12 Computing and Learning with Dynamic Synapses

Wolfgang Maass and Anthony M. Zador

12.1 Introduction

The models in all other chapters in this book assume that synapses are *static*, i.e., that they change their "weight" only on the slow time scale of learning. We will discuss in this chapter experimental data which show that this assumption is not justified for biological neural systems. As a matter of fact, this assumption is also unjustified for all hardware implementations of artificial neural nets where the sizes of synaptic "weights" are stored by analog techniques (see chapter 3). The consequences of this are threefold:

i) It is not clear whether implementations of pulsed neural nets in wetware or silicon are able to carry out computations in a way that is predicted by currently existing theoretical models for pulsed neural nets with static synapses.

ii) The inherent temporal dynamics of synaptic weights may not just be a curse, but also a blessing: dynamic synapses provide novel computational units for neural computation in the time series domain, i.e., in an environment where the inputs (and possibly also the outputs) of the networks are functions of time.

iii) One has to revise the foundations of learning in neural nets. In fact, it is not even clear anymore which are the parameters of networks of spiking neurons in which their "program" is stored. In particular all classical learning algorithms for neural nets that provide rules for tuning synaptic "weights" become dubious in this context. Furthermore, in view of point ii) it is not even clear what the goal of a learning algorithm for pulsed neural nets should be: the goal to learn a *function* (from input-vectors to output-vectors of numbers) or a *functional* (from input-functions to output-functions)?

In this chapter we will review in Section 12.2 results about the dynamic behaviour of biological synapses, we will survey in section 12.3 the available quantitative models for the temporal dynamics of biological synapses, and we will discuss possible computational uses of that dynamics in Section 12.4. Some consequences of synaptic dynamics for learning in neural nets will be discussed in Section 12.5.

12.2 Biological Data on Dynamic Synapses

In most models for networks of spiking neurons one assumes that the weight (or "efficacy") of a synapse is a parameter that changes only on the slow time scale of learning. It has however been known for quite some time (see for example [Katz, 1966, Magleby, 1987, Zucker, 1989] that synaptic efficacy changes with activity.

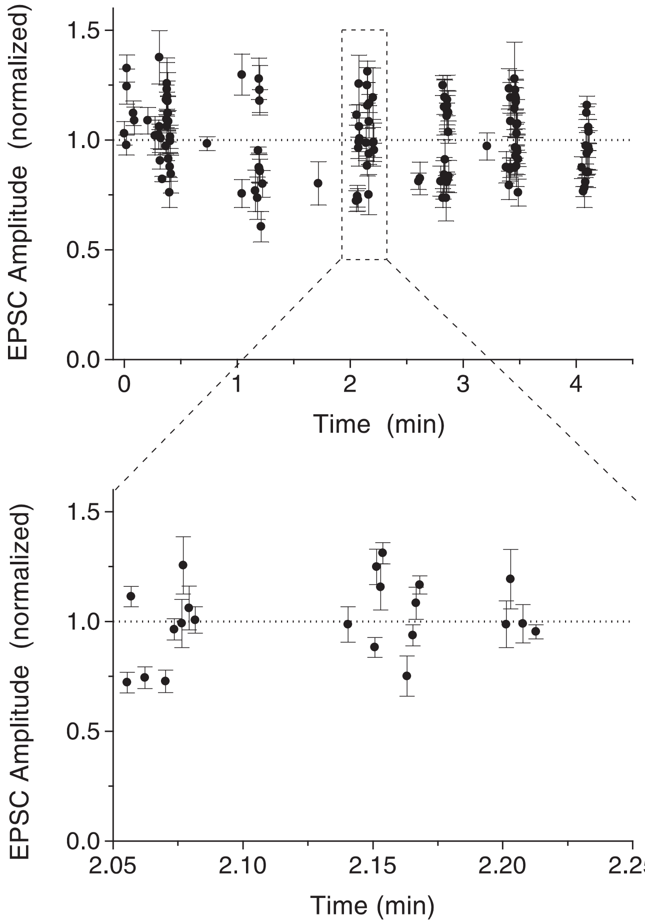

Figure 12.1. Synaptic response depends on the history of prior usage: Excitatory postsynaptic currents (EPSCs) recorded from a CA1 pyramidal neuron in a hippocampal slice in response to stimulation of the Schaffer collateral input. The stimulus is a spike train recorded *in vivo* from the hippocampus of an awake behaving rat, and "played-back" at a reduced speed *in vitro*. The presynaptic spikes have an average interspike interval of 1,950 msec that varies from a low at 35 msec to a maximum at 35 sec. The normalized strength of the EPSC varies in a deterministic manner depending on the prior usage of the synapse. For a constant synaptic weight, the normalized amplitudes should all fall on the dashed line. (A) EPSC as a function of time. The mean and standard deviation (4 repetitions) are shown. Note the response amplitude varies rapidly by more than two-fold. (B) AN excerpt is shown at a high temporal resolution. Unpublished data from L. E. Dobrunz and C. F. Stevens.

