A Model for Fast Analog Computation Based on Unreliable Synapses*

Wolfgang Maass and Thomas Natschläger

Institute for Theoretical Computer Science, Technische Universität Graz
Klosterwiesgasse 32/2, A-8010 Graz, Austria
email: (maass, tnatschl)@igi.tu-graz.ac.at

Abstract

We investigate through theoretical analysis and computer simulations the consequences of unreliable synapses for fast analog computations in networks of spiking neurons, with analog variables encoded by the current firing activities of pools of spiking neurons. Our results suggest a possible functional role for the well-established unreliability of synaptic transmission on the network level. We also investigate computations on time series and Hebbian learning in this context of space-rate coding in networks of spiking neurons with unreliable synapses.

1 Introduction

This article explores links between two levels of modeling computation in biological neural systems: the level of individual synapses and spiking neurons on one hand, and the network level on the other hand. Such links are of interest under the hypothesis that important aspects of the computational function of neuronal activity can only be understood on the larger scale of networks consisting of hundreds and more neurons. One particular challenge is the task to provide models for fast analog computation in neuronal systems that are consistent with experimental data. It has been demonstrated by Thorpe et al. (1996) and others that biological neural systems involving 10 or more synaptic stages are able to carry out complex computations within 100 to 150 ms.

*Research for this article was partially supported by the ESPRIT Working Group NeuroCOLT, No. 8556, and the Fonds zur Förderung der wissenschaftlichen Forschung (FWF), Austria, project P12153.

This cannot be be explained through models based on an encoding of analog variables through firing rates of spiking neurons, since the firing rates in these neural systems are typically well below 100 Hz and interspike intervals are highly variable Koch (1999). In addition it has recently been argued that synaptic depression makes firing rates above 20 Hz indistinguishable for the postsynaptic neurons Abbott et al. (1997).

One approach for explaining the possibility of fast analog computation relies on the assumption that the relevant analog variables are encoded in small temporal differences among the firing times of neurons Thorpe et al. (1996); Hopfield (1995); Maass (1997). These models are able to explain the possibility of fast analog computation in networks where neuronal firing and synaptic transmission is highly reliable, or where the average firing times of pools of neurons encode analog variables on a time scale of a few ms's. Although some evidence for such coding has been found, a more common type of coding encountered so far in vertebrate cortex is a population coding where information about the stimulus or subsequent responses of the organism are encoded in a space-rate code, i.e., in the fractions of neurons in various pools that fire within some short time interval (say, of length Δ between 5 and 10 ms). We will refer to this coding scheme as *space-rate coding* in the following. We will analyze in this article possible links between well-known response characteristics of individual neurons and synapses on one hand, and computational functions of networks of neurons with regard to space-rate coding on the other hand, as indicated in Table 1. In particular we will address the possible computational role of the unreliability of synaptic transmission, which has so far been ignored in this context.

In section 2 we will show that unreliability of synaptic transmission can be used as an essential ingredient for a model for fast analog computation in space-rate coding that spreads firing activity uniformly among all neurons in a pool. We exhibit a suitable tool from probability theory for analyzing large scale effects of unreliable synaptic transmission (the Berry-Esseen Theorem), and we analyze inherent noise sources of this computational model. Our theoretical analysis is complemented by results of computer simulations. In section 3 we turn to the question what other types of computations (besides approximating arbitrary continuous functions between vectors of analog input- and output values) can be induced by synaptic unreliability in combination with other details of the neuronal hardware on the larger scale of networks of neurons with space-rate coding. We show that for computations on time series the interaction of the time courses of EPSP's and IPSP's, refractory effects of neurons, and recurrent excitation and inhibition provides in combination with unreliable synapses the means for realizing in space-rate coding a rich class of linear filters with finite and infinite impulse response. Finally in section 4 we briefly address the relationship between temporal learning rules for individual neurons and Hebbian learning in space-rate coding.

Table 1: Relationships between details of neuronal hardware and possible computational functions on the network level that are addressed in this article.

neuron level	**network level**
synaptic unreliability	graded response of firing activity in pools of neurons, yielding the power to approximate arbitrary continuous functions in space-rate coding
decaying parts of EPSP's and IPSP's in combination with refractory effects and recurrent connections	emulation of arbitrary given linear filters with finite and infinite impulse response with regard to space-rate coding
time-sensitive rule for LTP	Hebbian learning for space-rate coding

2 A Model for Fast Analog Computation in a Space-Rate Code

We say the analog variable $x \in [0,1]$ is encoded by a pool U of N neurons in a space-rate code if during a short time interval of length Δ a total of Nx neurons fire. If the time interval is short enough, say $\Delta = 5$ ms, one can assume that each neuron fires at most once during this time interval.

Although there exists substantial empirical evidence that many cortical systems encode relevant analog variables by such space-rate code, it has remained unclear how networks of spiking neurons *compute* in terms of such a code. Some of the difficulties become apparent if one just wants to understand for example how the trivial linear function $f(x) = x/2$ can be computed by such a network if the input $x \in [0,1]$ is encoded by a space-rate code in a pool U of neurons and the output $f(x) \in [0, 1/2]$ is supposed to be encoded by a space-rate code in another pool V of neurons. If one assumes that all neurons in V have the same firing threshold and that reliable synaptic connections from all neurons in U to all neurons in V exist with approximately equal weights, a firing of a fraction x of neurons in U during a short time interval will typically trigger *almost none* or *almost all* neurons in V to fire, since they all receive about the same input from U.

Several mechanisms have already been suggested that could in principle achieve a *smooth* graded response in terms of a space-rate code in V instead of a binary "all or none" firing, such as strongly varying firing thresholds or different numbers of synaptic connections from U for different neurons $v \in V$ Wilson and Cowan (1972). Neither option is completely satisfactory, since firing thresholds of biological neurons appear to be rather homogeneous and both options would fail to spread average activity over all neurons in V. Hence they would fail to spread the energy consumption in the neuronal ensembles and also would make the computation less robust against failures of individual neurons.

In this section we introduce an alternative model for analog computing in space-rate coding that takes the well-known stochastic properties of biological synapses into account. We show that the unreliability of synaptic transmission provides a very useful mechanism for reliable analog computation in space-rate coding. Our model does not require that the firing threshold of the neurons involved are different, and it automatically spreads the firing activity uniformly over all neurons within a pool. In subsection 2.3 inherent noise-sources of this model for analog computation are investigated. In subsection 2.5 it is shown that our model induces an analog version of the familiar model of a synfire chain, that is computationally more powerful and simultaneously more consistent with experimental data than the traditional "digital" version of a synfire chain.

2.1 Definition of the Computational Model

We start by addressing the most basic question concerning computations in space-rate coding: how can the firing activity in a pool V of neurons be related to the firing activity in n presynaptic pools $U_1, \ldots, U_n$ of neurons ? For simplicity we assume that n pools $U_1, \ldots, U_n$ consisting of N neurons each are given, and we also assume that all neurons in these pools have synaptic connections to all neurons in another pool V of N neurons.[1] We write x_i for the analog variable ranging over $[0, 1]$ that is encoded by space-rate coding in pool U_i, and y for the analog variable ranging over $[0, 1]$ that is encoded by space-rate coding in pool V. We assume throughout this article that for each pool U_i all neurons in U_i are excitatory, or all neurons in U_i are inhibitory. By choosing the number n of pools U_i sufficiently large one can approximate with this model another type of computational model where one has instead of discrete pools of neurons, that encode one analog variable each, a continuous spatial pattern of firing activity. We will address in subsection 2.4 the question what functions can be computed by *multi-layer* feedforward networks in terms of space-rate coding, and we will address in section 3 the question which additional computational capabilities arise when one takes *recurrent* connections involving excitatory and inhibitory neurons into account.

[1]Our results remain valid for connection patterns given by sparser random graphs.

In accordance with standard results from neurophysiology we assume in our model that an action potential ("spike") from a neuron $u \in U_i$ triggers with a certain probability r_{vu} ("non-failure probability") the release of one or several vesicles filled with neurotransmitter at one or several release sites of the synapses between neurons $u \in U_i$ and $v \in V$. The data from Markram et al. (1997a); Larkman et al. (1997); Dobrunz and Stevens (1997) strongly suggest that in the case of a release the amplitude of the resulting EPSP in neuron v is a stochastic quantity. Consequently we model the amplitude of the EPSP (or IPSP) in the case of a release by a random variable a_{vu} with probability density function ϕ_{vu}.

Empirical data show that the variance of the distribution ϕ_{vu} is typically rather high Markram et al. (1997a); Larkman et al. (1997); Stevens and Zador (1998). This high variability can be traced back to two sources:

a) It is uncertain whether a vesicle is released at a single synaptic release site in response to an action potential from the presynaptic neuron. Thus the amplitude of the PSP in the postsynaptic neuron v varies in dependence of the actual number of vesicle releases that take place at the different synaptic release sites between neurons u and v.

b) Even at a single release site the amplitude of the EPSP ("quantal size") caused by a vesicle release at this site varies from trial to trial Dobrunz and Stevens (1997); Auger et al. (1998).

Most of our results – in particular the results in the following subsection – hold for *arbitrary* probability density functions ϕ_{vu} and arbitrary values of the parameters r_{vu}. In subsection 2.3 we will address the question what effect specific probability density functions ϕ_{vu} may have on analog computations in terms of space rate coding.

The equation for the probability density function ϕ_{vu} for the case of multiple release sites and a Gaussian distribution of the quantal size at each release site – which is modeled in our computer simulations – is discussed in the Appendix. Figure 1C shows an example of ϕ_{vu} for a synapse with 5 release sites.

