

Fast Sigmoidal Networks via Spiking Neurons

Wolfgang Maass

Institute for Theoretical Computer Science, Technische Universitaet Graz, Graz, Austria

We show that networks of relatively realistic mathematical models for biological neurons in principle can simulate arbitrary feedforward sigmoidal neural nets in a way that has previously not been considered. This new approach is based on temporal coding by single spikes (respectively by the timing of synchronous firing in pools of neurons) rather than on the traditional interpretation of analog variables in terms of firing rates. The resulting new simulation is substantially faster and hence more consistent with experimental results about the maximal speed of information processing in cortical neural systems.

As a consequence we can show that networks of noisy spiking neurons are “universal approximators” in the sense that they can approximate with regard to temporal coding *any* given continuous function of several variables. This result holds for a fairly large class of schemes for coding analog variables by firing times of spiking neurons.

This new proposal for the possible organization of computations in networks of spiking neurons systems has some interesting consequences for the type of learning rules that would be needed to explain the self-organization of such networks.

Finally, the fast and noise-robust implementation of sigmoidal neural nets by temporal coding points to possible new ways of implementing feedforward and recurrent sigmoidal neural nets with pulse stream VLSI.

1 Introduction

Sigmoidal neural nets are the most powerful and flexible computational model known today. In addition they have the advantage of allowing “self-organization” via a variety of quite successful learning algorithms. Unfortunately the computational units of sigmoidal neural nets differ strongly from biological neurons, and it is particularly dubious whether sigmoidal neural nets provide a useful paradigm for the organization of *fast* computations in cortical neural systems.

Traditionally one views the firing rate of a neuron as the representation of an analog variable in analog computations with spiking neurons, in particular, in the simulation of sigmoidal neural nets by spiking neurons.

However, with regard to fast cortical computations, this view is inconsistent with experimental data. Perrett *et al.* (1982) and Thorpe and Imbert (1989) have demonstrated that visual pattern analysis and pattern classification can be carried out by humans in just 100 ms, in spite of the fact that it involves a minimum of 10 synaptic stages from the retina to the temporal lobe. The same speed of visual processing has been measured by Rolls and others in macaque monkeys. Furthermore they have shown that a single cortical area involved in visual processing can complete its computation in just 20 to 30 ms (Rolls 1994; Rolls and Tovee 1994). On the other hand, the firing rates of neurons involved in these computations are usually below 100 Hz, and hence at least 20 to 30 ms would be needed just to sample the current firing rate of a neuron. Thus a coding of analog variables by firing rates is quite dubious in the context of fast cortical computations.

Experimental evidence accumulated during the past few years indicates that many biological neural systems use the timing of single action potentials (or "spikes") to encode information (Abeles *et al.* 1993; Bialek and Rieke 1992; Bair *et al.* 1994; Ferster and Spruston 1995; Hopfield 1995; Kempter *et al.* 1996; Sejnowski 1995; Softky 1994; Thorpe and Imbert 1989; Rieke *et al.* 1996). In addition various experiments have shown that biological neurons are able to fire in vitro with high timing precision (Bryant and Segundo 1976; Segundo 1994; Mainen and Sejnowski 1995).

We show in this article that there exists a completely different way of simulating sigmoidal neural nets with spiking neurons that is based on temporal coding with single spikes (and on temporal coding by synchronous firings of pools of neurons in a more noise-robust interpretation). This simulation is based on the observation that in the presence of some other excitation that moves the membrane potential close to the firing threshold, individual excitatory postsynaptic potentials (EPSPs) or inhibitory postsynaptic potentials (IPSPs) (or volleys of synchronized postsynaptic potentials (PSPs)) are able to *shift* the firing time of a neuron. This mechanism is particularly easy to analyze if we work in a range where all PSPs can be approximated well by linear functions. For this range one can show that the resulting firing time is linearly related to the weighted sum of the firing times of the presynaptic neurons, with the weights corresponding to the efficacies ("strengths") of the involved synapses. We will explain this key observation in a bit more detail at the end of this section, after defining the formal model of a noisy spiking neuron. Although this model ignores many of the intricate details of a biological neuron (e.g., nonlinearities in dendritic integration), one may argue that it underestimates, rather than overestimates, the computational capabilities of a biological neuron. We are not making explicit use of the noise in spiking neurons. Rather we show that the computational mechanism is robust with respect to various types of noise.

A complementary approach for simulating artificial neural nets by spiking neurons with temporal coding has recently been proposed (Hopfield 1995). Hopfield's construction yields basically a simulation of radial basis

function (RBF) units, where the weights of the RBF units are stored in the delays between synapses and soma of a spiking neuron. Hence Hopfield's construction provides an efficient way of implementing a look-up table with spiking neurons (with some very nice invariance regarding the strength of the stimulus). However, in contrast to the construction considered here, his system is based on "grandmother neurons," and it is not geared toward providing an informative output in a situation where the input $\langle s_1, \dots, s_n \rangle$ does not match (up to a factor) one of the fixed set of stored patterns (because it is, for example, a superposition of several stored patterns). Furthermore Hopfield's construction provides no method for simulating multilayer neural nets. In addition, in contrast to our construction, it provides no computational or learning-related role to the efficacy (i.e. strength) of synapses between biological neurons.

We describe in the remainder of this section the precise models for sigmoidal neural nets and noisy spiking neurons that we consider, and at the end of this section describe the key mechanism of our simulation. The main construction of this article is given in Section 2, and our main result is stated in the theorem at the end of that section. In Section 3 we show that this result implies that networks of noisy spiking neurons are universal approximators. We also prove that this result holds for a fairly large class of schemes for temporal coding of analog variables. In Section 4 we briefly indicate some new perspectives about the organization of learning in biological neural systems that follow from this approach.

We point out that this is not an article about biology but about computational complexity theory. Its main results (given in Sections 2 and 3) are rigorous theoretical results about the computational power of common mathematical models for networks of spiking neurons. However, some informal comments have been added (after the theorem in Section 2, as well as in Sections 4 and 5) in order to facilitate a discussion of the biological relevance of this mathematical model and its theoretical consequences.

The computational unit of a sigmoidal neural net is a sigmoidal gate (σ -gate) G , that assigns to analog input numbers $x_1, \dots, x_{n-1} \in [0, \gamma]$ an output of the form $\sigma(\sum_{i=1}^{n-1} r_i \cdot x_i + r_n)$. The function $\sigma: \mathbf{R} \rightarrow [0, \gamma]$ is called the activation function of G , r_1, \dots, r_{n-1} are the weights of G , and r_n is the bias of G . These are considered adjustable parameters of G in the context of a learning process. The parameter $\gamma > 0$ determines the scale of the analog computations carried out by the neural net.

For convenience we assume that each σ -gate G has an additional input x_n with some constant value $c \in (0, \gamma]$ available. Hence after rescaling r_n , the function f_G that is computed by G can be viewed as a restriction of the function

$$f_G(x_1, \dots, x_n) = \sigma \left(\sum_{i=1}^n r_i \cdot x_i \right)$$

to arguments with $x_n = c$. The original choice for the activation function σ in Rumelhart *et al.* (1986) has been the logistic sigmoid function $\sigma(y) = 1/(1 + e^{-y})$. Many years of practical experience with sigmoidal neural nets have shown that the exact form of the activation function σ is not relevant for the computational power and learning capabilities of such neural nets, as long as σ is nondecreasing and almost everywhere differentiable, the limits $\lim_{y \rightarrow -\infty} \sigma(y)$ and $\lim_{y \rightarrow \infty} \sigma(y)$ have finite values, and σ increases approximately linearly in some intermediate range. Gradient-descent learning procedures such as backpropagation formally require that σ is differentiable everywhere, but practically one can just as well use the piecewise linear "linear-saturated" activation function $\pi_\gamma: \mathbf{R} \rightarrow [0, \gamma]$ defined by

$$\pi_\gamma(y) = \begin{cases} 0, & \text{if } y < 0 \\ y, & \text{if } 0 \leq y \leq \gamma \\ \gamma, & \text{if } y > \gamma. \end{cases}$$

As a model for a spiking neuron we take the common model of a leaky integrate-and-fire neuron with noise, in the formulation of the somewhat more general spike response model of Gerstner and van Hemmen (1994). The only specific assumption needed for the construction in this article is that postsynaptic potentials can be described (or at least approximated) by a linear function during some initial segment. Actually the constructions of this article appear to be of interest even if this assumption is not satisfied, but in that case they are harder to analyze theoretically.

We consider networks that consist of a finite set V of spiking neurons, a set $E \subseteq V \times V$ of synapses, a weight $w_{u,v} \geq 0$ and a response function $\varepsilon_{u,v}: \mathbf{R}^+ \rightarrow \mathbf{R}$ for each synapse $\langle u, v \rangle \in E$ (where $\mathbf{R}^+ := \{x \in \mathbf{R}: x \geq 0\}$), and a threshold function $\Theta_v: \mathbf{R}^+ \rightarrow \mathbf{R}^+$ for each neuron $v \in V$.

Each response function $\varepsilon_{u,v}$ models either an EPSP or an IPSP. The typical shape of EPSPs and IPSPs is indicated in Figure 1.