Phenomenon	Duration	Locus of Induction
Short-term Enhancement		
Paired-pulse facilitation (PPF)	100 $msec$	Pre
Augmentation	10 sec	Pre
Post-tetanic potentiation	1 min	Pre
Long-term Enhancement		
Short-term potentiation (STP)	15 min	Post
Long-term potentiation (LTP)	$>$ 30 min	Pre and post
Depression		
Paired-pulse depression (PPD)	100 $msec$	Pre
Depletion	10 sec	Pre
Long-term Depressions (LTD)	$>$ 30 min	Pre and post

Table 12.1. Different forms of synaptic plasticity
Synaptic plasticity occurs across many time scales. This table is a list of some of the better studied forms of plasticity. Included also are very approximate estimates of their associated decay constants, and whether the conditions required for induction depend on pre- or on postsynaptic activity, or on both. This distinction is crucial from a computational point of view, since Hebbian learning rules require a postsynaptic locus for the induction of plasticity. Note that for LTP and LTD, we are referring specifically to the form found at the Schaffer collateral input to neurons in the CA1 region of the rodent hippocampus; other forms have different requirements.

The responses shown in Figure 12.1 represent the complex interactions of many use-dependent forms of synaptic plasticity, many of which are listed in Table 12.2. Some involve an increase in synaptic efficacy (called "facilitation" or "enhancement"), while others involve a decrease ("depression"). They differ most strikingly in duration: some (e.g. facilitation) decay on the order of about 10 to 100 milliseconds, while others (e.g. long-term potentiation, or LTP) persist for hours, days or longer. The spectrum of time constants is in fact so broad that it covers essentially every time scale, from the fastest (that of synaptic transmission itself) to the slowest (developmental). The terms "paired-pulsed facilitation" and "paired-pulsed depression" in Table 12.2 refer to experiments where the stimulus consists of just two spikes, and the second spike causes a larger (smaller) postsynaptic response.

These forms of plasticity differ not only in time scale, but also in the conditions required for their induction. Some – particularly the shorter-lasting forms – depend only on the history of presynaptic stimulation, independent of the postsynaptic response. Thus facilitation, augmentation, and post-tetanic potentiation (PTP) occur after rapid presynaptic stimulation, with more vigorous stimulation leading to more persistent potentiation. Other forms of plasticity depend on some conjunction of pre- and postsynaptic activity: the most famous example is LTP, which obeys Hebb's rule in that its induction requires simultaneous pre- and postsynaptic activation.

In view of these data it becomes unclear which parameter should be referred to as the current weight or efficacy of a synapse, and what *learning* in a biological neural system means. For example [Markram and Tsodyks, 1996] have exhibited cases where pairing of pre- and postsynaptic firing (i.e., Hebbian learning) does not affect the average amplitudes of postsynaptic responses. Instead, it *redistributes* smaller and larger responses among the spikes of an incoming spike train (see Figure 12.2).

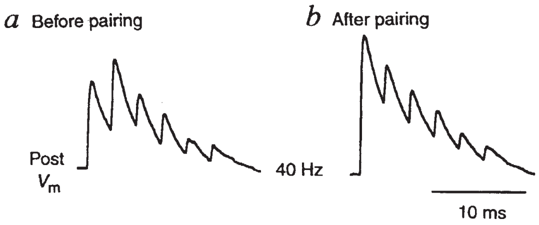

Figure 12.2. Postsynaptic responses to a regular 40 Hz spike train before and after Hebbian learning. From [Markram and Tsodyks, 1996].