2.2 Graded Response through Unreliable Synapses

In this subsection we introduce analytical tools for estimating the firing activity y in pool V in terms of the firing activities $x_1, \ldots, x_n$ in n presynaptic pools $U_1, \ldots, U_n$. We show that y can basically be approximated by a function of a weighted sum of the x_i, although some unexpected complications will turn up in this analysis. As the first step we demonstrate that the model defined above is indeed able to produce a graded output y encoded in a space-rate code in pool V such that the activity is uniformly distributed over all neurons $v \in V$. We prove this fact theoretically by employing a special version of the central limit theorem of probability theory, the Berry-Esseen Theorem. Our theoretical results are then compared with computer simulations of our model.

Consider an idealized mathematical model where all neurons which fire in the pool U_i fire synchronously at time T_{in}, and the probability that a neuron $v \in V$ fires (at time T_{out}) can be described by the probability that the sum h_v of the amplitudes of EPSP's and IPSP's resulting from firing of neurons in the pools $U_1, \ldots, U_n$ exceeds the firing threshold θ (which is assumed to be the same for all neurons $v \in V$). We assume in this section that the firing rates of neurons in pool V are relatively low, so that the impact of their refractory period can be neglected. We investigate refractory effects in section 3. The random variable (r.v.) h_v is the sum of random variables h_{vu} for all neurons $u \in \bigcup_{i=1}^n U_i$, where h_{vu} models the contribution of neuron u to h_v. We assume that h_{vu} is nonzero only if neuron $u \in U_i$ fires at time T_{in} (which occurs with probability x_i)[2] and if the synapse between u and v releases one or several vesicles (which occurs with probability r_{vu} whenever u fires). If both events occur then the value of h_{vu} is chosen according to some probability density function ϕ_{vu}. The functions ϕ_{vu}, as well as the parameters r_{vu}, are allowed to vary arbitrarily for different pairs u, v of neurons. For each neuron $v \in V$ we consider the sum $h_v = \sum_{i=1}^n \sum_{u \in U_i} h_{vu}$ of the r.v.'s h_{vu} and we assume that v fires at time T_{out} if and only if $h_v \geq \theta$. Although the r.v.'s h_{vu} may have quite different distributions (for example due to different ϕ_{vu} and r_{vu}), their stochastic independence allows us to approximate the firing probability $\mathrm{P}\{h_v \geq \theta\}$ through a normal distribution Φ. The Berry-Esseen Theorem Petrov (1995) implies that

$$|\mathrm{P}\{h_v \geq \theta\} - (1 - \Phi(\theta;\, \mu_v, \sigma_v))| \;\leq\; 0.7915 \frac{\rho_v}{\sigma_v^3} \;, \tag{1}$$

where $\Phi(\cdot;\, \mu_v, \sigma_v)$ denotes the normal distribution function with mean μ_v and variance σ_v^2. The three moments occurring in (1) can be related to the r.v.'s h_{vu} through the equations $\mu_v = \sum_{i=1}^n \sum_{u \in U_i} \mathrm{E}[h_{vu}]$, $\sigma_v^2 = \sum_{i=1}^n \sum_{u \in U_i} \mathrm{Var}[h_{vu}]$, and $\rho_v = \sum_{i=1}^n \sum_{u \in U_i} \mathrm{E}[|h_{vu} - \mathrm{E}[h_{vu}]|^3]$. According to the definition of the r.v. h_{vu} we have $\mathrm{E}[h_{vu}] = x_i r_{vu} \bar{a}_{vu}$ and $\mathrm{E}[h_{vu}^2] = x_i r_{vu} \hat{a}_{vu}$ where $\bar{a}_{vu} = \int a\phi_{vu}(a)da$ denotes the mean EPSP (IPSP) amplitude and $\hat{a}_{vu} = \int a\phi_{vu}(a)da$ denotes the

[2] This holds if the pool size is large enough such that we can treat x_i (y) as the probability that a neuron $u \in U_i$ ($v \in V$) will fire once during a certain input (output) interval of length Δ.

second moment. Hence we can assign to μ_v and σ_v in (1) the values

$$\mu_v = \sum_{i=1}^{n} \sum_{u \in U_i} x_i r_{vu} \bar{a}_{vu} \quad , \tag{2}$$

$$\sigma_v^2 = \sum_{i=1}^{n} \sum_{u \in U_i} (x_i r_{vu} \hat{a}_{vu} - x_i^2 r_{vu}^2 \bar{a}_{vu}^2) \quad . \tag{3}$$

A closer look reveals that the right hand side of (1) scales like $N^{-1/2}$.[3] Hence, equation (1) implies that for large N we can approximate the firing probability $\mathrm{P}\{h_v \geq \theta\}$ by the term $1 - \Phi(\theta;\, \mu_v, \sigma_v)$, which smoothly grows with μ_v. The gain of this sigmoidal function depends on the size of σ_v. In particular, if synaptic transmission were reliable, this function would degenerate to a step function. With the definition of the formal weights $w_{vi} := \sum_{u \in U_i} r_{vu} \bar{a}_{vu}$ we have $\mu_v = \sum_{i=1}^{n} w_{vi} x_i$, and hence $1 - \Phi(\theta;\, \mu_v, \sigma_v)$ smoothly grows with the weighted sum $\sum_{i=1}^{n} w_{vi} x_i$ of the inputs x_i.

So far we have just considered the probability $\mathrm{P}\{h_v \geq \theta\}$ that a single neuron $v \in V$ will fire, but we are really interested in the expected fraction y of neurons in pool V which will fire in response to a firing of a fraction x_i of neurons in the pools U_i for $i = 1, \ldots, n$. According to equation (1) one can approximate y for sufficiently large pool sizes N by

$$y = \frac{1}{N} \sum_{v \in V} \mathrm{P}\{h_v \geq \theta\} = \frac{1}{N} \sum_{v \in V} 1 - \Phi(\theta;\, \mu_v, \sigma_v) \quad .$$

Hence y is approximated by an average of the N sigmoidal functions $1 - \Phi(\theta;\, \mu_v, \sigma_v)$. If the weights w_{vi} have similar values for different $v \in V$ one can expect that y grows smoothly with the weighted sum $\bar{\mu} = \sum_{i=1}^{n} w_i x_i$, where we write $w_i = \sum_{v \in V} w_{vi}/N$ for the "effective weights" w_i between the pools of neurons U_i and V.

In order to test these theoretical predictions for an idealized mathematical model we have carried out computer simulations of a more detailed model consisting of more realistic models for spiking neurons and time intervals I_{in} and I_{out} of length $\Delta = 5$ ms for space-rate coding. The resulting fraction y of neurons in pool V that fired during I_{out} was measured for a large variety of different inputs $\langle x_1, \ldots, x_6 \rangle \in [0,1]^6$ encoded through the fractions of firing neurons in $U_1, \ldots, U_6$ during an earlier time interval I_{in}.[4] It is shown in

[3] More precisely, the right hand side of (1) scales like $N^{-1/2}$ if for all N the average value of the terms $\mathrm{E}[|h_{vu} - \mathrm{E}[h_{vu}]|^3]$ is uniformly bounded from above and the average value of the terms $\mathrm{Var}[h_{vu}]$ is uniformly bounded from below by a constant > 0 for $i \in \{1, \ldots, n\}$ and $u \in \bigcup_{i=1}^{n} U_i$. The latter can only be achieved for inputs where $x_j > 0$ for some j, since otherwise $\mathrm{Var}[h_{vu}] = 0$ for all $u \in U_i$ and all $i \in \{1, \ldots, n\}$. But in the case $x_i = 0$ for all $i \in \{1, \ldots, n\}$ both $\mathrm{P}\{h_v \geq \theta\}$ and $1 - \Phi(\theta;\, \mu_v, \sigma_v)$ have value 0 if $\theta > 0$ and hence the left hand side of (1) has value 0.

[4] The firing times of the individual neurons are distributed uniformly over I_{in}.

Figure 1A that y can be approximated quite well by a suitable sigmoidal "activation function" σ (more precisely, by the function σ derived in equation (6)) applied to a weighted sum $\sum_{i=1}^{6} w_i x_i$ of the inputs $x_1, \ldots, x_6$. Note that σ has not been implemented explicitly in our computational model, but rather emerges *implicitly* through the large scale statistics of the firing activity.

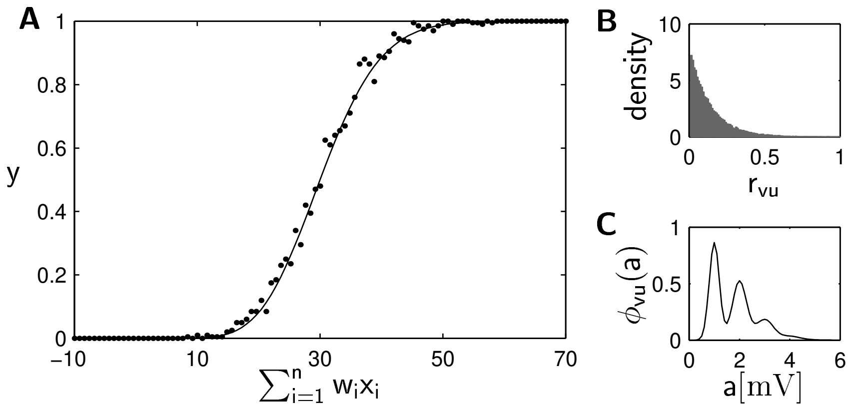

Figure 1: **A** Computer simulation of the model described in subsection 2.1 with a time interval Δ of length $\Delta = 5\,\text{ms}$ for space-rate coding and neurons modeled by the spike response model of Gerstner (1999b). We have chosen $n = 6$, a pool size $N = 200$ and $\langle w_1, \ldots, w_6 \rangle = \langle 10, -20, -30, 40, 50, 60 \rangle$ for the effective weights. Each dot is the result of a simulation with an input $\langle x_1, \ldots, x_6 \rangle$ selected randomly from $[0,1]^6$ in such a way that $\sum_{i=1}^{n} w_i x_i$ covers the range $[-10, 70]$ almost uniformly. The y-axis shows the fraction y of neurons in pool V that fire during a 5 ms time interval in response to the firing of a fraction x_i of neurons in pool U_i for $i = 1, \ldots, 6$ during an earlier time interval of length 5 ms. The solid line is a plot of $\sigma(\sum_{i=1}^{n} w_i x_i)$ as described in the text (cf. equation (6)). **B** Distribution of non-failure probabilities r_{vu} for synapses between the pools U_4 and V underlying this simulation. **C** Example of a probability density function ϕ_{vu} of EPSP amplitudes as used for this simulations. This corresponds to a synapse with 5 release sites and a release probability of 0.3. For details about the simulation see the Appendix.