If $F_u \subseteq \mathbf{R}^+$ is the set of firing times of a neuron u , then the potential at the trigger zone of neuron v at time t is given by

$$P_v(t) := \sum_{u: \langle u, v \rangle \in E} \sum_{s \in F_u: s < t} w_{u,v} \cdot \varepsilon_{u,v}(t - s).$$

Furthermore one considers a threshold function $\Theta_v(t - t')$ that quantifies the "reluctance" of v to fire again at time t if its last previous firing was at time t' . Thus $\Theta_v(x)$ is extremely large for small x and then approaches $\Theta_v(0)$ for larger x . In a noise-free model, a neuron v fires at time t as soon as $P_v(t)$ reaches $\Theta_v(t - t')$.

The precise form of this threshold function Θ_v is not important for the constructions in this article, since we consider here only computations that rely on the timing of the first spike in a spike train. Thus it suffices to assume that $\Theta_v(t - t') = \Theta_v(0)$ for sufficiently large values of $t - t'$ and

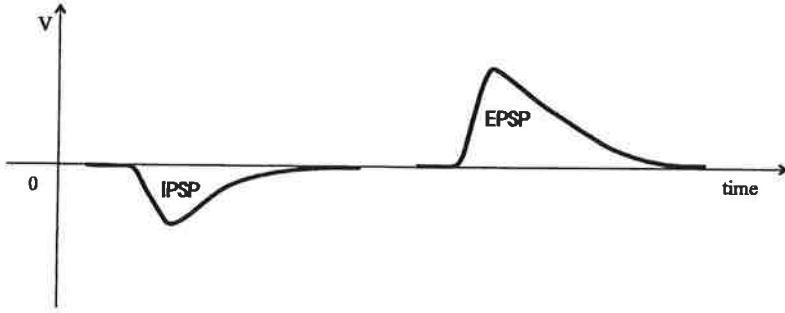


Figure 1: The typical shape of inhibitory and excitatory postsynaptic potentials at a biological neuron. (We assume that the resting membrane potential has the value 0.)

that $\inf\{\Theta_v(x): x \in (0, \gamma]\}$ is larger than the potentials $P_v(t)$ that occur in the construction of Section 2 for $t \in [T_{out} - \gamma, T_{out}]$. The latter condition (which amounts to the assumption of a sufficiently long refractory period) will prevent iterated firing of neuron v during the critical time interval $[T_{out} - \gamma, T_{out}]$.

The construction in Section 2 is robust with respect to several types of noise that make the model of a spiking neuron biologically more realistic. As in the model for a leaky integrate-and-fire neuron with noise, we allow that the potentials $P_v(t)$ and the threshold functions $\Theta_v(t - t')$ are subject to some additive noise. Hence $P_v(t)$ is replaced by

$$P_v^{noisy}(t) := P_v(t) + \alpha_v(t),$$

and $\Theta_v(t - t')$ is replaced by

$$\Theta_v^{noisy}(t - t') := \Theta_v(t - t') + \beta_v(t),$$

where $\alpha_v(t)$ and $\beta_v(t)$ describe the impact of some unknown (or even adversarial) source of noise (which might, for example, result from synapse failures).

One assumes in most previous theoretical studies that $\alpha_v(t)$, $\beta_v(t)$ are distributed according to some specific probability distribution (e.g., white noise), whereas our subsequent constructions allow that $\alpha_v(t)$, $\beta_v(t)$ are some arbitrary functions with bounded absolute value (e.g., “systematic noise”).

In a simpler model for a noisy spiking neuron, one assumes that a neuron v fires exactly at those time points t when $P_v^{noisy}(t)$ reaches from below the value $\Theta_v^{noisy}(t - t')$. We consider in this article a biologically more realistic

model, whereas in Gerstner and van Hemmen (1994), the size of the difference $P_v^{noisy}(t) - \Theta_v^{noisy}(t - t')$ governs just the probability that neuron v fires. The choice of the exact firing times is left up to some unknown stochastic processes, and it may, for example, occur that v does not fire in a time interval I during which $P_v^{noisy}(t) - \Theta_v^{noisy}(t - t') > 0$, or that v fires "spontaneously" at a time t when $P_v^{noisy}(t) - \Theta_v^{noisy}(t - t') < 0$.

For the subsequent constructions we need only the following assumption about the firing mechanism: For any time interval I of length greater than 0, the probability that v fires during I is arbitrarily close to 1 if $P_v^{noisy}(t) - \Theta_v^{noisy}(t - t')$ is sufficiently large for $t \in I$ (up to the time when v fires), and the probability that v fires during I is arbitrarily close to 0 if $P_v^{noisy}(t) - \Theta_v^{noisy}(t - t')$ is sufficiently negative for all $t \in I$.

It turns out that it suffices to assume only the following rather weak properties of the response functions $\varepsilon_{u,v}$: Each response function $\varepsilon_{u,v}: \mathbf{R}^+ \rightarrow \mathbf{R}$ is either excitatory or inhibitory. All excitatory response functions $\varepsilon_{u,v}(t)$ have the value 0 for $t \in [0, d_{u,v}]$, and the value $t - d_{u,v}$ for $t \in [d_{u,v}, d_{u,v} + \Delta]$, where $d_{u,v} \geq 0$ is some fixed delay and $\Delta > 0$ is some other constant. Furthermore we assume that $\varepsilon_{u,v}(t) \geq \varepsilon_{u,v}(d_{u,v} + \Delta)$ for all $t \in [d_{u,v} + \Delta, d_{u,v} + \Delta + \gamma]$, where γ with $0 < \gamma \leq \Delta/2$ is another constant. With regard to inhibitory response functions $\varepsilon_{u,v}(t)$, we assume that $\varepsilon_{u,v}(t) = 0$ for $t \in [0, d_{u,v}]$ and $\varepsilon_{u,v}(t) = -(t - d_{u,v})$ for $t \in [d_{u,v}, d_{u,v} + \Delta]$. Furthermore we assume that $\varepsilon_{u,v}(t) = 0$ for all sufficiently large t .

Finally we need a mechanism for increasing the firing threshold $\Theta := \Theta_v(0)$ of a "rested" neuron v (at least for a short period). One biologically plausible assumption that would account for such an increase is that neuron v receives a large number of IPSPs from randomly firing neurons that arrive on synapses that are far away from the trigger zone of v , so that each of them has barely any effect on the dynamics of the potential at the trigger zone, but together they contribute a rather steady negative summand BN^- to the potential at the trigger zone. Other possible explanations for the increase of the firing threshold Θ could be based on the contribution of inhibitory interneurons whose IPSPs arrive close to the soma and are time locked to the onset of the stimulus, or on long-lasting inhibitions such as those mediated by $GABA_B$ receptors.

Formally we assume that each neuron v receives some negative (i.e., inhibitory) potential $BN^- < 0$ that can be assumed to be constant during the time intervals that are considered in the following arguments.

In comparison with other models for spiking neurons, this model allows more general noise than the models considered in Gerstner and van Hemmen (1994) and Maass (1995). On the other hand this model is somewhat less general than the one considered in Maass (1996a).

Having defined the formal model, we can now explain the key mechanism of the constructions in more detail. It is well known that incoming

EPSPs and IPSPs are able to *shift* the firing time of a biological neuron. We explore this effect in the mathematical model of a spiking neuron, showing that in principle it can be used to carry out complex analog computations in temporal coding. Assume that a spiking neuron v receives PSPs from presynaptic neurons a_1, \dots, a_n , that w_i is the weight (efficacy) for the synapse from a_i to v , and that d_i is the time delay from a_i to v . Then there exists a range of values for the parameters where the firing time t_v of neuron v can be written in terms of the firing times t_{a_i} of the presynaptic neurons a_i as

$$t_v = \frac{\Theta - BN^-}{\sum_{i=1}^n w_i} + \frac{\sum_{i=1}^n w_i \cdot (t_{a_i} + d_i)}{\sum_{i=1}^n w_i}. \quad (1.1)$$

Hence in principle a spiking neuron is able to compute in *temporal coding* of inputs and outputs a *linear* function (where the efficacies of synapses encode the coefficients of the linear function, as in rate coding of analog variables). The calculations at the beginning of Section 2 show that this holds precisely if there is no noise and the n PSPs are at time t_v all in their initial linearly rising or linearly decreasing phase. However, for a biological interpretation, it is interesting to know that even if the firing times t_{a_i} (or more precisely their effective values $t_{a_i} + d_i$) lie further apart, this mechanism computes a meaningful approximation to a linear function. It employs (through the natural shape of PSPs) an interesting adaptation of outliers among the $t_{a_i} + d_i$: Input neurons a_i that fire too late (relative to the average) lose their influence on the determination of t_v , and input neurons a_i that fire extremely early have the same impact as neurons a_i that fire somewhat later (but still before the average). Remark 2 in Section 2 provides a more detailed discussion of this effect.

The goal of the next section is to prove a rigorous theoretical result about the computational power of formal models for networks of spiking neurons. We are not claiming that this construction (which is designed exclusively for that purpose) provides a blueprint for the organization of fast analog computations in biological neural systems. However, it provides the first theoretical model that is able to explain the possibility of fast analog computations with noisy spiking neurons. Some remarks about the possible biological relevance of details of this construction can be found after the theorem in Section 2.

2 The Main Construction

Consider an arbitrary π_γ -gate G , for some $\gamma > 0$, which computes a function $f_G: [0, \gamma]^n \rightarrow [0, \gamma]$. Let $r_1, \dots, r_n \in \mathbf{R}$ be the weights of G . Thus we have

$$f_G(s_1, \dots, s_n) = \begin{cases} 0, & \text{if } \sum_{i=1}^n r_i \cdot s_i < 0 \\ \sum_{i=1}^n r_i \cdot s_i, & \text{if } 0 \leq \sum_{i=1}^n r_i \cdot s_i \leq \gamma \\ \gamma, & \text{if } \sum_{i=1}^n r_i \cdot s_i > \gamma \end{cases}$$

for arbitrary inputs $s_1, \dots, s_n \in [0, \gamma]$.