In all of the abovementioned data the postsynaptic neuron was connected by several (typically about a half-dozen or more) synapses to the presynaptic neuron, and the recordings actually show the *superposition* of the responses of these multiple synapses for each spike of the presynaptic neuron. It is experimentally quite difficult to isolate the response of a *single* synapse, and data have become available just very recently [Dobrunz and Stevens, 1997]. The results are quite startling. Those single synapses (or more precisely: synaptic release sites) in the central nervous system that have been examined so far exhibit a *binary* response to each spike from the presynaptic neuron: either the synapse releases a single neurotransmitter-filled vesicle or it does not respond at all. In the case when a vesicle is released, its content enters the synaptic cleft and opens ion-channels in the postsynaptic membrane, thereby creating an electrical pulse in the postsynaptic neuron. It is shown in the lower panel of Figure 12.3 that the mean-size of this pulse in the postsynaptic neuron does not vary in a systematic manner for different spikes in a spike train from the presynaptic neuron. The *probability*, however, that a vesicle is released by a synapse varies in a systematic manner for different spikes in such spike train.

The stochastic nature of synaptic function is the basis of the *quantal hypothesis* [Katz, 1966], which states that vesicles of neurotransmitter are released in a probabilistic fashion following the invasion of the presynaptic terminal by an action potential. The basic quantal model, first developed to explain results at the neuromusular junction has been validated at many different synapses, including glutamatergic and gabaergic terminals in the mammalian cortex.

The history-dependence of synaptic release probability was first studied at the neuromuscular junction. Depending on experimental conditions, these synapses show some combination of facilitation (a use-dependent increase in release probability) and depression (a use-dependent decrease in release probability). Subsequent studies have revealed that use-dependent changes in presynaptic release probability underlie many forms of short-term plasticity, including some seen in the hippocampus and neocortex (reviewed in [Magleby, 1987, Zucker, 1989, Fisher et al., 1997, Zador and Dobrunz, 1997]). These changes occur on many different time scales, from seconds to hours or longer (reviewed in [Zador, 1998]).

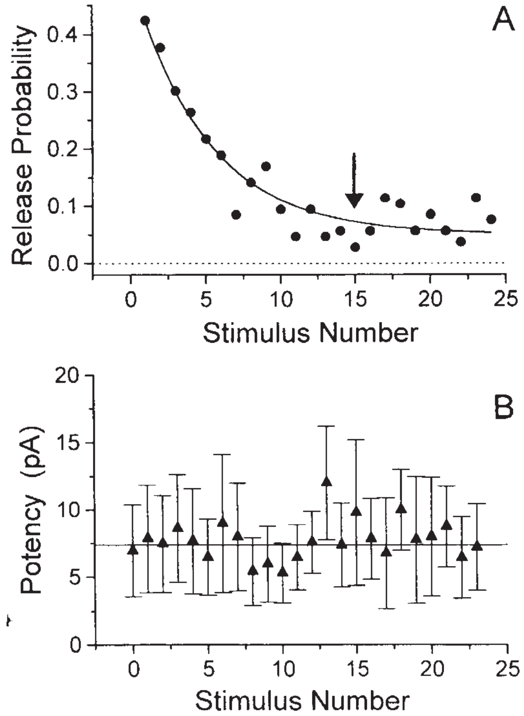

Figure 12.3. The upper panel shows the temporal evolution of release probabilities for a train of a 10 Hz spike train with regular interspike interval in a synapse from rat hippocampus. The lower panel shows the mean size and standard deviation of the amplitude of the postsynaptic pulses for the same spike train in those cases when a vesicles was released. From [Dobrunz and Stevens, 1997]

On first sight the binary response (release/failure of release) of a single synapse might appear to be inconsistent with the multitude of reproducible response sizes shown in Figure 12.1. However one should keep in mind that Figure 12.1 shows the superposition of responses from *many* synapses. Hence for each pulse the size of the postsynaptic response scales with the *number* of individual synapses that release a vesicle. For each spike the expected number of released vesicles equals the *sum of the release probabilities* at the individual synapses times the response at each synapse.

In this way a multi-synaptic connection between two neurons may be subject to an even richer dynamics and learning capability than the preceding models suggest, since there exists evidence that individual synapses have quite different temporal dynamics. While the heterogeneity of release properties at different release sites has long been suspected (e.g. [Atwood et al., 1978, Brown et al., 1976], only recently has direct evidence for such heterogeneity become available at central synapses.