The deviation of the data points in Figure 1A from the sigmoidal function $\sigma(\sum_{i=1}^{n} w_i x_i)$ (solid line) can be traced back to two independent sources of noise. One source of noise are stochastic fluctuations due to the finite pool size $N = 200$. This type of noise is guaranteed to disappear for $N \to \infty$ with $N^{-1/2}$ according to equation (1). Another source of noise is of a more systematic nature and is predicted by our preceding theoretical analysis (cf. equation (1)). For large pool sizes N the firing probability of a neuron v in pool V converges to $1 - \Phi(\theta; \mu_v, \sigma_v)$. This function depends not just on the weighted sum $\mu_v = \sum_{i=1}^{n} w_{vi} x_i$, but also on another function σ_v of the input vector $\langle x_1, \ldots, x_n \rangle$. This fact, which causes a noise for computing in space-rate coding that dominates the sampling noise already for moderate pool sizes

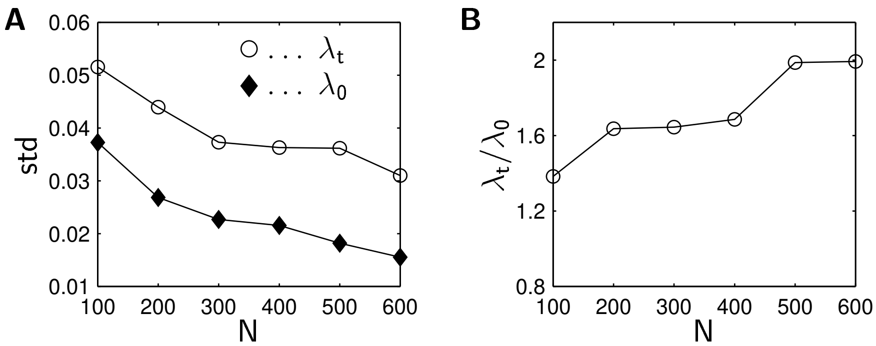

Figure 2: Quantitative study of the noise underlying the computation in our model. The model is constructed in exactly the same way as described in Figure 1, with the same values of the weights w_i. Panel A shows the empirical standard deviation (std) of the outputs y of 200 individual simulations in dependence of the pool size N for two cases: First we used the same input $\boldsymbol{x}^{(0)} = \langle x_1^{(0)}, \ldots, x_n^{(0)} \rangle$ for all of the 200 simulations and measured the standard deviation λ_0 of the observed outputs. Thus λ_0 describes the contribution of the stochastic fluctuations to the total noise. In order to test to what extent the response depends just on $\bar{\mu}$ we generated 200 different inputs $\boldsymbol{x} = \langle x_1, \ldots, x_n \rangle$ such that $\sum_{i=1}^n w_i x_i$ always was equal to $\bar{\mu}^{(0)} = \sum_{i=1}^n w_i x_i^{(0)} = 24$ and measured the standard deviation λ_t of the resulting output y, which describes the amount of the total noise. Panel B shows the ratio λ_t/λ_0 between the total noise and the stochastic fluctuations.

(cf. Figure 2) will be addressed in the next subsection. In the following the term *systematic noise* refers to the fact that the output y in space-rate code does not merely depend on a single variable $\bar{\mu} = \sum_{i=1}^n w_i x_i$ but is actually a function which depends in a more complex way on the input $\langle x_1, \ldots, x_n \rangle$.

2.3 Analysis of the Systematic Noise

At the end of the preceding subsection we had pointed out that the output y of our model for computing in space-rate coding depends not just on the weighted sum $\sum_{i=1}^n w_i x_i$ of the inputs $x_1, \ldots, x_n$. In this subsection we provide tools for analyzing on what other quantities the output y may depend, and we exhibit a specific type of synaptic connectivity that minimizes the impact of this type of systematic noise.

Equation (1) shows that if one wants to approximate the firing probability $\mathrm{P}\{h_v \geq \theta\}$ of an arbitrary neuron v in pool V by a function of the weighted sum μ_v of the firing probabilities x_i in pools U_i, it is indeed *necessary* that σ_v also can be approximated by a function of μ_v. This follows easily from the equation $\Phi(\theta;\, \mu_v, \sigma_v) = \Phi((\theta - \mu_v)/\sigma_v;\, 0, 1)$. In order to analyze the possibilities for approximating σ_v by a function of μ_v we restate σ_v^2 as

$$\sigma_v^2 = \sum_{i=1}^n \hat{w}_{vi} x_i - \sum_{i=1}^n x_i^2 \sum_{u \in U_i} r_{vu}^2 \bar{a}_{vu}^2 \quad \text{with} \quad \hat{w}_{vi} := \sum_{u \in U_i} r_{vu} \hat{a}_{vu} \ . \tag{4}$$

We will focus on scenarios where the second term $\sum_{i=1}^n x_i^2 \sum_{u \in U_i} r_{vu}^2 \bar{a}_{vu}^2$ of equation (4) can be neglected for sufficiently large N. This can be achieved if either the non-failure probabilities r_{vu} or the mean EPSP (IPSP) amplitudes $\bar{a}_{vu}$ scale such that the weights w_{vi} remain bounded for large N.[5] Under this assumption one can approximate σ_v^2 by some second order polynomial $A\mu_v^2 + B\mu_v + C$ in μ_v if the angle φ between the weight vector $\boldsymbol{w}_v = \langle w_{v1}, \ldots, w_{vn} \rangle$ and the "virtual" weight vector $\hat{\boldsymbol{w}}_v = \langle \hat{w}_{v1}, \ldots, \hat{w}_{vn} \rangle$ that is defined in equation (4) is rather small.[6]

One can easily show that σ_v^2 can be expressed *exactly* as a function of μ_v if and only if $\varphi = 0$ or $\varphi = \pi$. Hence it is worthwhile to analyze what type of synaptic connectivity structure can achieve this. Obviously $\varphi = 0$ ($\varphi = \pi$) holds if and only if there exists a constant $\gamma > 0$ ($\gamma < 0$) such that $\hat{w}_{vi} = \gamma w_{vi}$ for all $i \in \{1, \ldots, n\}$. Note that this implies that all pools U_i consist of excitatory ($\varphi = 0$) or inhibitory ($\varphi = \pi$) neurons, since $\hat{w}_{vi} \geq 0$ by definition of $\hat{w}_{vi}$. In the following we consider the case where all neurons are excitatory. The condition $\hat{w}_{vi} = \gamma w_{vi}$ for all $i \in \{1, \ldots, n\}$ with a common constant $\gamma > 0$ is quite difficult to achieve in the case of heterogeneous weights w_{vi} for different i. However, one scenario in which this can in principle be achieved even with heterogeneous nonnegative weights w_{vi} is that where there exists for each neuron $u \in \bigcup_{i=1}^n U_i$ just a single release site between u and v with a common stereotyped probability density function ϕ for the amplitudes a_{vu} of the EPSP's caused by a synaptic release. For such architecture the non-failure probabilities r_{vu} (equal to the release probability in that case) can be chosen *arbitrarily* in order to achieve a desired heterogeneous assignment of nonnegative weights w_{vi} for $i \in \{1, \ldots, n\}$, while the condition $\hat{w}_{vi} = \gamma w_{vi}$ for all $i \in \{1, \ldots, n\}$ will automatically be satisfied with a common constant $\gamma > 0$. This follows easily from the fact that for a common stereotyped probability density function ϕ one gets $w_{vi} = \bar{a} \sum_{u \in U_i} r_{vu}$ and $w_{vi} = \hat{a} \sum_{u \in U_i} r_{vu}$ with $\bar{a} = \int a\phi(a)da$ and $\hat{a} = \int a^2 \phi(a)da$.

A connection from a neuron $u \in U_i$ to a neuron $v \in V$ via a single release site seems also to be advantageous with regard to synaptic plas-

[5] The term $\sum_{i=1}^n x_i^2 \sum_{u \in U_i} r_{vu}^2 \bar{a}_{vu}^2$ converges to 0 for $N \to \infty$ if for example the empirical standard deviation of the terms $r_{vu}\bar{a}_{vu}$ over $u \in U_i$ does not grow much faster than the mean $w_{vi}/N = \sum_{u \in U_i} r_{vu}\bar{a}_{vu}/N$ for $N \to \infty$, for example if $\left(\frac{1}{N}\sum_{u \in U_i}(r_{vu}\bar{a}_{vu} - \frac{w_{vi}}{N})^2\right)^{1/2} \leq \frac{w_{vi}}{N}$. This appears to be a quite reasonable assumption, and it furthermore implies that $\sum_{u \in U_i} r_{vu}^2 \bar{a}_{vu}^2 \leq \frac{2}{N} \cdot w_{vi}^2$. Hence the second term $\sum_{i=1}^n x_i^2 \sum_{u \in U_i} r_{vu}^2 \bar{a}_{vu}^2$ in (4) converges to 0 for $N \to \infty$ if the "weights" w_{vi} remain bounded for $N \to \infty$.