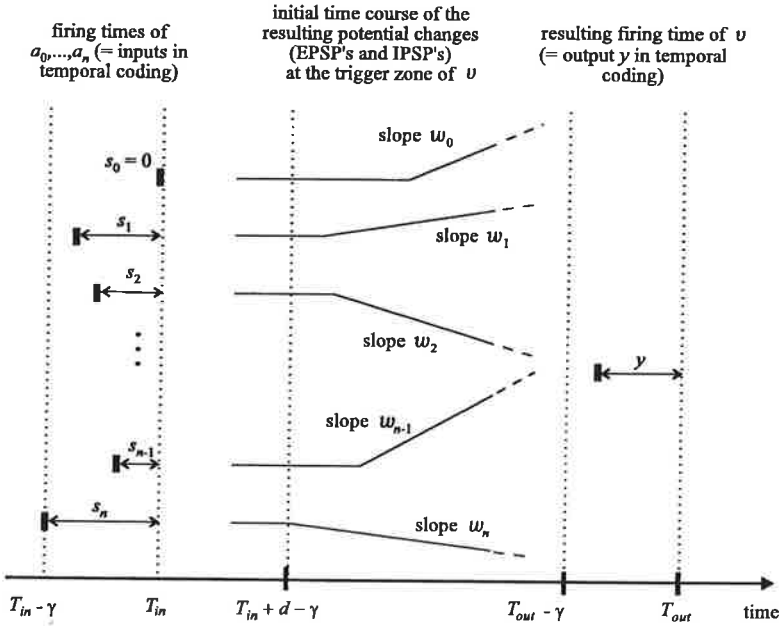


Figure 2: The simulation of a sigmoidal gate by a spiking neuron v in temporal coding

For the sake of simplicity we first consider the case of spiking neurons without noise (i.e., $\alpha_v = \beta_v \equiv 0$ and each neuron v fires whenever $P_v(t)$ crosses $\Theta_v(t - t')$ from below). Then we describe the changes that are needed in this construction for the general case of noisy spiking neurons.

We construct for a given π_γ -gate G and for an arbitrary given parameter $\varepsilon > 0$ with $\varepsilon < \gamma$ a network $\mathcal{N}_{G,\varepsilon}$ of spiking neurons that approximates f_G with precision $\leq \varepsilon$; that is, the output $\mathcal{N}_{G,\varepsilon}(s_1, \dots, s_n)$ of $\mathcal{N}_{G,\varepsilon}$ satisfies $|\mathcal{N}_{G,\varepsilon}(s_1, \dots, s_n) - f_G(s_1, \dots, s_n)| \leq \varepsilon$ for all $s_1, \dots, s_n \in [0, \gamma]$.

In order to be able to scale the size of weights according to the given gate G , we assume that $\mathcal{N}_{G,\varepsilon}$ receives an additional input s_0 that is given like the other input variables s_1, \dots, s_n in temporal coding. Thus we assume that there are $n + 1$ input neurons a_0, \dots, a_n with the property that a_i fires at time $T_{in} - s_i$ (where T_{in} is some constant). We will discuss at the end of this section (in Remarks 5 and 6) biologically more plausible variations of the construction where a_0 and T_{in} are not needed.

We construct a spiking neuron v in $\mathcal{N}_{G,\varepsilon}$ that receives $n + 1$ PSPs $h_0(t), \dots, h_n(t)$ from the $n + 1$ input neurons a_0, \dots, a_n , which result from the firing of a_i at time $T_{in} - s_i$ (see Figure 2). In addition v receives some auxiliary PSPs from other spiking neurons in $\mathcal{N}_{G,\varepsilon}$, whose timing depends only on T_{in} .

The firing time t_v of this neuron v will provide the output $\mathcal{N}_{G,\varepsilon}(s_1, \dots, s_n)$ of the network $\mathcal{N}_{G,\varepsilon}$ in temporal coding; that is, v will fire at time $T_{out} - \mathcal{N}_{G,\varepsilon}(s_1, \dots, s_n)$ for some T_{out} that does not depend on s_1, \dots, s_n . Let $w_{a_i,v}$ be the weight of the synapse from input neuron a_i to neuron v , $i = 0, \dots, n$. We assume that the “delay” $d_{a_i,v}$ between a_i and v is the same for all input neurons a_0, \dots, a_n , and we write d for this common delay. Thus we can describe for $i = 0, \dots, n$ the impact of the firing of a_i at time $T_{in} - s_i$ on the potential at the trigger zone of neuron v at time t by the EPSP or IPSP $h_i(t) = w_{a_i,v} \cdot \varepsilon_{a_i,v}(t - (T_{in} - s_i))$, which has on the basis of our assumptions the value

$$h_i(t) = \begin{cases} 0, & \text{if } t - (T_{in} - s_i) < d \\ w_i \cdot (t - (T_{in} - s_i) - d), & \text{if } d \leq t - (T_{in} - s_i) \leq d + \Delta, \end{cases}$$

where $w_i = w_{a_i,v}$ in the case of an EPSP and $w_i = -w_{a_i,v}$ in the case of an IPSP.

We assume that neuron v has not fired for a sufficiently long time, so that its threshold function $\Theta_v(t - t')$ can be assumed to have a constant value Θ when the PSPs $h_0(t), \dots, h_n(t)$ arrive at the trigger zone of v . Furthermore we assume for the moment that besides these $n+1$ PSPs only BN^- influences the potential at the trigger zone of v . This contribution BN^- is assumed to have a constant value in the time interval considered here. Then if no noise is present, the time t_v of the next firing of v can be described by the equality

$$\Theta = \sum_{i=0}^n h_i(t_v) + BN^- = \sum_{i=0}^n w_i \cdot (t_v - (T_{in} - s_i) - d) + BN^-, \quad (2.1)$$

provided that

$$d \leq t_v - (T_{in} - s_i) \leq d + \Delta \text{ for } i = 0, \dots, n. \quad (2.2)$$

We assume from now on that a fixed value $s_0 \equiv 0$ is chosen for the extra input s_0 . Then equation 2.1 is equivalent to

$$t_v = \frac{\Theta - BN^-}{\sum_{i=0}^n w_i} + T_{in} + d - \frac{\sum_{i=1}^n w_i \cdot s_i}{\sum_{i=0}^n w_i}. \quad (2.3)$$

This t_v satisfies equation 2.2 if $-s_j \leq t_v - T_{in} - d \leq \Delta - s_j$ for $j = 0, \dots, n$; hence for any $s_j \in [0, \gamma]$ if

$$0 \leq \frac{\Theta - BN^- - \sum_{i=1}^n w_i \cdot s_i}{\sum_{i=0}^n w_i} \leq \Delta - \gamma. \quad (2.4)$$

We set $w_i := \lambda \cdot r_i$ for $i = 1, \dots, n$, where r_1, \dots, r_n are the weights of the simulated π_γ -gate G , and $\lambda > 0$ is some not-yet-determined factor (that we

will later choose sufficiently large in order to make sure that neuron v fires closely to the time t_v given by equation 2.3 even in the presence of noise). We choose w_0 so that $\sum_{i=0}^n w_i = \lambda$. This implies that equation 2.3 with

$$T_{out} := \frac{\Theta - BN^-}{\lambda} + T_{in} + d$$

is now equivalent to

$$t_v = T_{out} - \sum_{i=1}^n r_i \cdot s_i, \quad (2.5)$$

and equation 2.4 is equivalent to

$$0 \leq \frac{\Theta - BN^-}{\lambda} - \sum_{i=1}^n r_i \cdot s_i \leq \Delta - \gamma. \quad (2.6)$$

Hence, provided that γ , λ , and BN^- are chosen in relationship to Θ and Δ so that

$$\gamma \leq \frac{\Theta - BN^-}{\lambda} \leq \Delta - \gamma, \quad (2.7)$$

we have satisfied equation 2.4, and therefore achieved that the firing time t_v of neuron v provides in temporal coding the output $f_G(s_1, \dots, s_n) = \sum_{i=1}^n r_i \cdot s_i$ of the simulated π_γ -gate G for all inputs $s_1, \dots, s_n \in [0, \gamma]$ with $\sum_{i=1}^n r_i \cdot s_i \in [0, \gamma]$.

In order to simulate G also for $s_1, \dots, s_n \in [0, \gamma]$ with other values of $\sum_{i=1}^n r_i \cdot s_i$, we make sure that v fires at a time with distance at most ε to $T_{out} - \gamma$ if $\sum_{i=1}^n r_i \cdot s_i$ is larger than γ and that v fires with distance at most ε to time T_{out} if $\sum_{i=1}^n r_i \cdot s_i < 0$. For that purpose, we add inhibitory and excitatory neurons that fire at times that depend on time T_{in} but not on the input s_1, \dots, s_n . The activity of these auxiliary inhibitory and excitatory neurons may also shift the firing time t_v of v if $\sum_{i=1}^n r_i \cdot s_i \in [0, \gamma]$, but at most by ε .

