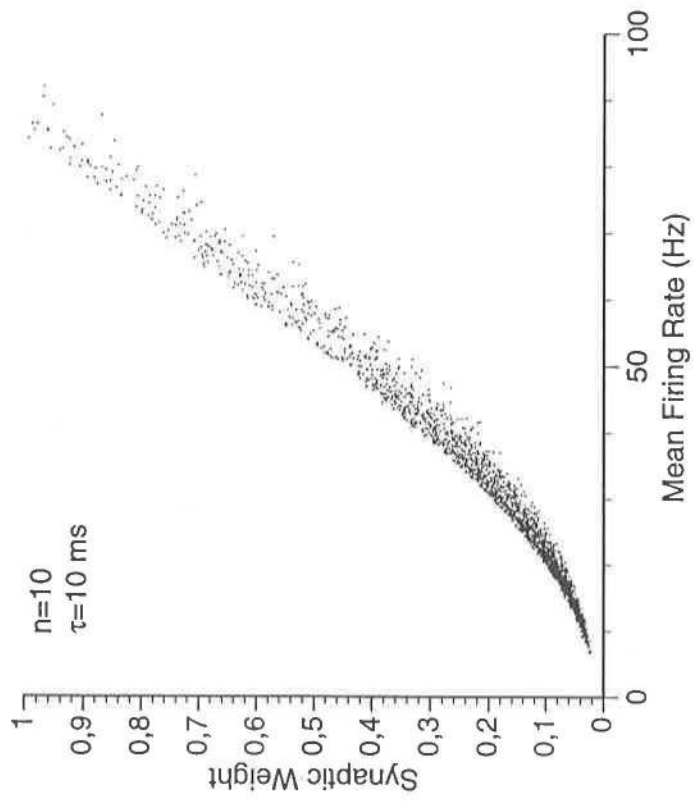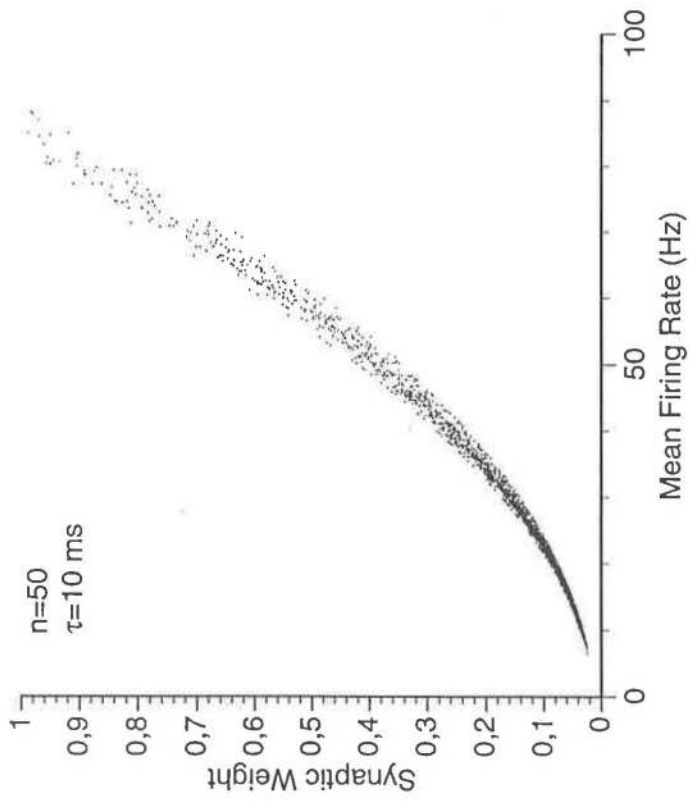
Figure 1A
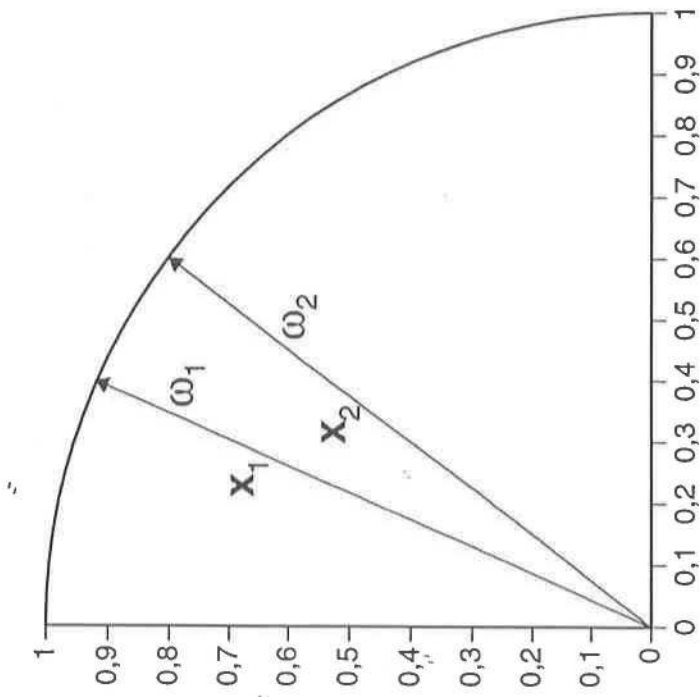