[6] φ is defined by $\cos\varphi = (\boldsymbol{w}_v \cdot \hat{\boldsymbol{w}}_v)/(\|\boldsymbol{w}_v\| \cdot \|\hat{\boldsymbol{w}}_v\|)$.

ticity. In the case of a single release site it is a reasonable assumption that the non-failure probabilities r_{vu} (equal to the release probability in that case) and the distribution of the EPSP amplitudes ϕ_{vu} can be chosen independently.[7] This implies that one can approximate the effective weights w_{vi} by $\frac{1}{N}\left(\sum_{u\in U_i}\bar{a}_{vu}\right)\left(\sum_{u\in U_i} r_{vu}\right)$ and the "virtual weights" $\hat{w}_{vi}$ by $\frac{1}{N}\left(\sum_{u\in U_i}\hat{a}_{vu}\right)\left(\sum_{u\in U_i} r_{vu}\right)$. Thus the condition $\hat{w}_{vi} = \gamma w_{vi}$ for all $i \in \{1,\ldots,n\}$ reduces to

$$\frac{1}{N}\sum_{u\in U_i}\hat{a}_{vu} = \frac{\gamma}{N}\sum_{u\in U_i}\bar{a}_{vu} \text{ for all } i \in \{1,\ldots,n\} \ . \tag{5}$$

Assume that for a given set of effective and virtual weights the relation $\hat{w}_{vi} = \gamma w_{vi}$ holds for all $i \in \{1,\ldots,n\}$, and subsequently the weights w_{vi} are subject to change due to some plasticity mechanism. In order to maintain the property that $\hat{w}_{vi} = \gamma w_{vi}$ for all i, such a change is best implemented by changes in the non-failure probabilities r_{vu}, since (5) holds independently of their values. If the weight change would be implemented by changes in the mean EPSP amplitudes $\bar{a}_{vu}$ then the average of the second moments $\hat{a}_{vu}$ would have to change proportional to keep (5) valid. If one considers the well known relation $\hat{a}_{vu} = \bar{a}_{vu}^2 + \tilde{a}_{vu}^2$, where $\tilde{a}_{vu}^2$ is the variance of the distribution ϕ_{vu} ($\tilde{a}_{vu}^2 = \int \phi_{vu}(z)(z-\bar{a}_{vu})^2\,dz$), it becomes even more clear that such a change would involve complex changes in the distributions ϕ_{vu}. Hence in the case of connections with single release sites it is advantageous to implement weight changes through changes in the non-failure probabilities rather than through changes in the distribution of the EPSP amplitudes.

Thus our preceding analysis exhibits a possible computational advantage of synaptic connections consisting of single release sites. On first sight such connectivity structure appears to be quite undesirable because of the high unreliability of such synaptic connections. However our preceding arguments shows that such connectivity structure supports precise analog computations in space-rate coding better than multiple release sites. Together with equation (5) it induces an angle $\varphi = 0$ between the vectors $\langle w_{v1},\ldots,w_{vn}\rangle$ and $\langle \hat{w}_{v1},\ldots,\hat{w}_{vn}\rangle$, thereby allowing that the resulting firing activity y in pool V only depends on $\mu_v = \sum_{i=1}^n w_{vi}x_i$ for arbitrary values $x_1,\ldots,x_n \in [0,1]$ of presynaptic pool activities (see Figure 3).

Now we consider the case where the angle φ between the weight vector $\boldsymbol{w}_v = \langle w_{v1},\ldots,w_{vn}\rangle$ and the "virtual" weight vector $\hat{\boldsymbol{w}}_v = \langle \hat{w}_{v1},\ldots,\hat{w}_{vn}\rangle$ is rather small but $\neq 0$. We will show that in this case we can still approximate σ_v^2 well by the expression $B_0\mu_v + C_0$ for $B_0 := \bar{\boldsymbol{w}}_v \cdot \hat{\boldsymbol{w}}/\|\bar{\boldsymbol{w}}_v\|^2$ and $C_0 := \frac{1}{2}\sum_{i=1}^n(\hat{w}_{vi} - B_0 w_{vi})$. In fact, this specific linear expression in μ_v provides an approximation to σ_v^2 that is in a heuristic sense optimal among all second order

[7] For synapses with more than one release site the assumption that r_{vu} can be chosen independently from ϕ_{vu} does not hold (see the Appendix for an example).

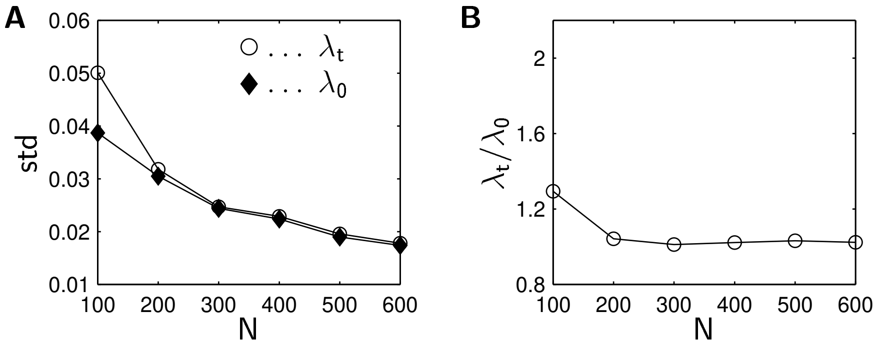

Figure 3: Quantitative study of the noise underlying the computation in our model if all connections just consist of a single release site. The model is constructed in a similar way as described in Figure 1, with $\langle w_1, \ldots, w_6 \rangle = \langle 10, 20, 30, 40, 50, 60 \rangle$ for the effective weights. The simulation protocol is the same as described in the caption of Figure 2. As you can see in panel B the ratio between the stochastic fluctuations (λ_0) and the total noise (λ_t) has a value of about 1 for $N \geq 200$. This shows that the output y of the network depends just on the weighted sum $\bar{\mu} = \sum_{i=1}^n w_i x_i$.

polynomials $A\mu_v^2 + B\mu_v + C$ in μ_v. To see this, we choose the parameters A, B and C such that the average error $E = \int_{[0,1]^n} \left(\sigma_v^2 - (A\mu_v^2 + B\mu_v + C)\right)^2 \,\mathrm{d}\boldsymbol{x}$ of such approximation is minimized. As shown in the Appendix the conditions $\frac{\delta E}{\delta A} = 0, \frac{\delta E}{\delta B} = 0$, and $\frac{\delta E}{\delta C} = 0$ imply for $N \to \infty$ that $A = 0$, $B = B_0$ and $C = C_0$. For these values of A, B and C the average error E has the minimum value $E_{min} = \int_{[0,1]^n} (\sigma_v^2 - (B_0\mu_v + C_0))^2 \mathrm{d}\boldsymbol{x} = \|\hat{\boldsymbol{w}}_v\|^2(1 - \cos^2\varphi)/12$. E_{min} decreases with decreasing φ, especially $E_{min} = 0$ if $\varphi = 0$. This observation indicates that one can approximate the firing probability $\mathrm{P}\{h_v \geq \theta\}$ of neuron v in pool V by the sigmoidal function $1 - \Phi\big(\theta;\, \mu_v, (B_0\mu_v + C_0)^{1/2}\big)$ of the single variable μ_v if the angle φ is not too large. Furthermore if we assume that the values w_{vi} and $\hat{w}_{vi}$ do not differ too much for different neurons $v \in V$ we can approximate the output y of the network by

$$y \approx 1 - \Phi\big(\theta;\, \bar{\mu}, (B_0\bar{\mu} + C_0)^{1/2}\big) \tag{6}$$

for $\bar{\mu} = \sum_{i=1}^n w_i x_i$ and $w_i = (1/N)\sum_{v \in V} w_{vi}$. This theoretical prediction is supported by our computer simulations: The function of $\bar{\mu} = \sum_{i=1}^n w_i x_i$ that appears on the right hand side of (6) is drawn as a solid line in Figure 1A.

Complications arise if some of the weights w_{vi} are positive and some of the weights w_{vi} are negative. It is obvious that in this case it is impossible to satisfy the condition $\hat{w}_{vi} = \gamma w_{vi}$ with a common constant $\gamma \neq 0$ for all $i \in \{1, \ldots, n\}$. Hence σ_v^2 depends not only on μ_v and thus also y depends not only on μ_v. This indicates that the systematic noise is larger in the case where one considers a combination of inhibitory and excitatory input. However, the simulations reported in Figure 1 show that even in this case the output y of the network approximates a sigmoidal function of $\bar{\mu} = \sum_{i=1}^n w_i x_i$.

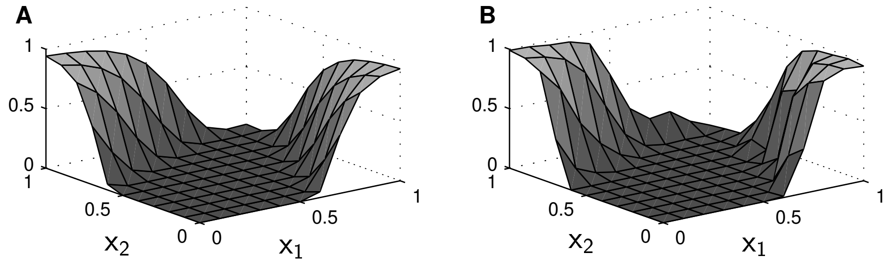

Figure 4: **A**) Plot of a function $f(x_1, x_2) : [0,1]^2 \rightarrow [0,1]$ which interpolates XOR. Since f interpolates XOR it cannot be computed by a single sigmoid unit. **B**) Computation of f by a 3-layer network in space-rate coding with spike response model neurons ($N = 200$) according to our model (for details see the Appendix).