According to the previously described construction, we have by equation 2.5 that $t_v = T_{out}$ if $s_i = 0$ for $i = 1, \dots, n$ (and therefore $\sum_{i=1}^n r_i \cdot s_i = 0$). Furthermore the parameters have been chosen so that equation 2.2 is satisfied for this case, which implies that each of the PSPs $h_i(t)$ is at time T_{out} for $s_i = 0$ still within the initial segment of length Δ of its first nonzero segment. This implies that for *any* value of $s_i \in [0, \gamma]$ the PSP $h_i(t)$ is at time $T_{out} - \gamma$ not further advanced than the end of the initial segment of length Δ of its first nonzero segment. Thus

$$\begin{aligned} |h_i(t)| &= |w_i \cdot (t - (T_{in} - s_i) - d)| \\ &\leq |w_i| \cdot (T_{out} - \gamma - (T_{in} - \gamma) - d) \leq |w_i| \cdot \Delta \end{aligned}$$

for all $t \leq T_{out} - \gamma$ and all $s_i \in [0, \gamma]$. This implies that for any $s_1, \dots, s_n \in [0, \gamma]$ (even if $\sum_{i=1}^n r_i \cdot s_i \notin [0, \gamma]$) we have $|\sum_{i=0}^n h_i(t)| \leq W \cdot \Delta$ for all $t \leq T_{out} - \gamma$, where $W := \sum_{i=0}^n |w_i|$.

Hence in order to prevent a firing of v before time $T_{out} - \gamma$, it suffices to employ auxiliary neurons in $\mathcal{N}_{G,\varepsilon}$ that send IPSPs to v that lower the potential at the trigger zone of v during the interval $[T_{out} - \gamma - \Delta, T_{out} - \gamma]$ by more than $W \cdot \Delta + BN^- - \Theta$. In general, these IPSPs also influence the potential at the trigger zone of v shortly after time $T_{out} - \gamma$. Since these auxiliary IPSPs will be independent of the input s_1, \dots, s_n , they may delay the firing of v even when $\sum_{i=1}^n r_i \cdot s_i \in [0, \gamma]$. In order to approximate the given π_γ -gate G with an error of at most ε , we assume that these auxiliary IPSPs have vanished at the trigger zone of v by time $T_{out} - \gamma + \varepsilon$. It is obvious that all these conditions can be achieved with a sufficient number (and/or weights) of auxiliary synchronized IPSPs from inhibitory neurons in $\mathcal{N}_{G,\varepsilon}$ whose firing time depends only on T_{in} . This can be satisfied using only the very weak condition that each IPSP is continuous and of value $\neq 0$ before it eventually vanishes (see the precise condition in Section 1).

We now want to make sure that v fires at the latest by time $T_{out} - \gamma + \varepsilon$ if $\sum_{i=1}^n r_i \cdot s_i > \gamma - \varepsilon$. Since all auxiliary IPSPs have vanished by time $T_{out} - \gamma + \varepsilon$, we have $P_v(T_{out} - \gamma + \varepsilon) = \sum_{i=0}^n h_i(T_{out} - \gamma + \varepsilon) + BN^-$. Hence it suffices to show that the latter is $\geq \Theta$. Consider the set I of those $i \in \{1, \dots, n\}$ with $r_i > 0$, that is, those i where $w_{a_i,v} \cdot \varepsilon_{a_i,v}(t)$ represents an EPSP. By choosing suitable values s'_i in the interval $[0, s_i]$ for $i \in I$ and by setting $s'_i := s_i$ for $i \in \{1, \dots, n\} - I$, one can achieve that $\sum_{i=1}^n r_i \cdot s'_i = \gamma - \varepsilon$. According to equation 2.5, the potential P_v reaches the value Θ at time $T_{out} - \gamma + \varepsilon$ for the input $\langle s'_1, \dots, s'_n \rangle$, and according to equations 2.2, 2.4, 2.6, and 2.7, each PSP $w_{a_i,v} \cdot \varepsilon_{a_i,v}$ is at time $T_{out} - \gamma + \varepsilon$ still within Δ of the beginning of its nonzero phase. If we now change s'_i to s_i for $i \in I$, the EPSP $w_{a_i,v} \cdot \varepsilon_{a_i,v}$ will be advanced in time by $s_i - s'_i \in [0, \gamma]$. Hence each of the EPSPs $w_{a_i,v} \cdot \varepsilon_{a_i,v}$ for $i \in I$ is for input s_i at time $T_{out} - \gamma + \varepsilon$ within $\Delta + \gamma$ of the beginning of its rising phase. Since we have assumed in Section 1 that $w_{a_i,v} \cdot \varepsilon_{a_i,v}(t) \geq w_{a_i,v} \cdot \varepsilon_{a_i,v}(d + \Delta)$ for all $t \in [d + \Delta, d + \Delta + \gamma]$, $P_v(T_{out} - \gamma + \varepsilon)$ has for input $\langle s_1, \dots, s_n \rangle$ a value that is at least as large as for input $\langle s'_1, \dots, s'_n \rangle$, and therefore a value $\geq \Theta$. This implies that in the case $\sum_{i=1}^n r_i \cdot s_i > \gamma - \varepsilon$ the neuron v will fire within the time interval $[T_{out} - \gamma, T_{out} - \gamma + \varepsilon]$.

In an analogous manner we can achieve with the help of EPSPs from auxiliary excitatory neurons in $\mathcal{N}_{G,\varepsilon}$ (whose firing time depends only on T_{in} , not on the input values s_1, \dots, s_n) that neuron v fires at the latest at time T_{out} , even if $\sum_{i=1}^n r_i \cdot s_i < 0$. The preceding analysis implies that for any value $s_i \in [0, \gamma]$ and any $t \leq T_{out}$, the absolute value of the PSP $w_{a_i,v} \cdot \varepsilon_{a_i,v}(t)$ can be bounded by $|w_{a_i,v}| \cdot \rho$, where $\rho := \sup\{|\varepsilon_{a_i,v}(t)| : i \in \{0, \dots, n\} \text{ and } t \in [d, d + \Delta + \gamma]\}$ is the maximum absolute value that any $\varepsilon_{a_i,v}(t)$ for $i = 0, \dots, n$ can reach during the initial segment of length $\Delta + \gamma$ of its nonzero segment. Thus if the absolute value of these functions $\varepsilon_{a_i,v}(t)$ grows during $[d + \Delta, d + \Delta + \gamma]$ not faster than during their linear segment $[d, d + \Delta]$ (which

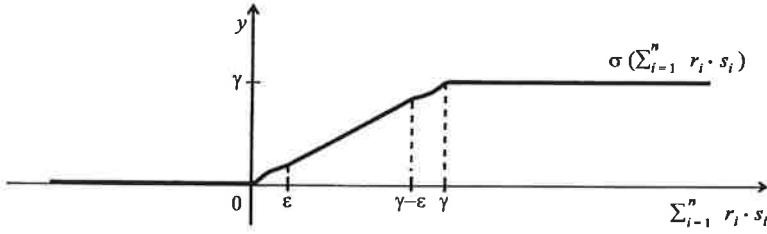


Figure 3: The resulting activation function σ of the implementation of a sigmoidal gate in temporal coding.

apparently holds for biological PSPs), we can set $\rho := \Delta + \gamma$. Consequently we can derive for any $t \leq T_{out}$ and any values $s_1, \dots, s_n \in [0, \gamma]$ the bound $|\sum_{i=0}^n h_i(t)| \leq \sum_{i=0}^n |w_{a_i, v}| \cdot \rho = W \cdot \rho$. Thus it suffices to make sure that EPSPs from auxiliary neurons in $\mathcal{N}_{G, \varepsilon}$ reach the trigger zone of neuron v shortly after time $T_{out} - \varepsilon$ and that their sum reaches a value $\Theta - BN^- + W \cdot \rho$ by time T_{out} . Then for any values of $s_1, \dots, s_n \in [0, \gamma]$ the potential $P_v(t)$ will reach the value Θ at the latest by time T_{out} , causing a firing of v by time T_{out} . These auxiliary EPSPs will in general have the side effect that they slightly advance for inputs $s_1, \dots, s_n \in [0, \gamma]$ with $\sum_{i=1}^n r_i \cdot s_i \in [0, \varepsilon]$ the firing time t_v , but the firing time will stay within the interval $[T_{out} - \varepsilon, T_{out}]$.

According to our assumption in Section 1, $\inf\{\Theta_v(x) : x \in (0, \gamma)\}$ is so large that v cannot fire more than once during $[T_{out} - \gamma, T_{out}]$. Hence our construction makes sure that v fires exactly once during $[T_{out} - \gamma, T_{out}]$ for any inputs $s_1, \dots, s_n \in [0, \gamma]$. In the following discussion we will denote this firing time t_v of v by $T_{out} - \mathcal{N}_{G, \varepsilon}(s_1, \dots, s_n)$.

We have now shown that under the assumption that no noise influences the firing time t_v of neuron v , we have $|\mathcal{N}_{G, \varepsilon}(s_1, \dots, s_n) - f_G(s_1, \dots, s_n)| \leq \varepsilon$ for all $s_1, \dots, s_n \in [0, \gamma]$. If one plots the variable $y = \mathcal{N}_{G, \varepsilon}(s_1, \dots, s_n)$ that neuron v outputs in temporal coding as a function of $\sum_{i=1}^n r_i \cdot s_i$, one sees that the effective “activation function” σ of this implementation of the π_γ -gate G has the form indicated in Figure 3. Like the piecewise linear activation function π_γ of the simulated gate G , this activation function σ is linear in the intervals $(-\infty, 0]$, $[\varepsilon, \gamma - \varepsilon]$, $[\gamma, \infty)$. However, in contrast to π_γ , it connects these linear segments by smooth segments in the intervals $[0, \varepsilon]$ and $[\gamma - \varepsilon, \gamma]$, whose precise shape depends on the exact shape and size of the previously discussed auxiliary EPSPs and IPSPs.