Three main lines of evidence support this heterogeneity. First, the activity-dependent synaptic channel blocker MK-801 blocks synapses at different rates, as would be expected from a population of synapses with many different release probabilities [Hessler et al., 1993, Rosenmund et al., 1993, Castro-Alamancos and Connors, 1997, Manabe and Nicoll, 1994] Second, minimal stimulation [Allen and Stevens, 1994, Dobrunz and Stevens, 1997, Stratford et al., 1996] and paired recordings [Stratford et al., 1996, Markram and Tsodyks, 1996, Bolshakov and Siegelbaum, 1995] indicate that different connections have different properties. Finally, visualization of the vesicular marker FM-143 [Ryan et al., 1996, Murthy et al., 1997] indicates that release at different terminals is different.

12.3 Quantitative Models

A rather simple but very successful quantitative model for the size of the postsynaptic response caused by *multiple* synapses was proposed in [Varela et al. 1997]. The amplitude $A(t_i)$ of the postsynaptic pulse for the ith spike arriving at time t_i is modeled as a product

$$A(t_i) = A_0 \cdot F(t_i) \cdot D_1(t_i) \cdot D_2(t_i)$$

of a constant A_0 and three functions F, D_1, and D_2. The function F models the effect of facilitation: a fixed amount Δ is added to F for each presynaptic spike. Between spikes the value of F decays exponentially back to its initial value. The functions D_1 and D_2 model synaptic depression in a dual fashion: for each presynaptic spike the current value of D_i is multiplied by a factor $d_i \in (0, 1)$, and the value of D_i recovers exponentially (with some time constant τ_i) back to the initial value of $D_i;\ i = 1, 2$. It turns out that two terms D_1 and D_2 with efficient constants d_i and τ_i provide a substantially better fit to experimental data than just a single one. Sometimes even three terms are used. Figure 12.4 shows that with suitable choice of parameters this model can predict quite well the amplitude of multi-synaptic responses to a 4 Hz Poisson train. A somewhat related but more complex quantitative model for the response of multiple synapses was proposed in [Tsodyks and Markram, 1997] for the case of depressing synapses and extended in [Markram and Tsodyks, 1997] to allow also facilitation effects.

We had indicated already at the end of the preceding section that the actual dynamics of a synapse in the central nervous system of a biological organism differs from the preceding models since it is stochastic, and the synaptic response to each spike apparently ranges over just two discrete values (release/failure of release). Based on earlier quantitative models for partial aspects of synaptic dynamics (see for example [Dobrunz and Stevens, 1997]), a computational model for individual dynamic stochastic synapses was proposed in [Maass and Zador, 1998]. In this model a spike train is represented as in Chapter 1 by a sequence $\underline{t}$ of firing times, i.e. as an increasing sequence of numbers $t_1 < t_2 < \ldots$ from $\mathbf{R}^+ := \{z \in \mathbf{R} : z \geq 0\}$. For each spike train $\underline{t}$ the output of synapse S consists of the sequence $S(\underline{t})$ of those $t_i \in \underline{t}$ on which vesicles are "released" by S , i.e. of those $t_i \in \underline{t}$ which cause an excitatory or inhibitory postsynaptic

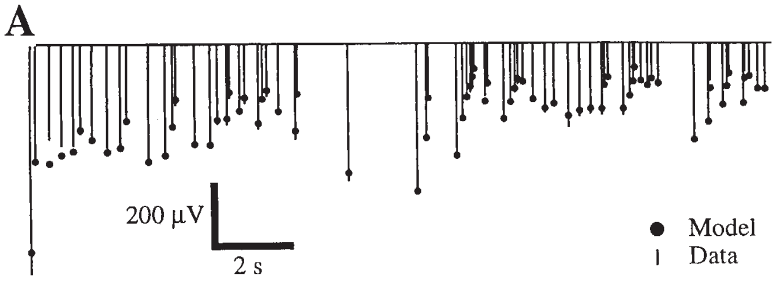

Figure 12.4. Measured and predicted amplitude (field potentials) of the postsynaptic response to a 4 Hz Poisson spike train. Predictions result from fit of model with three terms F, D_1, D_2. From [Varela et al. 1997]