Figure 1B
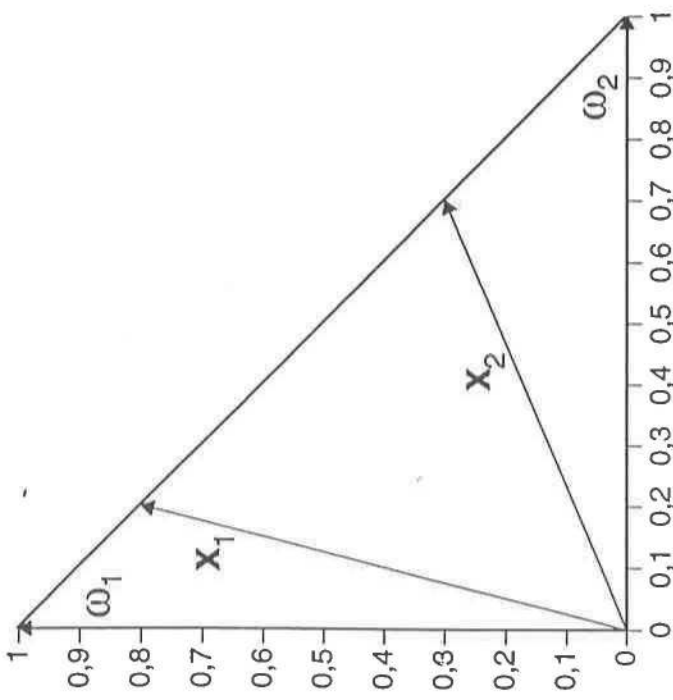
Figure 1C
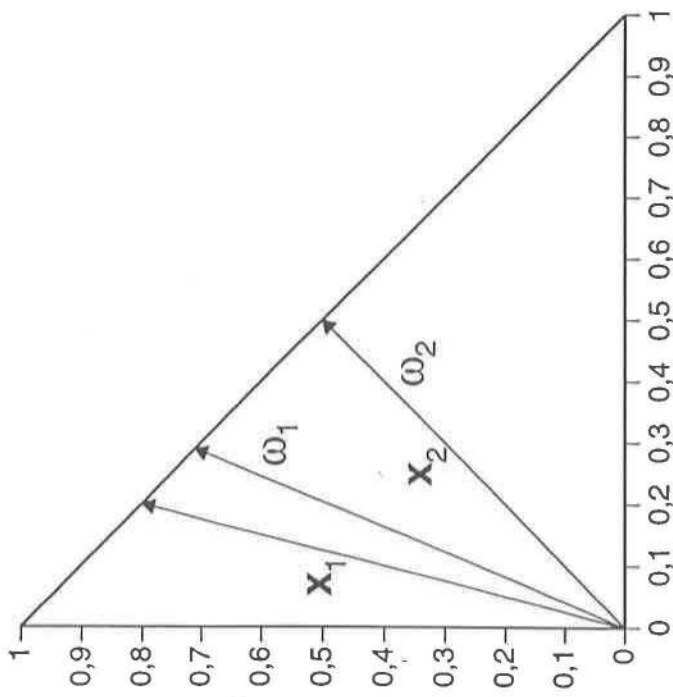
Figure 2A