2.4 Multi-Layer Computations

The preceding arguments show that approximate computations of functions of the form $\langle x_1, \ldots, x_n \rangle \rightarrow y = \sigma(\sum_{i=1}^{n} w_i x_i)$, with inputs and output in space-rate code, can be carried out within 10 ms by a network of spiking neurons. Hence the universal approximation theorem for multi layer perceptrons implies that *arbitrary continuous functions* $f : [0,1]^n \rightarrow [0,1]^m$ can be approximated with a computation time of not more than 20 ms by a network of spiking neurons with 3 layers. Thus our model provides a possible theoretical explanation for the empirically observed very fast multi-layer computations in biological neural systems that were mentioned in the introduction.

Results of computer simulations of the computation of a specific function f in space-rate coding that requires a multi-layer network because it interpolates the boolean function XOR is shown in Figure 4.

2.5 Consequences for Synfire Chains

On the side we would like to point out that our computational model also throws new light on the familiar model of a synfire chain Abeles (1991). There one considers a longer chain of pools of neurons, with large diverging/converging connectivity between adjacent pools in the chain. Abeles et al. (1993) had pointed out that this architecture has the property that relatively synchronous firing in the first pool triggers even more synchronous firing in the subsequent pools of the chain. We have shown in this article that if one takes synaptic unreliability into account, a synchronous firing of a certain fraction x of neurons in the first pool will cause a certain fraction y of neurons in the subsequent pools to fire synchronously. Whereas Abeles (1991) considered only the situation where x and y are close to 1, our analysis preceding shows that more subtle information processing can be carried out by a synfire chain: One can arrange that y is a smooth function of x, for a wide range of values for x. Such "graded activation" of synfire chains allows substantially more complex computations in networks of linked and/or reverberating synfire chains, since it supports the implementation of "graded pointers" from one synfire chain to other ones.

The original version and our analog version of a synfire chain (based on unreliable synapses) make slightly different predictions regarding the firing behavior of individual neurons in the synfire chain. In the analog version of a synfire chain only a certain fraction y of neurons will fire in a pool V of a chain, where y may assume arbitrary values between 0 and 1. Furthermore for repeated activations of the synfire chain with the same firing activity x in the first pool the actual set of neurons in V that fire will change from trial to trial. Hence precisely timed firing patterns among neurons in different pools of the chain would not occur *every* time when the first pool is activated with the same x, but just with a certain probability that is somewhat higher than for completely randomly firing neurons. This prediction appears to be consistent with published data Abeles et al. (1993).

3 Analog Computation on Time Series in a Space-Rate Code

We have shown in section 2 that biological neural systems with space-rate coding have at least the computational power of multi layer perceptrons. In this section we will demonstrate that they have strictly more computational power. This becomes clear if one considers computations on *time series*, rather than on *static batch inputs* as in section 2.

We now analyze the behavior of our computational model if the firing probabilities in the pools U_i change with time. Writing $x_i(t)$ ($y(t)$) for the probability that a neuron in pool U_i (V) fires during the t-th time window of length Δ (e.g. for Δ in the range between 1 and 5 ms), our computational model from section 2 maps a vector of n analog time series $\{x_i(t)\}_{t\in\mathbb{N}}$ onto an output time series $\{y(t)\}_{t\in\mathbb{N}}$ (where ($\mathbb{N} = \{0, 1, 2, \dots\}$).

As an example consider a network which consists of one presynaptic pool U_1 connected to the output pool V with the same type of synapses as discussed in section 2. In addition there are *feedback connections* between individual neurons $v \in V$. The results of simulations reported in Figures 5 and 6 show that this network computes an interesting map in the time series domain: The space-rate code in pool V represents a sigmoidal function σ (like in section 2) applied to the output of a *bandpass filter*. Figure 5B shows the response of such network to a sine wave with some bursts of activity added on top (Figure 5A).

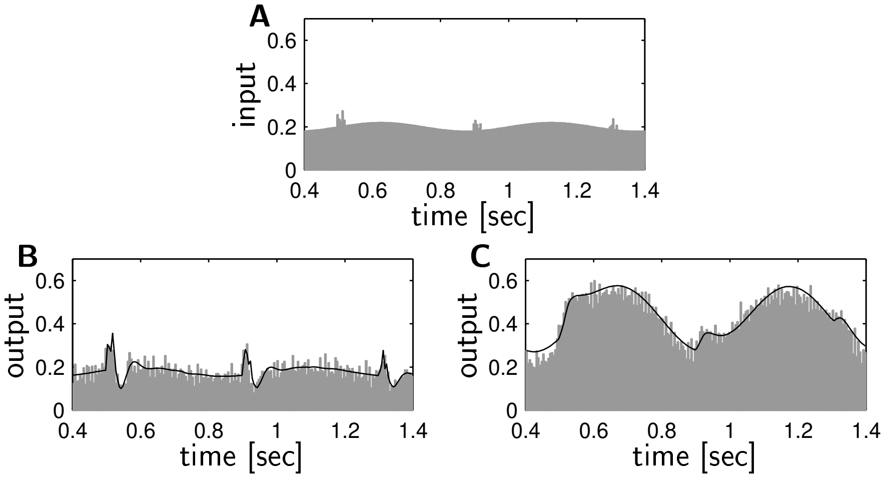

Figure 5: Response of two different networks to the same input (panel A). Panel B shows the response of a network which approximates a bandpass filter , whereas panel C shows the response of a network which approximates a lowpass filter. The gray shaded bars in panels B and C show the actual measured fraction of neurons which fire in pool V of the network during a time interval of length 5 ms in response to the input activity in pool U_1 shown in panel A. The solid lines in panels B and C are plots of the output predicted by equation (10). For details about the parameters see the Appendix.

Figures 6A and 6B show the frequency response of the bandpass filter which is implemented by this network of spiking neurons (see the appendix for the definition of the frequency response). Figure 5C shows the output of another network of spiking neurons (which approximates a *lowpass* filter) to the same input (shown in Figure 5A). The theoretical analysis of these networks will be presented in subsection 3.3.

To analyze theoretically the computational power of such networks of spiking neurons in the time-series domain we use the spike-response model Gerstner (1999b). We will focus on the situation where one can neglect the nonlinear effects of the sigmoidal function σ, i.e. we consider time series of the form $x_i(t) = x_0 + \tilde{x}_i(t)$ and $y(t) = y_0 + \tilde{y}(t)$ where the magnitudes of the signals $\tilde{x}_i(t)$ and $\tilde{y}(t)$ are small enough so that σ can be approximated in this range by a linear function $x \mapsto Kx$.

3.1 Taking the Time Course of Postsynaptic Potentials into Account

We model the effect of a firing of a neuron $u \in U_i$ at time k on the membrane potential of a neuron $v \in V$ at time t by a "response function" $\varepsilon_i(t-k)$, as in the spike-response model Gerstner (1999b). Such response function has the typical shape of an EPSP or IPSP. It is usually described mathematically as a difference of two exponentially decaying functions (see the Appendix). If one takes this time course of EPSP's and IPSP's into account, but ignores the impact of v's firing on its own membrane potential, our arguments from section 2 imply that one can write the expected value $\mu(t)$ of the membrane potentials of neurons $v \in V$ at time t as

$$\mu(t) = \sum_{i=1}^{n} w_i \sum_{k=0}^{t} \varepsilon_i(k) x_i(t-k) . \tag{7}$$

The "weights" w_i result from the release probabilities and distributions of PSP-amplitudes as described in section 2. We will ignore the impact of synaptic dynamics and assume that these weights w_i do not depend on t.

3.2 Taking Refractory Effects into Account

When the resulting firing probabilities in pool V become sufficiently large, the refractory effects of neurons $v \in V$ start to affect the resulting computation on time series. We then have to replace (7) by the equation

$$\mu(t) = \sum_{i=1}^{n} w_i \sum_{k=0}^{t} \varepsilon_i(k) x_i(t-k) + \sum_{k=1}^{t} \eta(k) y(t-k) \ , \tag{8}$$

where $\eta(k)$ typically assumes a very large negative value for a few ms, and then decays exponentially with a relatively small time constant Gerstner (1999b). This yields a special case of an *infinite impulse response* time-invariant linear filter, as we will show in the next subsection in a more general context. Since different specific firing properties of specific neurons, such as adaptation or rhythmic firing, correspond in this spike response model to different refractory functions $\eta(k)$ of specific neurons, a rich class of different infinite impulse response filters may arise in this way in biological neural systems.