We now consider the general case where the firing time t_v of v is subject to two types of noise, as described in Section 1. We show that by choosing a sufficiently large parameter λ in the preceding construction, we can achieve that in spite of these two types of noise the actual firing time \hat{t}_v of neuron

v will be arbitrarily close to the previously considered deterministic firing time t_v of neuron v , with probability arbitrarily close to 1.

Assume that some arbitrary $\varepsilon, \delta > 0$ are given. In the preceding construction, we had chosen values $w_i := \lambda \cdot r_i$ for the strengths of the synapses to neuron v , where $\lambda > 0$ was some not-yet-determined factor. According to equation 2.3, a change in the value of this parameter λ can result only in a shift of t_v by an amount that is independent from the input variables s_1, \dots, s_n . Furthermore if one chooses the contribution BN^- of inhibitory background noise as a function of λ so that $\Theta - BN^- = \lambda \cdot \tilde{c}$ for some constant \tilde{c} , the resulting firing time t_v is completely independent of the choice of λ . On the other hand, the parameter λ occurs as a factor in the non-constant term $\sum_{i=0}^n \lambda \cdot r_i \cdot |\varepsilon_{a_i, v}(t - (T_{in} - s_i))| = \sum_{i=0}^n h_i(t)$ in the potential $P_v(t) = \sum_{i=0}^n h_i(t) + BN^-$ (we ignore the effect of auxiliary PSPs for the moment). Hence by choosing λ sufficiently large, one can make sure that in the deterministic case $P_v(t)$ has an arbitrarily large derivative at the time t_v when it crosses $\Theta_v(t - t')$.

Furthermore if we choose γ, λ , and BN^- so that equation 2.7 can be replaced by

$$2\gamma \leq \frac{\Theta - BN^-}{\lambda} \leq \Delta - \gamma - \varepsilon, \quad (2.8)$$

then it is guaranteed that the linearly increasing potential $P_v(t)$ (with slope proportional to λ) will rise with the same slope throughout the interval $[t_v - \gamma, t_v + \varepsilon]$.

If we now keep this setting of the parameters but replace the deterministic neuron v by a stochastic neuron v that is subject to the two types of noise that were specified in Section 1, the following can be observed: If $|\alpha(t)| \leq \alpha$ and $|\beta(t)| \leq \beta$ for all t , then the time interval around t_v during which $P_v(t)$ is within the interval $[\Theta - \alpha - \beta, \Theta + \alpha + \beta]$ becomes arbitrarily small for sufficiently large λ .

Furthermore if λ is sufficiently large (and BN^- is adjusted along with λ so that $\frac{\Theta - BN^-}{\lambda} = \tilde{c}$ for some constant $\tilde{c} > 0$ that is independent of λ), then $P_v(t) + \alpha + \beta - \Theta$ is arbitrarily negative during $[T_{out} - \gamma, t_v - \varepsilon]$. Thus the probability that v fires during $[T_{out} - \gamma, t_v - \varepsilon]$ can be brought arbitrarily close to 0. In addition, by making λ sufficiently large, one can achieve that $P_v(t) - \alpha - \beta - \Theta$ is arbitrarily large throughout the time interval $[t_v + \varepsilon/2, t_v + \varepsilon]$. Hence the probability that v fires by time $t_v + \varepsilon$ can be brought arbitrarily close to 1.

If one increases the number or the weights of the previously described auxiliary PSPs in an analogous manner, one can achieve for the noisy neuron model that with probability $\geq 1 - \delta$ the neuron v of network $\mathcal{N}_{G, \varepsilon, \delta}$ fires exactly once during $[T_{out} - \gamma, T_{out}]$, and at a time $\tilde{t}_v = T_{out} - \mathcal{N}_{G, \varepsilon, \delta}(s_1, \dots, a_n)$ with $|\mathcal{N}_{G, \varepsilon, \delta}(s_1, \dots, a_n) - f_G(s_1, \dots, s_n)| \leq 2\varepsilon$, for all $s_1, \dots, s_n \in [0, \gamma]$.

The previously described construction can easily be adapted to allow shifts in the arrival times of auxiliary PSPs at v of up to $\varepsilon/2$. Hence we can allow that these PSPs also come from noisy spiking neurons.

In order to simulate an arbitrary given feedforward network N of π_γ -gates with precision ε and probability $\geq 1 - \delta$ of correctness, one applies the preceding construction separately to each gate G in N . For any given $\varepsilon, \delta > 0$ one determines for each π_γ -gate G in N (starting at the output gates of N) suitable values $\varepsilon_G, \delta_G > 0$, so that to approximate N within ε with probability $\geq 1 - \delta$, it suffices that each gate G in N is approximated within $2\varepsilon_G$ with probability $\geq 1 - \delta_G$ by a network $\mathcal{N}_{G, \varepsilon_G, \delta_G}$ of noisy spiking neurons. In this way, one can achieve that the network $\mathcal{N}_{N, \varepsilon, \delta}$ of noisy spiking neurons that is composed of these networks $\mathcal{N}_{G, \varepsilon_G, \delta_G}$ approximates in temporal coding with probability $\geq 1 - \delta$ the output of N within ε , for any given network input $x_1, \dots, x_m \in [0, \gamma]$. Actually it would be more efficient if the modules $\mathcal{N}_{G, \varepsilon_G, \delta_G}$ are not disjoint but share the neurons that generate the auxiliary EPSPs and IPSPs. Note that the problem of achieving reliable digital and analog computation with noisy neurons has already been considered in von Neumann (1956) for other types of formal neuron models.

Thus we have shown:

Theorem. *For any given $\varepsilon, \delta > 0$ one can simulate any given feedforward sigmoidal neural net N consisting of π_γ -gates (for some sufficiently small γ that depends on the chosen model for a spiking neuron) by a network $\mathcal{N}_{N, \varepsilon, \delta}$ of noisy spiking neurons in temporal coding. More precisely, for any network input $x_1, \dots, x_m \in [0, \gamma]$ the output of $\mathcal{N}_{N, \varepsilon, \delta}$ differs with probability $\geq 1 - \delta$ by at most ε from that of N . Furthermore the computation time of $\mathcal{N}_{N, \varepsilon, \delta}$ depends on neither the number of gates in N nor the parameters ε, δ , but only on the number of layers of the sigmoidal neural network N .*

Remark 1. One can exploit the concrete shape of PSPs in biological neurons in order to arrive at an alternative approximation of sigmoidal neural nets by spiking neurons in temporal coding *without* simulating explicitly the activation function π_γ of the sigmoidal neural net. In this alternative approach, one can *delete* from the previously described construction the auxiliary neurons that prevent a firing of neuron v before time $T_{out} - \gamma$ and force it to fire at the latest by time T_{out} . Furthermore in this interpretation we no longer have to assume that the initial linear segments of all incoming PSPs overlap, and hence less precision in the firing times is needed.

In case v fires substantially after time T_{out} , the resulting PSP at a postsynaptic neuron v' (which simulates a sigmoidal gate on the next layer of the simulated sigmoidal neural net) still has its initial value 0 at the time $t_{v'}$ when v' fires. Dually, if v fires before time $T_{out} - \gamma$, the resulting PSP may be at time $t_{v'}$ already near its saturation value (where it increases or decreases more slowly than during its initial “linear phase”). Thus in either case, the concrete functional form of EPSPs and IPSPs in a biological neuron v' mod-

ulates the input from a presynaptic neuron v in a way that corresponds to the application of a saturating sigmoidal activation function to the output of v in temporal coding. Furthermore, larger differences than γ between the firing times of presynaptic neurons v can be tolerated in this context.

This implicit implementation of an activation function is, however, mathematically imprecise. A closer look shows that the amount by which the output of v in temporal coding is adjusted depends on the size of the output of v relative to the size of the outputs of the other presynaptic neurons of v' (in temporal coding). The output of v is changed by a larger amount if it differs more strongly from the median output of the other presynaptic neurons of v' (but it is always moved in the direction of the median). This context-dependent modulation of the output of v is hard to exploit for precise theoretical results, but it appears to be useful for practically relevant computations such as pattern recognition.

Thus we arrive in this way at a variation of the implementation of a sigmoidal neural net, where two biologically dubious components of our main construction (the auxiliary neurons that force v to fire exactly once during $[T_{out} - \gamma, T_{out}]$, as well as the requirement that the initial linear segments of all relevant PSPs have to overlap) are deleted, but the resulting network of spiking neurons can still carry out complex and practical multilayer computations in temporal coding.

Remark 2. Our construction shows as a special case that a linear function of the form $\underline{s} \mapsto \underline{w} \cdot \underline{s}$ for real-valued inputs $\underline{s} = \langle s_1, \dots, s_n \rangle$ and a stored vector $\underline{w} = \langle w_1, \dots, w_n \rangle$ can be computed very efficiently (and very quickly) by a spiking neuron v in temporal coding. In this case, no auxiliary neurons are needed.