potential (EPSP or IPSP, respectively). The map $\underline{t} \to S(\underline{t})$ may be viewed as a stochastic function that is *computed* by synapse S . Alternatively one can characterize the output $S(\underline{t})$ of a synapse S through its *release pattern* $\underline{q} = q_1 q_2 \ldots \in \{R, F\}^*$, where R stands for release and F for failure of release. For each $t_i \in \underline{t}$ one sets $q_i = R$ if $t_i \in S(\underline{t})$, and $q_i = F$ if $t_i \notin S(\underline{t})$. The central equation in this model gives the probability $p_S(t_i)$ that the i^{th} spike in a presynaptic spike train $\underline{t} = (t_1, \ldots, t_k)$ triggers the release of a vesicle at time t_i at synapse S,

$$p_S(t_i) = 1 - e^{-C(t_i) \cdot V(t_i)} \; . \tag{12.1}$$

The release probability is assumed to be nonzero only for $t \in \underline{t}$, so that releases occur only when a spike invades the presynaptic terminal (*i.e.* the spontaneous release probability is assumed to be zero). The functions $C(t) \geq 0$ and $V(t) \geq 0$ describe, respectively, the states of facilitation and depletion at the synapse at time t .

The dynamics of facilitation are given by

$$C(t) \; = \; C_0 + \sum_{t_i < t} c(t - t_i) \; , \tag{12.2}$$

where C_0 is some parameter ≥ 0 that can for example be related to the resting concentration of calcium in the synapse. The exponential response function $c(s)$ models the response of $C(t)$ to a presynaptic spike that had reached the synapse at time $t - s$: $c(s) = \alpha \cdot e^{-s/\tau_C}$, where the positive parameters τ_C and α give the decay constant and magnitude, respectively, of the response. The function C models in an abstract way internal synaptic processes underlying presynaptic facilitation, such as the concentration of calcium in the presynaptic terminal. The particular exponential form used for $c(s)$ could arise for example if presynaptic calcium dynamics were governed by a simple first order process.

The dynamics of depletion are given by

$$V(t) \; = \; \max(\, 0 \, , \; V_0 - \sum_{t_i :\; t_i < t \text{ and } t_i \in S(\underline{t})} v(t - t_i)) \; , \tag{12.3}$$

for some parameter $V_0 > 0$. $V(t)$ depends on the subset of those $t_i \in \underline{t}$ with $t_i < t$ on which vesicles were actually released by the synapse, i.e. $t_i \in S(\underline{t})$. The function $v(s)$ models the response of $V(t)$ to a preceding release of the same synapse at time $t - s \leq t$. Analogously as for $c(s)$ one may choose for $v(s)$ a function with exponential decay where $\tau_V > 0$ is the decay constant. The function V models in an abstract way internal synaptic processes that support presynaptic depression, such as depletion of the pool of readily releasable vesicles. In a more specific synapse model one could interpret V_0 as the maximal number of vesicles that can be stored in the readily releasable pool, and $V(t)$ as the expected number of vesicles in the readily releasable pool at time t.

In summary, the model of synaptic dynamics presented here is described by five parameters: C_0, V_0, τ_C, τ_V and α. The dynamics of a synaptic computation and its internal variables $C(t)$ and $V(t)$ are indicated in Figure 12.5.