3.3 Adding Recurrent Connections in Pool V

We now assume that in addition to the model defined in section 2 there are m different kinds of recurrent connections among neurons in pool V, with different response functions $\rho_j(k)$ and weights $\tilde{w}_j$, $j = 1, \ldots, m$. We assume that the large scale structure of these recurrent connections is of the same type as those between pools U_i and V, i.e. we assume that these are connections with unreliable synapses. Therefore the weights $\tilde{w}_j$ also result from release probabilities and PSP amplitude distributions as described in section 2. We then arrive at the following equation for the average membrane potential of neurons in pool V:

$$\mu(t) = \sum_{i=1}^{n} w_i \sum_{k=0}^{t} \varepsilon_i(k) x_i(t-k) + \sum_{k=1}^{t} \eta(k) y(t-k) + \sum_{j=1}^{m} \tilde{w}_j \sum_{k=1}^{t} \rho_j(k) y(t-k) \ .$$

If t is sufficiently large such that $\varepsilon_i(t') = \rho_j(t') = 0$ for $t' \geq t$ then for $x_i(t) = x_0 + \tilde{x}_i(t)$ and $y(t) = y_0 + \tilde{y}(t)$ one can rewrite $\mu(t)$ as $\mu(t) = \mu_0 + \tilde{\mu}(t)$ with

$$\mu_0 = x_0 \sum_{i=1}^{n} w_i \sum_{k=0}^{\infty} \varepsilon_i(k) + y_0 \sum_{k=1}^{\infty} \eta(k) + y_0 \sum_{j=1}^{m} \tilde{w}_j \sum_{k=1}^{\infty} \rho_j(k) \ \text{ and}$$

$$\tilde{\mu}(t) = \sum_{i=1}^{n} w_i \sum_{k=0}^{t} \varepsilon_i(k) \tilde{x}_i(t-k) + \sum_{k=1}^{t} \eta(k) \tilde{y}(t-k) + \sum_{j=1}^{m} \tilde{w}_j \sum_{k=1}^{t} \rho_j(k) \tilde{y}(t-k) \ .$$

We have shown in section 2 that under certain conditions one can approximate the firing activity $y(t)$ in pool V by a function of the form $\sigma(\mu(t))$, for some suitable sigmoidal function σ (see equation (6)). Since we consider only small signals $\tilde{x}_i(t)$ and $\tilde{y}(t)$ we can make the approximation $\tilde{y}(t) = K \cdot \tilde{\mu}(t)$ for some constant K which depends on the sigmoidal function σ. This yields for $\tilde{y}(t)$ the recursive equation

$$\begin{aligned} \tilde{y}(t) = K \cdot \sum_{i=1}^{n} w_i \sum_{k=0}^{t} \varepsilon_i(k) \tilde{x}_i(t-k) \\ + K \sum_{k=1}^{t} \eta(k) \tilde{y}(t-k) + K \sum_{j=1}^{m} \tilde{w}_j \sum_{k=1}^{t} \rho_j(k) \tilde{y}(t-k) \ . \end{aligned} \tag{9}$$

We will now show that equation (9) describes a very powerful computational model for computations on time series. This becomes clear if one just considers the special case where $x_i(t) = x(t)$ (hence $\tilde{x}_i(t) = \tilde{x}(t)$ for all $i \in \{1, \ldots, n\}$). We then have

$$\tilde{y}(t) = \sum_{k=0}^{t} b_k \tilde{x}(t-k) - \sum_{k=1}^{t} a_k \tilde{y}(t-k) \tag{10}$$

with $b_k = K \sum_{i=1}^{n} w_i \varepsilon_i(k)$ and $a_k = -K\eta(k) - K \sum_{j=1}^{m} \tilde{w}_j \rho_j(k)$. For arbitrary real valued numbers b_k and a_k for $k \in \mathbb{N}$ with $b_k = a_k = 0$ for $k > k_0$ (where k_0 is some sufficiently large constant)[8] this is is the general form of an *infinite impulse response (IIR)* time-invariant linear filter, see Haykin (1996); Back and Tsoi (1991). Furthermore if each response function $\varepsilon_i(\cdot)$ and $\rho_j(\cdot)$ is represented as a difference of two exponentially decaying functions and $\eta(\cdot)$

[8]Since in biological systems the responses to individual spikes vanish after some time is is a reasonable approximation to set $\varepsilon_i(k) = \rho_j(k) = \eta(k) = 0$ for $k > k_0$.

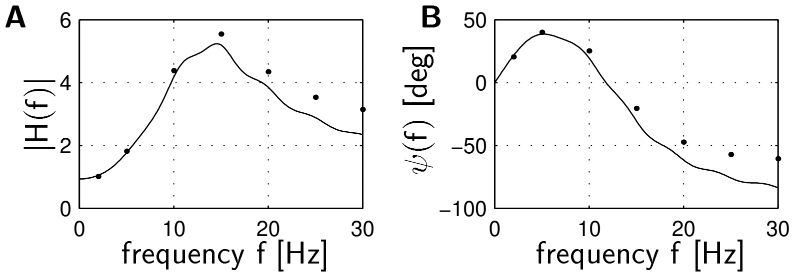

Figure 6: Amplitude response $|H(f)|$ (panel A) and phase response $\psi(f)$ (panel B) of the filter (bandpass) implicitly implemented by the network of spiking neurons described at the beginning of section 3 (see the Appendix for the definition of $|H(f)|$, $\psi(f)$, and the parameters of the network). Solid lines are plots of the amplitude and phase response predicted by the theory presented in subsection 3.3. Dots are simulation results for the approximating network of spiking neurons (see the Appendix for details).

is also represented as an exponentially decaying function, then the kernels $K\sum_{i=1}^{n} w_i \varepsilon_i(k)$ and $-K\eta(k) - K\sum_{j=1}^{m} \tilde{w}_j \rho_j(k)$ in equation (10) are weighted sums of exponentially decaying functions of the form $e^{-k/\tau}$ for various different values of τ. In that case one can easily show that the map from arbitrary input functions $\tilde{x}(t)$ to the output $\tilde{y}(t)$ can in principle approximate any given time-invariant linear filter with infinite impulse response (IIR) for any finite number t of discrete time steps.[9] In the case where $a_k = 0$ for all $k \in \mathbb{N}$, i.e. $\tilde{y}(t) = \sum_{k=0}^{t} b_k x(t-k)$, the resulting mapping from $\tilde{x}(t)$ to $\tilde{y}(t)$ is a time-invariant linear filter with *finite impulse response (FIR)*. Hence if there are no recurrent connections and one can neglect the refractory effects our model can still approximate any given FIR filter.

One can build a large class of practically relevant filters with the help of such FIR and IIR filters, including good approximations to lowpass and bandpass filters (Figure 6). Note that different architectures of recurrent circuits with different delays and different types of neurons involved in the recurrent loop give rise to a large variety of different coefficient vectors $\langle a_1, \ldots, a_{k_0}\rangle$. The variety of different vectors $\langle b_0, \ldots, b_{k_0}\rangle$ appears to be more limited in comparison, since time courses of PSP's in isolated neurons are relatively uniform. However, it is shown in Bernander et al. (1991) that the level of background

[9]For a rigorous proof of that result one just needs to observe that any function $g : \{0, \ldots, k_0\} \to \mathbb{R}$ can be approximated arbitrarily closely by functions of the form $f(k) = \sum_{i=1}^{n} w_i \varepsilon_i(k)$, where each function $\varepsilon_i(\cdot)$ is a difference of two exponentially decaying functions. This approximation result relies on two facts: the fact that any continuous function can be approximated arbitrarily closely by a polynomial on any bounded interval (Approximation Theorem of Weierstrass), and the fact that the variable transformation $u = -\log x$ for $u \in [0, \infty)$ transforms an exponentially decaying function $e^{-u/\tau}$ over $[0, \infty)$ into the function $x^{1/\tau}$ over $(0, 1]$. It becomes clear from this proof that if suffices to consider just time constants $\tau \leq 1$ for that purpose.

activity in cortical neurons may change the time constants of PSP's considerably. Hence such mechanism supports the implementations of a large variety of vectors $\langle b_0, \ldots, b_{k_0} \rangle$.

The preceding analysis implies that assuming that one can approximate any given real-valued sequences b_k and a_k on any finite interval, the output of the network of spiking neurons can approximate in space-rate coding on any finite time interval the saturated version of any given linear time invariant filter with finite and infinite impulse response. One may view this result as an *Universal Approximation Theorem* for linear filters by networks of spiking neurons.

Back and Tsoi (1991) have shown that an artificial neural network consisting of IIR filters as "synapses" and two layers of sigmoidal gates as "neurons" can be adjusted (via gradient descent for the parameters involved) to approximate for low pass filtered white noise as input in its output a very complex given time series. The combination of our results from section 2 and the results of this section shows that in principle an arbitrary network of the type considered by Back and Tsoi (1991) can be implemented by a network of spiking neurons in space-rate coding. Note however that in this implementation each IIR filter is implemented not by a single synapse but by a pool of spiking neurons.

An interesting aspect of this analysis is that these computations in the time series domain with space-rate coding make essential use of two aspects of neuronal dynamics that had received little attention in preceding computational models for neural systems: the specific forms of the refractory functions η and of the *decaying* parts of PSP's.

4 Hebbian Learning for Space-Rate Coding

The model for space-rate coding based on unreliable synapses that we had analyzed in section 2 has the characteristic property that the actual sets of neurons that fire will vary from trial to trial, even for the same input. In this section we show that this property is quite desirable from the point of view of learning. Most theoretically attractive learning rules for a computational unit in an artificial neural net, such as for example the Hebb rule, require a weight change Δw_i that depends on the *product* $x_i y$ of the ith analog input x_i and the analog output y of the unit. If these values x_i and y are encoded in a network of spiking neurons by a space-rate code, it is not clear how this product $x_i y$ can be "computed" locally by a biological synapse, or by the two spiking neurons that the synapse connects.

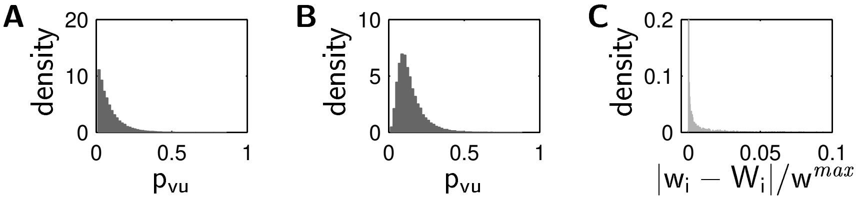

Figure 7: **A** and **B**: Typical change in the distribution of release probabilities at synapses between U_1 and V during learning as described in the text, (A) before and (B) after learning. **C**) Distribution of the normalized errors $|w_i^{new} - W_i^{new}|/|w^{max}|$ for all $i \in \{1, \ldots, 6\}$ where W_i^{new} is the ideal value of the updated weight according to the Hebb rule and $w_i^{new} = (1/N) \sum_{v \in V} \sum_{u \in U_i} r_{vu} \bar{a}_{vu}$ is the "effective weight" resulting from applications of Markram's rule in our simulated network of spiking neurons ($|w^{max}| = 100$ in this case; the mean is 0.0062). See the Appendix for details.