Quick computation of linear functions is relevant in many biological contexts, such as coordinate transformations between different frames of reference or the analysis of a complex stimulus \underline{s} in terms of many stored patterns $\underline{w}, \underline{w}', \dots$. For example, in an olfactory neural system (see, e.g., Hopfield 1991, 1995) the stimulus \underline{s} may be thought of as a superposition of various stored basic odors $\underline{w}, \underline{w}', \dots$. In this case the output $y = \underline{w} \cdot \underline{s}$ of neuron v in temporal coding may be interpreted as the amount by which the basic odor \underline{w} (which is stored in the efficacies of the synapses of v) is present in the stimulus \underline{s} . Furthermore another neuron \tilde{v} on the next layer might receive as its input $\underline{y} = \langle y, y', \dots \rangle$ from several such neurons v, v', \dots ; that is, \tilde{v} receives the “mixing proportions” $y = \underline{w} \cdot \underline{s}, y' = \underline{w}' \cdot \underline{s}$ for various stored basic odors $\underline{w}, \underline{w}', \dots$, in temporal coding. This neuron \tilde{v} on the second layer can then continue the pattern analysis by computing for its input \underline{y} the inner product $\underline{W} \cdot \underline{y}$ with some stored “higher-order pattern” \underline{W} (e.g., the composition of basic odors characteristic for an individual animal). Such multilayer pattern analysis is facilitated by the fact that the neurons considered here encode their output in the same way in which their input is

encoded (in contrast to the approach in Hopfield 1995). One also gets in this way a very fast implementation of Linsker's network (Linsker 1988) with spiking neurons in temporal coding.

Remark 3. For biological neurons it is impossible to choose the parameter λ arbitrarily large. On the other hand, the experiments of Bryant and Segundo (1976), as well as the recent experiments of Mainen and Sejnowski (1995), suggest that biological neurons already exhibit very high precision in their firing times if the slope of the membrane potential $P_v(t)$ at the time t when it crosses the firing threshold is moderately large.

Remark 4. Analog computations with the type of temporal coding considered here become impossible if the jitter in the firing times is large relative to the size of the range $[0, \gamma]$ of analog variables in temporal coding. The following points should be taken into account in this context:

- Even if the jitter is so high that in temporal coding just two different outputs can be distinguished in a reliable manner, the computational power of the constructed network \mathcal{N} of spiking neurons is still enormous from the point of view of computational complexity theory. In this case, the network can simulate arbitrary threshold circuits (i.e., multilayer perceptrons whose gates give binary outputs) very fast. Threshold circuits are extremely powerful models for parallel digital computation, which can (in contrast to PRAMs and other models for currently available computer hardware) compute various nontrivial boolean functions that depend on a large number n of input bits with polynomially in n many gates and not more than four layers (Johnson 1990; Roychowdhury *et al.* 1994; Siu *et al.* 1995).
- Formally the value of γ in the preceding construction was required to be very small: it was bounded by a fraction of the length Δ of the rising segment of an EPSP (see inequalities 2.7 and 2.8). However, keep in mind that we were forced to resort to such small values for γ only because we wanted to prove a rigorous theoretical result. The considerations in Remark 1 suggest that in a *practical* context, one can still carry out meaningful computations if the linearly rising segments of incoming EPSP are spread out over a somewhat larger time interval than Δ , where they need no longer overlap. If one wants to compute specific values for the parameters γ and Δ , one runs into the problem that the value of Δ varies enormously among different biological neural systems, from about 1 to 3 ms for EPSPs resulting from AMPA receptors in cortex to about a second in dark-adapted toad photoreceptors.
- There exists an alternative interpretation of our construction in a biological neural system that focuses on a smaller spatial scale. In this interpretation, a "hot spot" in the dendritic tree of a biological neuron

assumes the role of a spiking neuron in our construction. Hot spots are patches of membrane with voltage-dependent channels that are known to occur at branching points of the dendritic tree (Shepherd 1994; Jaffe *et al.* 1992; Mel 1993; Softky 1994). They fire a dendritic spike if the membrane potential reaches a certain threshold value. From this point of view, a single biological neuron may be viewed as a network of spiking neurons, which according to our construction, can simulate in temporal coding a multilayer sigmoidal neural net. The timing precision of such circuits is possibly very high since it does not involve synapses (see also Softky 1994). Furthermore in this interpretation the computation is less affected by possible failures of synapses.

- In another biological interpretation one can replace each neuron v in $\mathcal{N}_{N,\varepsilon,\delta}$ by a pool P_v of neurons (as in a synfire chain; see Abeles *et al.* 1993). The firing time t_v of neuron v is then replaced by the mean firing time of neurons in P_v , and an EPSP from v is replaced by a sum of EPSPs from neurons in P_v . This interpretation has the advantage that it is less affected by jitter in the firing times of individual neurons and by stochastic failures of individual synapses. Furthermore, the rising segment of a sum of EPSPs from P_v is likely to be substantially longer than that of an individual EPSP. Hence in this interpretation, the parameter γ can be chosen substantially larger than for single neurons.

Remark 5. The auxiliary neuron a_0 that provides in the construction a reference spike at time T_{in} is not really necessary. Without such auxiliary neuron a_0 , the weights w_i of the inputs s_i are automatically normalized so that they sum up to 1 (see equation 2.3 with $\sum_{i=0}^n w_i$ replaced by $\sum_{i=1}^n w_i$). Such normalization is disadvantageous for simulating an arbitrary given sigmoidal gate G , but it may be desirable in a biological context.

Remark 6. Our construction was based on a specific temporal coding in terms of reference times T_{in} and T_{out} . Some biological systems may provide such reference times that are time locked with the onset of a sensory stimulus. However, our construction can also be adapted to other types of temporal coding that require no reference times. For example, the basic equation at the end of Section 1—equation 1.1—also yields a mechanism for carrying out analog computations with regard to the scheme of “competitive temporal coding” discussed in Thorpe and Imbert (1989). In this coding scheme the firing time t_{a_i} of neuron a_i encodes the analog variable $(t_{a_i} + d_i) - \min_{j=1,\dots,n} (t_{a_j} + d_j)$, where $d_j := d_{a_j,v}$ is the delay between neuron a_j and v . For $T := \min_{j=1,\dots,n} (t_{a_j} + d_j)$ we have according to equation 1.1,

$$t_v = \frac{\Theta - BN^-}{\sum_{i=1}^n w_i} + T + \frac{\sum_{i=1}^n w_i \cdot ((t_{a_i} + d_i) - T)}{\sum_{i=1}^n w_i}. \quad (2.9)$$

In this coding scheme the firing time t_v encodes the analog value $t_v - \min_{\tilde{v} \in L} t_{\tilde{v}}$, where L is the layer of neurons to which v belongs. Thus, equation 2.9 provides the mechanism for computing a segment of a linear function for inputs and outputs in competitive temporal coding. Hence our construction also provides a method for simulating multilayer neural nets with regard to this alternative coding scheme. No reference times T_{in} or T_{out} are needed for that.

In this way, one can implement, for example, a multilayer pattern analysis as described in Remark 2 without making any use of reference times. If one wants to simulate gates that employ a sigmoidal activation function by spiking neurons with competitive temporal coding, one can add lateral excitation among all neurons on the same layer L . In this way the neurons in L are forced to fire within a rather small time window (corresponding to the bounded output range of a sigmoidal activation function). In other applications, it appears to be more advantageous to employ instead lateral inhibition. This also has the effect of preventing the value of $t_v - \min_{\tilde{v} \in L} t_{\tilde{v}}$ from becoming too large.

Remark 7. For the sake of simplicity we have considered in the preceding construction only feedforward computations on networks of spiking neurons. However, the same simulation method can be used to simulate recurrent sigmoidal neural nets by recurrent networks of spiking neurons. For example, in this way one gets a novel implementation of a *Hopfield net* (with synchronous updates). At each “round” of the computation, the output of a unit of the Hopfield net is encoded by the firing time of a corresponding spiking neuron relative to that of other neurons in the network. For example, one may assume that a neuron gives the highest possible output value if it is among the first neurons that fire at that round. Each spiking neuron simulates in temporal coding a sigmoidal unit whose inputs are the firing times of the other spiking neurons in the network at the previous round. One can employ here competitive temporal coding (see Remark 6); hence no reference times or external clock are needed.

A stable state of the Hopfield net is reached in this implementation (with competitive temporal coding) if and only if all neurons fire “regularly” (i.e., at regular intervals with a common interspike interval) but generally with different phases. These phases encode the output values of the individual neurons, and together these phase differences represent a “recalled” stored pattern of the Hopfield net. Thus, each stored pattern of the Hopfield net is realized by a different assignment of phase differences in some stable oscillation.

This novel implementation of a Hopfield net with spiking neurons in temporal coding has, apart from its high computation speed, another feature that is possibly of biological interest: whereas the input to such network of spiking neuron can be transient (encoded in the relative timing of the firing of each neuron in the network at the first round), its output is available over

a longer time period, since it is encoded in the phases of the neurons in a stable global oscillation of the network. Hence, this implementation makes it possible that even in the rapidly fluctuating environment of temporal coding with single spikes, the outputs from different neural subsystems (which may operate at different time scales) can be collected and integrated by a larger neural system.

3 Universal Approximation Properties of Networks of Noisy Spiking Neurons in Temporal Coding

One of the most interesting and useful features of sigmoidal neural nets N is the fact that if $[0, \gamma]$ is the range of their activation functions, they can approximate for any given natural numbers n and k any given continuous function F from $[0, \gamma]^n$ into $[0, \gamma]^k$ within any given $\varepsilon > 0$ (with regard to uniform convergence, i.e., the L_∞ norm). Furthermore it suffices to consider for this purpose feedforward nets N with just one hidden layer of neurons and (roughly) any activation function σ that is not a polynomial (Leshno *et al.* 1993). In addition, many years of experiments with backpropagation and other learning rules for sigmoidal neural nets have shown that for most concrete application problems, sigmoidal neural nets with relatively few hidden units allow a satisfactory approximation of the underlying target function F .