Figure 2 B

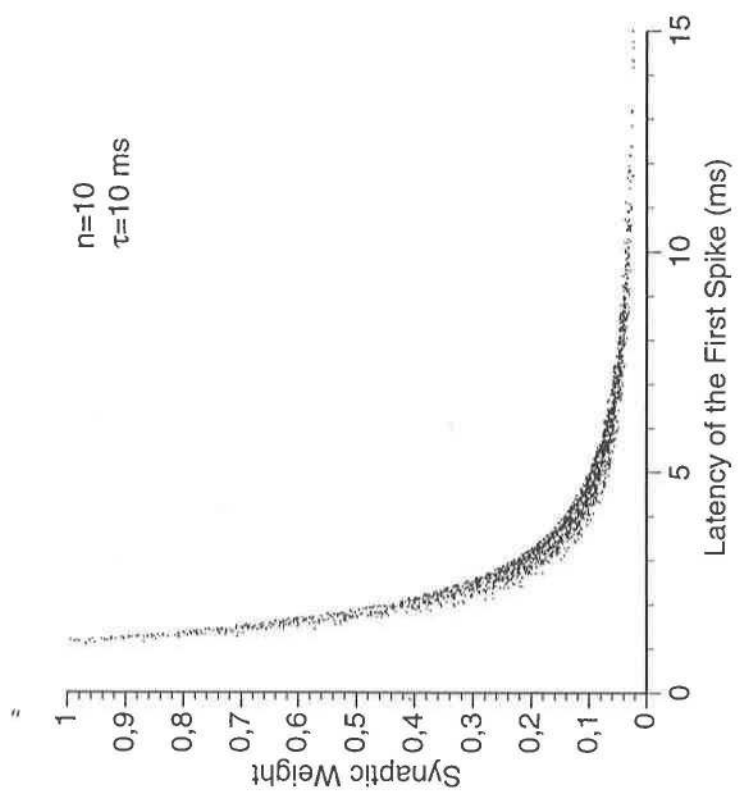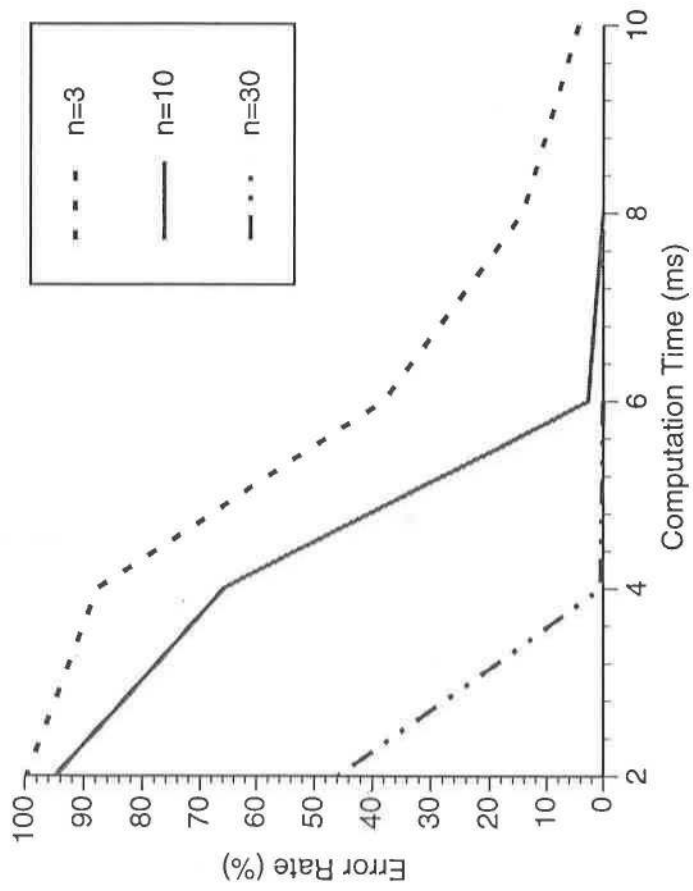Figure 3A

Figure 3B

Figure 3C

n=10
τ=10 ms

Synaptic Weight

Latency of the First Spike (ms)

Figure 4

Figure 5A

Figure 5B