On the other hand Markram et al. (1997b) have shown that biological synapses regulate their efficacies in dependence of the *temporal relationship* between pre- and postsynaptic firing. For example a biological synapse may increase its release probability by an amount β if the postsynaptic neuron v fires within 100 ms after the presynaptic neuron u. When we apply this rule in the context of our computational model from section 2, we see that the fraction of synapses between neurons $u \in U_i$ and $v \in V$ so that v fires within 100 ms after the firing of u has an expected value of $x_i y$. Hence the expected fraction of synapses between U_i and V whose release probability is increased by β according to the rule from Markram et al. (1997b) is $x_i y$. This causes an average change in release probability between the pools U_i and V proportional to $x_i y$, although the value of the product $x_i y$ is *nowhere explicitly computed* in the network.

A closer look shows that this average change in release probability only has the desired effect of increasing the response y after a few repetitions under a special condition, namely if the actual sets of neurons in the pools that fire vary from trial to trial. Such trial-to-trial variability of the specific neurons that fire is inherent to our model for analog computation in space-rate coding presented in section 2, in contrast to previous models that ignore the large scale effects of synaptic unreliability.

We have compared the relationship between weight changes resulting from Hebb's learning rule for an artificial neural network and local changes of synaptic efficacy based on temporal relationship of pre- and postsynaptic firing according to Markram et al. (1997b) for our computational model from section 2 through computer simulations (see the Appendix). A typical change in the distribution of release probabilities between pools U_1 and V resulting from repeated applications of Markram's rule is shown in the Figures 7A and 7B. After 10 learning steps the weights $w_i := (1/N) \sum_{v \in V} \sum_{u \in U_i} r_{vu} \bar{a}_{vu}$ for all

$i \in \{1, \ldots, n\}$ resulting from Markram's rule differ on average by 0.0062 (normalized error, see Figure 7C) from the weights W_i resulting from applications of Hebb's rule for an artificial neural network with the corresponding parameter values and the same initial weights. For the application of Hebb's learning rule y was computed by $\sigma(\sum_{i=1}^{n} w_i x_i)$ for the sigmoidal function σ given by the right hand side of equation (6), with the same inputs x_i that are used as firing probabilities for the pools $U_1, \ldots, U_n$.

These results show that due to the large trial-to-trial variability of neuronal firing in our model from section 2 iterated local applications of Markram's temporal learning rule implement with high fidelity Hebb's learning rule for space-rate coding on the network level.

5 Discussion

We have addressed in this article specific links between details of realistic models for biologic neurons and synapses on one hand and resulting large scale effects for computations with populations of neurons on the other hand. In particular we have investigated possible macroscopic effects of the well-known stochastic nature of synaptic transmission. This aspect has been neglected in the previous analytical studies of the dynamics of populations of spiking neurons (see for example the references in the recent survey articles Gerstner (1999a,b)).

In section 2 we have shown that the unreliability of synaptic transmission suffices to explain the possibility of fast analog computations in space-rate coding on the network level. In fact, theoretically any given bounded continuous function can be approximated arbitrary closely by such networks, with a computation time of not more than 20 ms. In contrast to other possible explanations of analog computation in space-rate coding, this model spreads firing activity uniformly among the neurons, and therefore is able to explain the large trial-to-trial variability in neuronal firing patterns. Whereas it is occasionally conjectured that synaptic unreliability "averages out" on the network level, and hence causes no significant macroscopic effects, we have shown in section 2 through rigorous theoretical analysis and extensive computer simulations that synaptic unreliability may be essentially involved in generating a "graded response" for computations on the network level. Furthermore we have shown that most of these results are quite robust, since they hold for arbitrary assignments of synaptic failure probabilities and arbitrary distributions of PSP amplitudes.

We also have investigated a specific source of systematic noise for analog computation in space-rate coding that arises in this model. In contrast to sampling errors, this systematic noise is not likely to disappear when the pool size N is chosen very large. However our computer simulations suggest that the computational effect of this systematic noise is rather small. We have shown in subsection 2.3 that this source of systematic noise is removed if synaptic connections consist of single release sites and the "weights" of synaptic connections are encoded in release probabilities of synapses, rather than in the amplitudes of postsynaptic responses caused by the releases. On the side we have shown in subsection 2.5 that our computational model also gives rise to an *analog* version of the familiar model of a synfire chain. This analog version appears to be computationally more powerful than the classical binary version, and its firing behavior appears to be more consistent with experimental data.

In section 3 we have addressed the question what additional computational operations on the network level are supported by other prominent features of biological neurons and micro-circuits such as the refractory behavior of neurons and local recurrent connections. For that purpose we have moved from computations on static inputs towards an analysis of computations on time series, which is arguable a biologically particular relevant computational domain. Various filtering properties of biological neural systems have already been addressed in Koch (1999). We have exhibited in this article specific macroscopic effects for analog computations on time series in space-rate coding that are caused by unreliable synapses in combination with the decaying parts of postsynaptic potentials and refractory effects on the neuron level, and in combination with recurrent connections in micro-circuits. We have shown that through these features a rich repertoire of linear filters, especially arbitrary time-invariant linear filters with finite and infinite impulse response, can be approximated on the level of space-rate coding in populations of neurons. In combination with the results from section 2 this shows that for computations on time series in principle the full computational power of the artificial neural networks proposed in Back and Tsoi (1991) can be emulated by networks of biologically realistic neurons with space-rate coding.

In section 4 we have investigated macroscopic effects of a local learning rule for spiking neurons that is empirically supported by the experiments of Markram et al. (1997b). Whereas this learning rule is usually only discussed in the context of temporal coding in small neuronal circuits, we have shown that the same learning rule gives rise to a Hebbian learning rule on the network level with space-rate coding (hence without a coding of relevant information in firing times). This link between the Markram rule and Hebbian learning for space-rate coding appears to be obvious on first sight. A closer look shows however that it requires the large trial-to-trial variability of neuronal firing that is provided – in contrast to previous models – by the model for analog computation with space-rate coding based on unreliable synapses that was presented in section 2.

Acknowledgments

We would like to thank Peter Auer for helpful discussions and an anonymous referee for helpful comments.

Appendix

Synapse model. In all the simulations reported in this paper a connection between a neuron $u \in U_i$ and $v \in V$ is modeled by d_{vu} independent release sites, where d_{vu} may vary for different pairs u, v of neurons. The release probability p_{vu} is assumed to be equal at all d_{vu} release sites of one connection. In the case of a release at a single release site the amplitude of the resulting EPSP (IPSP), also known as quantal size, is drawn from a Gaussian distribution with mean $\bar{q}_{vu}$ and variance $\tilde{q}_{vu}$. For such a connection the probability density function ϕ_{vu} of EPSP (IPSP) amplitudes in the case of a release is given by (we skip the indices vu for sake of brevity)

$$\phi(a) = \frac{1}{r}\sum_{k=1}^{d}\binom{d}{k}p^k(1-p)^{d-k}\varphi\Big(a;\, k\bar{q}, \sqrt{k\tilde{q}^2}\Big) \quad ,$$

where $\varphi(\cdot;\, \bar{q}, \tilde{q})$ is the normal probability density function with mean $\bar{q}$ and variance $\tilde{q}$ and $r = 1-(1-p)^n$ is the probability that at least at one of the release sites some vesicle is released, i.e. r is the "non-failure probability". The mean $\bar{a} := \int a\phi(a)da$ and the second moment $\hat{a} := \int a^2\phi(a)da$ of such a distribution are given by $\bar{a} = \bar{q}dp$ and $\hat{a} = \frac{1}{r}\left(\bar{q}^2(dp^2(d-1)+dp)+\tilde{q}^2dp\right)$. Note that $d = 1$ yields $\bar{a} = \bar{q}$ and $\hat{a} = \bar{q}^2 + \tilde{q}^2$.

Neuron model. For all simulations we have used the spike response model as described in Gerstner (1999b). The time course of an EPSP (IPSP) is modeled by $\frac{a}{\tau_1-\tau_2}(e^{-t/\tau_1} - e^{-t/\tau_2})$ where τ_1 and τ_2 describes the rise and fall time of the PSP, and a defines the amplitude.[10] If not stated otherwise we have set $\tau_1 = 5\,\text{ms}$, $\tau_2 = 12\,\text{ms}$ for the EPSP's and $\tau_1 = 10\,\text{ms}$, $\tau_2 = 12\,\text{ms}$ for the slower IPSP's. The refractory behavior is modeled by the function $\eta(t) = -Re^{-t/\tau_r}$ where $\tau_r = 4\,\text{ms}$ was used for all simulations. R was chosen to be equal to the threshold of the neuron.

Synaptic parameters. The parameters d_{vu}, p_{vu}, $\bar{q}_{vu}$ and $\tilde{q}_{vu}$ for a connection between a neuron $u \in U_i$ and $v \in V$ are chosen independently from distributions reported in the literature. The number of release sites n_{vu} was chosen randomly from the set $\{1, 2, 3, 4, 5\}$ Markram et al. (1997a). The release probabilities p_{vu} are

[10]If the two time constants of these two exponentially decaying functions converge to a common value, then their difference converges to an "α-function", which is another common description of the time course of a PSP.

drawn from an exponential distribution Huang and Stevens (1997) with mean p_i (see Figure 1B for an example). The mean quantal sizes $\bar{q}_{vu}$ are drawn from a normal distribution with mean $\bar{q}_i$ and a variance of $0.1|\bar{q}_i|$. In accordance with the results reported in Auger et al. (1998) we have chosen the standard deviation of the quantal sizes $\tilde{q}_{vu} = 0.05\bar{q}_{vu}$. p_i and $\bar{q}_i$ were chosen to reflect different "effective weights" w_i between pool U_i and pool V as specified in the caption of Figure 1.