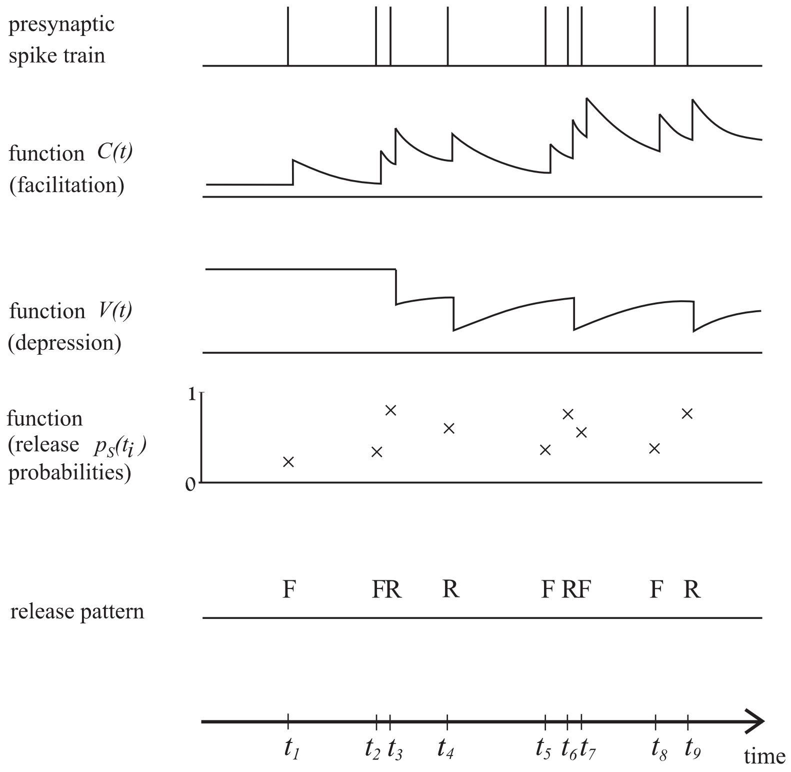

Figure 12.5. Synaptic computation on a spike train $\underline{t}$, together with the temporal dynamics of the internal variables C and V of our model. Note that $V(t)$ changes its value only when a presynaptic spike causes release.

This model for the dynamics of a single stochastic synapse is closely related to the previously discussed model for the combined response of multiple synapses by [Varela et al. 1997], since Eq. (12.1) can be expanded to first order around $r(t) := C(t) \cdot V(t) = 0$ to give

$$p_S(t_i) = C(t_i) \cdot V(t_i) + O([C(t_i) \cdot V(t_i)]^2). \qquad (12.4)$$

According to Eq. (12.2) the dynamics of $C(t_i)$ is quite similar to that of the facilitation term $F(t_i)$ in [Varela et al. 1997], and according to Eq. (12.3) the term $V(t_i)$ is closely related to their depression terms $D_j(t_i)$. This correspondence becomes even closer in variations of the previously described most basic model for an individual synapse that are discussed in [Maass and Zador, 1998].

In order to investigate macroscopic effects caused by stochastic dynamic synapses one can expand the previously described model to a mean field version that describes the mean response of a population of dynamic synapses that connect populations of neurons. In this approach [Zador et al., 1998] the input to a population of "parallel" synapses can be described by a continuous function $r(t)$ which represents the current firing activity in the preceding pool of neurons. The impact that this firing activity has on the next pool of neurons is described by a term $w \cdot p(t) \cdot r(t)$, where $p(t) = 1 - e^{C(t) \cdot V(t)}$ is a continuous function with values in $[0, 1]$ that represents the current mean release probability in the population of synapses that connect both pools of neurons. The dynamics of the auxiliary functions $C(t)$ and $V(t)$ is defined as in equation (12.2) and (12.3), but with an integral over the preceding time window and the input $r(t)$ instead of a sum over spikes. The latter may be viewed as a special case of such integral by integrating over a sum of δ-functions, as described in Chapter 1.

12.4 On the Computational Role of Dynamic Synapses

Quantitative models for biological dynamic synapses are relatively new, and the exploration of their possible computational use has just started. We will survey in this section some of the ideas that have emerged so far.

[Abbott et al., 1997] and [Tsodyks and Markram, 1997] point out that depression in dynamic synapses provides a dynamic gain control mechanism: a neuron i can detect whether a presynaptic neuron j suddenly increases its firing rate by a certain percentage – independently of the current firing rate of that presynaptic neuron j. Assume that some of the predecessors of neuron i fire at a high rate and others at a low rate. Then with *static* synapses the neuron i is rather insensitive to changes in the firing rates of slowly firing presynaptic neurons, since its membrane potential is dominated by the large number of EPSP's from rapidly firing presynaptic neurons. However according to the model by [Varela et al. 1997] described at the beginning of Section 3, one can achieve with multiple dynamic synapses that the amplitude of the EPSP's caused by a presynaptic neuron j with a firing rate r_j (and regular interspike intervals) scales like $1/r_j$. This implies that an increase in that firing rate by a fraction $p \cdot r_j$ causes an increase of the postsynaptic response by $\frac{p \cdot r_j}{r_j} = p$, independently of the current value of r_j. In this way the neuron becomes equally sensitive to changes by a percentage p in the firing rate of slowly firing and rapidly firing presynaptic neurons.

For realistic values of the parameters in the synapse model of [Varela et al. 1997] (that result from fitting this model to data from multiple synapses in slices of rat primary visual cortex) the previously described effect, whereby the amplitude of EPSP's from presynaptic neuron j scales like $1/r_j$, sets in for firing rates r_j above 10 Hz. A startling consequence of this effect is that the neuron i becomes insensitive to changes in the sustained firing rates r_j of presynaptic neurons j if these rates lie above 10 Hz. This implies that the traditional "non-spiking" model for biological neural computation, where biological neurons are modeled by sigmoidal neurons with inputs and outputs encoded by firing rates, becomes inapplicable for input firing rates above 10 Hz. On the other hand the typical membrane time constants of a biological neuron lie well below 100 msec, so hat this traditional model also becomes questionable for input firing rates below 10 Hz.