The result from the preceding section allows us to transfer these results to networks of spiking neurons with temporal coding. If some feedforward sigmoidal neural net N approximates an arbitrary given continuous function $F: [0, \gamma]^n \rightarrow [0, \gamma]^k$ within an ε (with regard to the L_∞ norm), then with probability $\geq 1 - \delta$ the network $\mathcal{N}_{N,\varepsilon,\delta}$ of spiking neurons that we constructed in Section 2 approximates the same F within 2ε (with regard to the L_∞ norm). Furthermore, if N has only a small number p of layers, the computation time of $\mathcal{N}_{N,\varepsilon,\delta}$ can be bounded (for biologically reasonable choices of the parameters involved) by $10 \cdot p$ ms.

Thus if one neglects the fact that the fan-in of biological neurons is bounded by some fixed (although rather large) constant, the preceding theoretical results suggest that networks of biological neurons can (in spite of their “slowness”) approximate arbitrary continuous functions $F: [0, \gamma]^n \rightarrow [0, \gamma]^k$ within any given ε with a computation time of not more than 20 ms.

Finally, we would like to point out that our approximation result holds not only for the particular way of encoding analog inputs and outputs by firing times that was considered in the previous section but basically for *any* coding that is continuously related to it. More precisely, let $n, \tilde{n}, k, \tilde{k}$ be arbitrary natural numbers. Let $Q: [0, \gamma]^n \rightarrow [0, 1]^{\tilde{n}}$ be any continuous function that specifies a method of “decoding” \tilde{n} variables ranging over $[0, 1]$ from the firing times of n neurons during some time window of length γ (whose

end point is marked by the firing time of the first one of these n neurons), and let $R: [0, 1]^{\tilde{k}} \rightarrow [0, \gamma]^k$ be any continuous and invertible function that describes a method for “encoding” \tilde{k} output variables ranging over $[0, 1]$ by the firing times of k neurons during some time window of length γ (whose end point is marked by the firing time of the first one of these k neurons). Then for any given continuous function $\tilde{F}: [0, 1]^{\tilde{n}} \rightarrow [0, 1]^{\tilde{k}}$ the composition $F: R \circ \tilde{F} \circ Q$ of these three functions is a continuous function from $[0, \gamma]^n$ into $[0, \gamma]^k$. According to our preceding argument, there exists a network $\tilde{\mathcal{N}}_{\varepsilon, \delta}$ of noisy spiking neurons (with one “hidden layer”) such that for any $\underline{x} \in [0, \gamma]^n$ one has for the output $\tilde{\mathcal{N}}_{\varepsilon, \delta}(\underline{x})$ of this network $\tilde{\mathcal{N}}_{\varepsilon, \delta}$ that $\|F(\underline{x}) - \tilde{\mathcal{N}}_{\varepsilon, \delta}(\underline{x})\| \leq \varepsilon$ with probability $\geq 1 - \delta$, where $\|\cdot\|$ can be any common norm. Hence $\tilde{\mathcal{N}}_{\varepsilon, \delta}$ approximates for arbitrary inputs the given function $\tilde{F}: [0, 1]^{\tilde{n}} \rightarrow [0, 1]^{\tilde{k}}$ for arbitrarily chosen continuous functions R, Q for coding and decoding of analog variables by firing times of spiking neurons with a precision of at least $\sup \{\|R^{-1}(\underline{y}) - R^{-1}(\underline{y}')\|: \underline{y}, \underline{y}' \in [0, 1]^k \text{ and } \|\underline{y} - \underline{y}'\| \leq \varepsilon\}$. Thus, if the inverse R^{-1} of the function R is uniformly continuous, one can approximate \tilde{F} with regard to neural coding and decoding described by R and Q with arbitrarily high precision by networks of noisy spiking neurons with just one hidden layer.

4 Consequences for Learning

In the traditional interpretation of (unsupervised) Hebbian learning, a synapse is strengthened if both the presynaptic and the postsynaptic neurons are simultaneously “active” (i.e., both give high output values in terms of their current firing rates). In the implementation of a sigmoidal neural net N by a network of spiking neurons $\mathcal{N}_{N, \varepsilon, \delta}$ in Section 2, the “weights” r_i of N are in fact modeled by the “strengths” w_i of corresponding synapses between spiking neurons. However, the information whether both the presynaptic and postsynaptic neurons give high output values in temporal coding can no longer be read off from their “activity” but only from the time difference T_i between their firing times.

This observation gives rise to the question of whether there are biological mechanisms known that support a modulation of the efficacy (i.e. “strength”) w_i of a synapse as a function of this time difference T_i . If one works in the linear range of the simulation $\mathcal{N}_{G, \varepsilon}$ of a π_γ -gate G according to Section 2 (where G computes the function $\langle s_1, \dots, s_n \rangle \mapsto \sum_{i=1}^n r_i \cdot s_i$) and $h_i(t)$ describes an EPSP, then for Hebbian learning it would be desirable to increase $w_i = \lambda \cdot r_i$ if s_i is close to $\sum_{i=1}^n r_i \cdot s_i$; that is, if the difference in firing times $T_i := t_v - (T_{in} - s_i) = T_{out} - \sum_{i=1}^n r_i \cdot s_i - T_{in} + s_i$ is close to $T_{out} - T_{in}$. On the other hand, one would like to “decrease” w_i if T_i is substantially smaller

or larger than $T_{out} - T_{in}$. Hence, a Hebbian-style unsupervised learning rule of the form

$$\begin{aligned}\Delta w_i &= \frac{1}{1 + \beta \cdot (T_i - (T_{out} - T_{in}))^2} - \rho \quad \text{or} \\ \Delta w_i &= e^{-\beta(T_i - (T_{out} - T_{in}))^2} - \rho\end{aligned}\tag{4.1}$$

(for suitable parameters $\beta, \rho > 0$) would be meaningful in this context.

Recent results from neurobiology (Stuart and Sakmann 1994) show that action potentials in neocortical pyramidal cells are actively (i.e., supported by voltage-dependent channels) propagated *backward* from the soma into the dendrites (see also Jaffe *et al.* 1992). Hence the time difference T_i between the firing of the presynaptic and the postsynaptic neurons is in principle available to each synapse. Furthermore new experimental results (Markram 1995; Markram and Sakmann 1995) show that in vitro, the efficacy of synapses of neocortical pyramidal neurons is in fact modulated as a function of this time difference T_i .

There exists one interesting structural difference between this interpretation of Hebbian learning in temporal coding and its traditional interpretation: the time difference T_i provides a synapse with the full information about the correlation between the output values of the pre- and postsynaptic neurons in temporal coding, no matter whether both neurons give high or low output values. However, in the traditional interpretation of Hebbian learning in terms of firing rates, the efficacy of a synapse is increased only if both neurons give high output values (in frequency coding).

An implementation of Hebbian learning in the temporal domain is also appealing in the context of pulse stream VLSI (i.e., “silicon spiking neurons”). These artificial neural nets are much faster than biological spiking neurons: they can work with interspike intervals in the microsecond range. If for a hardware implementation of a sigmoidal gate with pulse stream VLSI according to the construction of Section 2, a Hebbian learning rule can be applied in the temporal domain after each pulse, such a chip may be able to carry out so many learning steps per second that it could in principle (neglecting input-output constraints) overcome the main impediment of traditional artificial neural nets: their low learning speeds.

So far we have assumed in the construction of $\mathcal{N}_{N,\epsilon,\delta}$ in Section 2 that the time delays d_i between the presynaptic neurons a_i and the postsynaptic neuron v (i.e., the time needed until an action potential from a_i can influence the potential at the trigger zone of v) were the same for all neurons a_i . Differences among these delays have the effect of providing an additive correction $s_i \mapsto s_i + d_i$ to the variable that is communicated in temporal coding from a_i to v . Hence, they also have the ability to give different “weights” to different input variables. In a biological context, they appear to be useful for providing to the network a priori information about its computational

task, so that Hebbian learning can be viewed as “fine-tuning” on top of this preprogrammed information. If there exist biological mechanisms for modulating such delays (see Hopfield 1995; Kempter *et al.* 1996), they would provide in addition to a short-term memory via synaptic modulation a separate mechanism for storing and adapting long-term memory via differences in the delays.

5 Conclusions

We have shown in this article that there exists a rather simple way to compute linear functions and to simulate feedforward as well as recurrent sigmoidal neural nets in temporal coding by networks of noisy spiking neurons. In contrast to the traditionally considered implementation via frequency coding, the new approach yields a computation speed that is faster by several orders of magnitude. In fact, to the best of our knowledge, it provides the first theoretical model that is able to explain the experimentally observed speed of fast information processing in the cortex on the basis of relatively slow spiking neurons as computational units.

Further experiments will be needed to determine whether this theoretical model is biologically relevant. One problem is that we do not know in which way batch inputs (consisting of many analog variables in parallel) are encoded by biological neural systems. The existing results on neural coding (see Rieke *et al.* 1996) address only the coding of time series, that is, sequential analog inputs. However, if further experiments showed that the input-dependent firing times in visual cortex, as reported in Bair *et al.* (1994), vary in a continuous (i.e., piecewise continuous) manner in response to smooth changes of complex inputs, this would provide some support to the style of theoretical models for analog computation with spiking neurons that is considered here.

Furthermore, a biological realization of recurrent neural nets (e.g., Hopfield nets) in temporal coding with spiking neurons, as proposed in Remark 7 in Section 2, would predict the occurrence of oscillations in neural systems (especially systems involved in pattern recognition and working memory), where the *phase* of individual neurons (or of pools of neurons) with regard to this oscillation is *input dependent*.