Second order approximation of σ_v^2. We define the average error E between σ_v^2 and the second order polynomial $A\mu_v^2 + B\mu_v + C$ as $E := \frac{1}{2}\int_{[0,1]^n}(\sigma_v^2 - (A\mu_v^2 + B\mu_v + C))^2\, \mathrm{d}\boldsymbol{x}$. We want to find the values A_0, B_0 and C_0 for the parameters A, B and C such that E assumes its minimum E_{min}. Therefore we simultaneously solve the equations $\frac{\delta E}{\delta A} = 0$, $\frac{\delta E}{\delta B} = 0$ and $\frac{\delta E}{\delta C} = 0$. After some calculations one gets

$$A_0 = -\frac{1}{N}\frac{\sum_i w_i^4}{\sum_i w_i^4 + (5/2)\sum_i\sum_{j\neq i} w_i w_j} \quad ,$$

$$B_0 = \frac{\sum_i \hat{w}_i w_i}{\sum_i w_i^2} - \frac{1}{N}\frac{\sum_i w_i^3}{\sum_i w_i^2} - \frac{A_0}{\sum_i w_i^2}\left(\sum_i w_i^3 + \sum_i\sum_{j\neq i} w_i^2 w_j\right), \text{ and}$$

$$C_0 = \frac{1}{2}\sum_i(\hat{w}_i - B_0 w_i) - \frac{1}{3N}\sum_i w_i^2 - \frac{A_0}{4}\sum_i\sum_{j\neq i} w_i w_j - \frac{A_0}{3}\sum_i w_i^2 \ .$$

In the limit $N \to \infty$ these equations reduce to $A_0 = 0$, $B_0 = \frac{\sum_i \hat{w}_i w_i}{\sum_i w_i^2}$ and $C_0 = \frac{1}{2}\sum_i(\hat{w}_i - Bw_i)$. These solutions lead to a minimal value of the average error E of $E_{min} = \frac{(\sum_i \hat{w}_i^2)(\sum_i w_i^2) - (\sum_i \hat{w}_i w_i)^2}{12\sum_i w_i^2}$. With the definitions $\bar{\boldsymbol{w}}_v = \langle w_{v1}, \ldots, w_{vn}\rangle$, $\hat{\boldsymbol{w}}_v = \langle \hat{w}_{v1}, \ldots, \hat{w}_{vn}\rangle$ and $\cos\varphi = (\boldsymbol{w}_v \cdot \hat{\boldsymbol{w}}_v)/(\|\boldsymbol{w}_v\| \cdot \|\hat{\boldsymbol{w}}_v\|)$ this reduces to $E_{min} = \frac{\|\hat{\boldsymbol{w}}\|^2}{12}(1 - \cos^2\varphi)$.

Frequency response of a time-invariant linear filter. The transfer function (in terms of the z-transformation Haykin (1996)) of a time-invariant linear filter that transforms a time series $x(t)$ into a time series $y(t) = \sum_{k=0}^{k_0} b_k x(t-k) - \sum_{k=1}^{k_0} a_k y(t-k)$ is given by $H(z) = (\sum_{k=0}^{k_0} b_k z^{-k})/(1 + \sum_{k=1}^{k_0} a_k z^{-k})$ where z is a complex variable. Setting $z = e^{j2\pi f}$ ($j = \sqrt{-1}$) we get the filter frequency response denoted by $H(e^{j2\pi f})$, where f denotes the frequency in Hertz. Expressing $H(e^{j2\pi f})$ in its polar form $|H(f)|e^{j\psi(f)}$, one defines the frequency response of the filter in terms of two components: the amplitude response $|H(f)|$ and the phase response $\psi(f)$. These two quantities have the following meaning: Obviously, if one applies a sine wave with frequency f at the input, the output of a linear filter is also a sine wave, but with different amplitude and phase. The amplitude response $|H(f)|$ is the ratio between the amplitudes of the output and the input sine wave. The phase response $\psi(f)$ is the difference between the phases of the output and the input sine waves. In that way one can measure for each frequency f the amplitude and phase response as it was done in the simulations reported in Figure 6.

Details to simulations reported in Figures 5 and 6. The network which approximates a lowpass filter (cf. Figure 5C) consists of a single excitatory presynaptic pool U_1 and the postsynaptic pool V. The effective weight between U_1 and V is $w_1 = 20$. For the time constants of the EPSP's we have chosen $\tau_1 = 25$ ms and $\tau_2 = 26$ ms. For the network which approximates a bandpass filter (cf. Figure 5B) there were in addition inhibitory recurrent connections from pool V to pool V with an effective weight $\tilde{w}_1 = -40$ ($\tau_1 = 30$ ms and $\tau_2 = 26$ ms). The effective weight from pool U_1 to pool V was $w_1 = 130$ ($\tau_1 = 10$ ms and $\tau_2 = 11$ ms).

Details to simulations reported in Figure 7. We performed 300 simulations for the model from section 2 for $n = 6, N = 200$ with different values of β and different sets of initial release probabilities p_{vu} (and hence different sets of initial weights $w_i = (1/N) \sum_{v \in V} \sum_{u \in U_i} r_{vu} \bar{a}_{vu}$). Throughout these simulations we used synapses with just a single release site. Each of these 300 simulations consisted of 10 individual trials with different inputs $\langle x_1^{(l)}, \ldots, x_6^{(l)} \rangle$, $l = 1, \ldots, 10$, where the inputs $x_i^{(l)}$ are drawn from normal distributions with mean $\bar{x}_i$ and variance $\hat{x}_i^2$ ($\bar{x}_i \in [0, 1]$ and $\hat{x}_i \in [0.1, 0.3]$ were chosen uniformly from the given interval for each simulation). In each trial we simulated a network of spiking neurons and applied locally Markram's rule. The release probability p_{vu} of a synapse between neuron $u \in U_i$ and $v \in V$ was increased by β if the postsynaptic neuron fired within 20 ms after the presynaptic neuron. After these 10 trials we compared the resulting effective weights $w_i^{new} = (1/N) \sum_{v \in V} \sum_{u \in U_i} r_{vu} \bar{a}_{vu}$ with the weights $W_i^{new} = W_i^{old} + N \beta \bar{a}_i \sum_{l=1}^{10} x_i^{(l)} y^{(l)}$ predicted by the Hebb rule, where $\bar{a}_i$ is the mean PSP amplitude resulting from spikes of neurons $u \in U_i$ and $y^{(l)}$ is given by $y^{(l)} = \sigma(\sum_{i=1}^{6} w_i x_i^{(l)})$ for the sigmoidal function given by the right hand side of equation (6).

References

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and gain control. *Science*, 275:220–224.

Abeles, M. (1991). *Corticonics: Neural Circuits of the Cerebral Cortex.* Cambridge University Press.

Abeles, M., Bergman, H., Margalit, E., and Vaadia, E. (1993). Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. of Neurophysiology*, 70(1):1629–1638.

Auger, C., Kondo, S., and Marty, A. (1998). Multivesicular release at single functional synaptic sites in cerebellar stellate and basket cells. *The Journal of Neuroscience: The official Journal of the Society for Neuroscience*, 18(12).

Back, A. D. and Tsoi, A. C. (1991). FIR and IIR synapses, a new neural network architecture for time series modeling. *Neural Computation*, 3:375–385.

Bernander, O., Douglas, R. J., Martin, K. A. C., and Koch, C. (1991). Synaptic background activity determines spatial-temporal integration in pyramidal cells. *Proceedings of the National Academy of Sciences*, 88:11569–11573.

Dobrunz, L. and Stevens, C. (1997). Heterogeneity of release probability, facilitation ad depletion at central synapses. *Neuron*, 18:995–1008.

Gerstner, W. (1999a). Populations of spiking neurons. In Maass, W. and Bishop, C., editors, *Pulsed Neural Networks*, pages 261–293. MIT-Press, Cambridge.

Gerstner, W. (1999b). Spiking neurons. In Maass, W. and Bishop, C., editors, *Pulsed Neural Networks*, pages 3–53. MIT-Press, Cambridge.

Haykin, S. (1996). *Adaptive Filter Theory.* Prentice-Hall, New Jersy.

Hopfield, J. J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature*, 367:33–36.

Huang, E. P. and Stevens, C. F. (1997). Estimating the distribution of synaptic reliabilities. *J. of Neurophysiology*, 78(6):2870–2880.

Koch, C. (1999). *Biophysics of Computation: Information Processing in Single Neurons.* Oxford University Press.

Larkman, A. U., Jack, J. J. B., and Stratford, K. J. (1997). Quantal analysis of excitatory synapses in rat hippocampal CA1 in vitro during low-frequency depression. *J. of Physiology.*

Maass, W. (1997). Fast sigmoidal networks via spiking neurons. *Neural Computation*, 9:279–304.

Markram, H., Lübke, J., Frotscher, M., Roth, A., and Sakmann, B. (1997a). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurons in the developing rat neocortex. *J. of Physiology*, 500(2):409–440.

Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997b). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275:213–215.

Petrov, V. V. (1995). *Limit Theorems of Probability Theory.* Oxford University Press.

Stevens, C. F. and Zador, A. (1998). Input synchrony and the irregular firing of cortical neurons. *Nature Neuroscience.* In press.

Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381:520–522.

Wilson, H. R. and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysics Journal*, 12:1–24.