[Markram and Tsodyks, 1997] emphasize that different synapses in a neural circuit tend to have different dynamic features. In this way a spike train from a neuron whose axon makes on the order of 1000 synaptic contacts may convey different messages to the large number of postsynaptic neurons, with the synapses acting as filters that extract different special features from the spike trains. They also point out that in a circuit with excitatory and inhibitory neurons, where the response of the inhibitory neurons is delayed in a frequency-dependent manner via facilitating synapses, a frequency-dependent time window ($\sim 1/r^2$) is created for excitation to spread before inhibition is recruited.

Other possible computational uses of dynamic synapses can be derived from the synapse model of [Maass and Zador, 1998] described at the end of Section 3. The following result shows that by changing just two of the synaptic parameters of this model, a synapse S can choose virtually independently the release probabilities $p_S(t_1)$ and $p_S(t_2)$ for the first two spikes in a spike train.

Theorem 12.1 *Let $\langle t_1, t_2 \rangle$ be some arbitrary spike train consisting of two spikes, and let $p_1, p_2 \in (0,1)$ be some arbitrary given numbers with $p_2 > p_1 \cdot (1 - p_1)$. Furthermore assume that arbitrary positive values are given for the parameters α, τ_C, τ_V of a synapse S. Then one can always find values for the two parameters C_0 and V_0 of the synapse S so that $p_S(t_1) = p_1$ and $p_S(t_2) = p_2$.*

Furthermore the condition $p_2 > p_1 \cdot (1 - p_1)$ is necessary in a strong sense. If $p_2 \leq p_1 \cdot (1 - p_1)$ then no synapse S can achieve $p_S(t_1) = p_1$ and $p_S(t_2) = p_2$ for any spike train $\langle t_1, t_2 \rangle$ and for any values of its parameters $C_0, V_0, \tau_C, \tau_V, \alpha$. ■

One can use this result for a rigorous proof that a spiking neuron with dynamic synapses has more computational power than a spiking neuron with static synapses: Let T be a some given time window, and consider the computational task of detecting whether at least one of n presynaptic neurons $a_1, \ldots, a_n$ fire at least twice during T ("burst detection"). To make this task computationally feasible we assume that none of the neurons $a_1, \ldots, a_n$ fires outside of this time window.

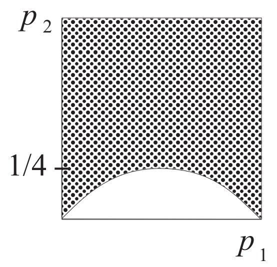

Figure 12.6. The dotted area indicates the range of pairs $\langle p_1, p_2 \rangle$ of release probabilities for the first and second spike through which a synapse can move (for any given interspike interval) by varying its parameters C_0 and V_0.

Theorem 12.2 *A single spiking neuron with dynamic stochastic synapses can solve this burst detection task (with arbitrarily high reliability). On the other hand no spiking neuron with static synapses can solve this task (for any assignment of "weights" to its synapses).* [1]

In order to show that a single spiking neuron with dynamic synapses can solve this burst detection task one just has to choose values for the parameters of its synapses S so that $p_S(t_1)$ is close to 1 and $p_S(t)$ is close to 0 for all $t \in [t_1, t_1 + T]$. We refer to [Maass and Zador, 1998] for a proof that a spiking neuron with static synapses cannot solve this burst detection task.

Another possible computational use of stochastic dynamic synapses is indicated by the following example. Two arbitrary Poisson spike trains A and B were chosen, that each consist of 10 spikes and hence represent the same firing rate.

Figure 12.7 compares the response of a synapse with fixed parameters to two spike trains A and B. The synaptic parameters were adjusted so that the average release probability (computed over the 10 spikes) was greater for A than for B. Figure 12.8 compares the response of a different synapse to the same two spike trains; in this case, the parameters were chosen so that the average response to B was greater than to A. These examples indicate that even in the context of rate coding, synaptic efficacy may not be well-described in terms of a single scalar parameter w. In the mean field version of this synapse model (described at the end of section 3) one can show that the internal synaptic parameters can be chosen in such a way that a population of synapses computes a function that either approximates an arbitrary given linear filter, or higher order terms in the Volterra-series expansion of a nonlinear filter.