A noteworthy feature of the constructions presented here is that the “weights” (efficacies) of synapses are in principle able to play in temporal coding the same role as in the more familiar context of *firing rate coding*. Hence, all theories and experimental results regarding adaptation of neural circuits via synaptic plasticity can in principle also be applied to such computations in *temporal coding*.

Furthermore, a closer look shows that the networks that we have constructed for analog computations in *temporal coding* can compute the *same* analog function on a different time scale in *firing rate coding*. This possible *dual role* of the circuits of spiking neurons constructed here is of interest in

the context of biology. There exists empirical evidence that some neural systems carry out in addition to a very fast preliminary computation a more thorough subsequent computation in terms of firing rates, which has the ability to integrate relevant contributions from several neural subsystems but takes several hundred milliseconds longer (see, e.g., Heller *et al.* 1995).

So far it is not known whether there exists a biological neural system whose computational organization can be understood as an implementation of a sigmoidal neural net. However, one may argue that this conjecture can be supported by a heuristic “optimality principle” (whereby biological systems often find the theoretically best use for their “hardware”). Feedforward sigmoidal neural nets with few layers are the most powerful and fastest parallel computational model that is known to date, from the point of view of computational complexity theory and engineering applications. In addition, feedforward and recurrent sigmoidal neural nets are the only parallel computational models known that support self-organization (and hence do not require cumbersome and error-prone central “programming”). Hence if the “hardware” of biological neural systems allows in principle an efficient implementation of sigmoidal neural nets, it is not unreasonable to assume that this possibility has been realized by at least some biological neural systems (in spite of our lack of knowledge about the concrete functions that they compute).

Apart from their biological interpretation, the construction of this article appears to be of interest also in the context of hardware implementations of neural nets via pulse-stream VLSI (see, e.g., Pratt 1989; Horinchi *et al.* 1991; Murray and Tarassenko 1994; Jahnke *et al.* 1995). A simulation of a feedforward sigmoidal neural net by pulse stream VLSI along the lines of this construction offers the possibility of combining extremely high computation speed (simulating one parallel computation step of the sigmoidal neural net per pulse) with a noise-robust and fast transmission of internal analog variables via the relative timing of pulses from different gates. Furthermore, an implementation of Hopfield nets according to Remark 7 provides a new idea for the technological realization of content-addressable memory.

We have shown in this article that networks of spiking neurons with temporal coding have at least the same computational power as sigmoidal neural nets of roughly the same size and depth. Together with a recent separation result (Maass 1996b) this implies that networks of spiking neurons have in fact strictly more computational power.

Acknowledgments

I thank Martin Arndt, David Haussler, Geoffrey Hinton, John Hopfield, Berthold Ruf, and Gordon Shepherd for stimulating discussions related to the topic of this article. I also thank the anonymous referees for their helpful comments.

References

- Abeles, M., Bergman, H., Margalit, E., and Vaadia, E. 1993. Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. Neurophysiology* **70**, 1629–1638.
- Bair, W., Koch, C., Newsome, W., and Britten, K. 1994. Reliable temporal modulation in cortical spike trains in the awake monkey. In *Proc. Symposium on Dynamics of Neural Processing*, Washington, DC.
- Bialek, W., and Rieke, F. 1992. Reliability and information transmission in spiking neurons. *Trends in Neuroscience* **15**, 428–434.
- Bryant, H. L., and Segundo, J. P. 1976. *J. Physiology* **260**, 279–314.
- Ferster, D., and Spruston, N. 1995. Cracking the neuronal code. *Science* **270**, 756–757.
- Gerstner, W., and van Hemmen, J. L. 1994. How to describe neuronal activity: Spikes, rates, or assemblies? In *Advances in Neural Information Processing Systems* **6**, pp. 463–470. Morgan Kaufmann, San Mateo, CA.
- Heller, J., Hertz, J. A., Kjær, T. W., and Richmond, B. J. 1995. Information flow and temporal coding in primate pattern vision. *J. Computational Neuroscience* **2**(3), 175–193.
- Hopfield, J. J. 1991. Olfactory computation and object perception. *Proc. Nat. Acad. Sci. USA* **88**, 6462–6466.
- Hopfield, J. J. 1995. Pattern recognition computation using action potential timing for stimulus representations. *Nature* **376**, 33–36.
- Horinchi, T., Lazzaro, J., Moore, A., and Koch, C. 1991. A delay-line based motion detection chip. In *Advances in Neural Information Processing Systems* **3**, pp. 406–412. Morgan Kaufmann, San Mateo, CA.
- Jaffe, D. B., Johnston, D., Lasser-Ross, N., Lisman, J. E., Miyakawa, H., and Ross, W. N. 1992. The spread of Na^+ spikes determines the pattern of dendritic Ca^{2+} entry into hippocampal neurons. *Nature* **357**, 244–246.
- Jahnke, A., Roth, U., and Klar, H. 1995. Towards efficient hardware for spike-processing neural networks. In *Proc. of the World Congress on Neural Networks*, Washington, DC.
- Johnson, D. S. 1990. A catalog of complexity classes. In *Handbook of Theoretical Computer Science*, J. V. Leeuwen, ed., vol. A, pp. 67–162. MIT Press, Cambridge, MA.
- Kempter, R., Gerstner, W., van Hemmen, J. L., and Wagner, H. 1996. Temporal coding in the sub-millisecond range: Model of barn owl auditory pathway. In *Advances in Neural Information Processing Systems* **8**, pp. 124–130. MIT Press, Cambridge, MA.
- Leshno, M., Lin, V. Y., Pinkus, A., and Schocken, S. 1993. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Networks* **6**, 861–867.
- Linsker, R. 1988. Self-organization in a perceptual network. *Computer Magazine* **21**, 105–117.

- Maass, W. 1995. On the computational power of networks of spiking neurons. In *Advances in Neural Information Processing Systems 7*, pp. 183–190. MIT Press, Cambridge, MA.
- Maass, W. 1996a. On the computational power of noisy spiking neurons. In *Advances in Neural Information Processing Systems 8*, pp. 211–217. MIT Press, Cambridge, MA.
- Maass, W. 1996b. Networks of spiking neurons: The third generation of neural network models. To appear in *Neural Networks*.
- Mainen, Z. F., and Sejnowski, T. J. 1995. Reliability of spike timing in neocortical neurons. *Science* **268**, 1503–1506.
- Markram, H. 1995. Neocortical pyramidal neurons scale the efficacy of synaptic input according to arrival time: A proposed selection principle of the most appropriate synaptic information. In *Proc. of Cortical Dynamics*, pp. 10–11, Jerusalem.
- Markram, H., and Sakmann, B. 1995. Action potentials propagating back into dendrites trigger changes in efficacy of single-axon synapses between layer V pyramidal neurons. In *Proc. of the Conference of the Society for Neuroscience*, vol. 21.
- Mel, B. W. 1993. Synaptic integration in an excitable dendritic tree. *J. Neurophysiology* **70**, 1086–1101.
- Murray, A., and Tarassenko, L. 1994. *Analogue Neural VLSI: A Pulse Stream Approach*. Chapman & Hall, London.
- Perrett, D. I., Rolls, E. T., and Caan, W. C. 1982. Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research* **47**, 329–342.
- Pratt, G. A. 1989. *Pulse Computation*. PhD. thesis, MIT.
- Rieke, F., Warland, D., Ruyter van Steveninck, R., and Bialek, W. 1996. *SPIKES: Exploring the Neural Code*. MIT Press, Cambridge, MA.
- Rolls, E. T. 1994. Brain mechanisms for invariant visual recognition and learning. *Behavioural Processes* **33**, 113–138.
- Rolls, E. T., and Tovee, M. J. 1994. Processing speed in the cerebral cortex, and the neurophysiology of visual backward masking. *Proc. Roy. Soc. B* **257**, 9–15.
- Roychowdhury, V., Siu, K., and Orlitzky, A., eds. 1994. *Theoretical Advances in Neural Computation and Learning*. Kluwer, Boston.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. 1986. Learning representations by back-propagating errors. *Nature* **323**, 533–536.
- Segundo, J. P. 1994. Noise and the neurosciences: A long history, a recent revival, and some theory. In *Origins: Brain and Self Organization*, K. H. Pribram, ed., pp. 300–331. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Sejnowski, T. J. 1995. Time for a new neural code? *Nature* **376**, 21–22.
- Shepherd, G. M. 1994. *Neurobiology*. 3d ed. Oxford University Press, New York.
- Siu, K., Roychowdhury, V., and Kalaith, T. 1995. *Discrete Neural Computation*. Prentice Hall, Englewood Cliffs, NJ.
- Softky, W. 1994. Sub-millisecond coincidence detection in active dendritic trees. *Neuroscience* **58**, 13–41.
- Stuart, G. J., and Sakmann, B. 1994. Active propagation of somatic action potentials into neocortical pyramidal cell densities. *Nature* **367**, 69–72.

- Thorpe, S. J., and Imbert, M. 1989. Biological constraints on connectionist modelling. In *Connectionism in Perspective*, R. Pfeifer, Z. Schreter, F. Fogelman-Soulié, and L. Steels, eds., pp. 63–92. Elsevier, New York.
- von Neumann, J. 1956. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In *Automata Studies*, C. E. Shannon and J. MacCarthy, eds. Princeton University Press, Princeton, NJ.

Received December 15, 1995; accepted April 30, 1996.