Curiously enough the view of synapses as linear filters had already been explored before by [Back and Tsoi, 1991, Principe, 1994], and others in the context of artificial neural nets. They pointed to various potential computational uses of the resulting neural nets in the context of computations

[1] We assume here that neuronal transmission delays differ by less than $(n-1) \cdot T$, where by *transmission delay* we refer to the temporal delay between the firing of the presynaptic neuron and its effect on the postsynaptic target.

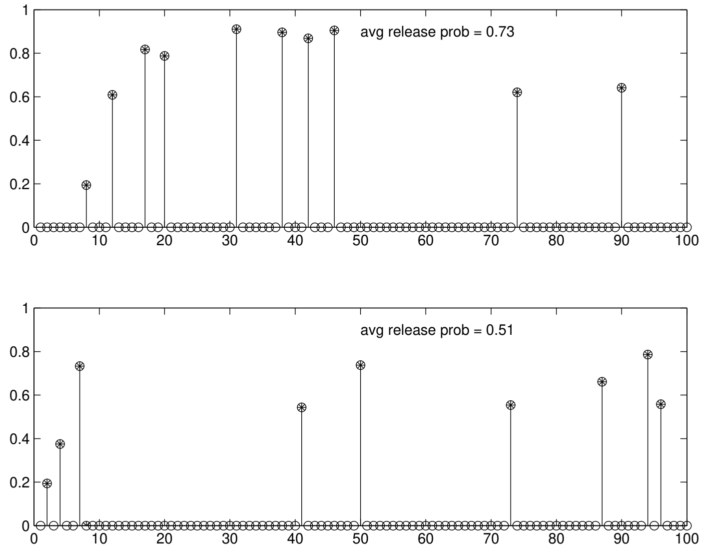

Figure 12.7. Example of a synapse whose average release probability is 22 % higher for spike train A, shown in the top panel. Release probabilities of the same synapse for spike train B are shown in the bottom panel. The fourth spike in spike train B has a release probability of nearly zero and so is not visible. The spike trains shown here are the same as in the next figure.

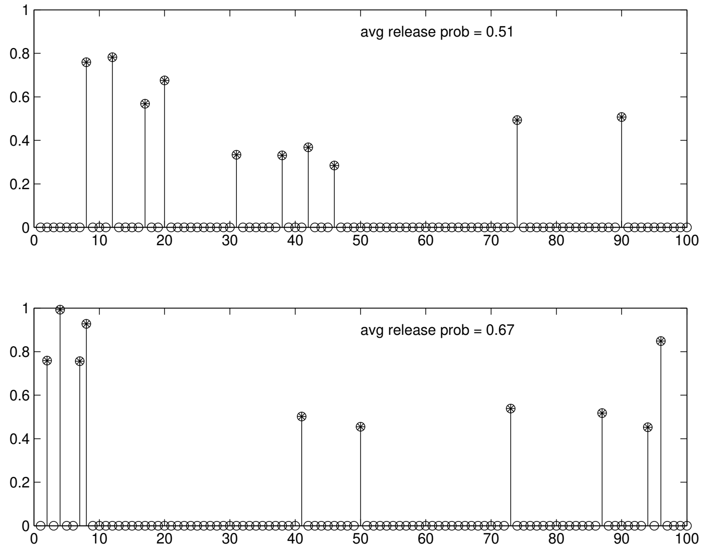

Figure 12.8. Example of a synapse whose average release probability is 16 % higher for spike train B, shown in bottom panel. Release probabilities of the same synapse for spike train A are shown again in the top panel. The spike trains shown here are the same as in the previous figure.

on time series without having at that time any indication that biological neural systems might be able to implement such sets.

12.5 Implications for Learning in Pulsed Neural Nets

The preceding experimental data and associated models show that it is quite problematic to view a biological synapse as a trivial computational unit that simply multiplies its input with a fixed scaler w, its synaptic "weight". Instead, a biological synapse should be viewed as a rather complex nonlinear dynamical system. It is known that the "hidden parameters" that regulate the dynamics of a biological synapse vary from synapse to synapse [Dobrunz and Stevens, 1997], and that at least some of these hidden parameters can be changed through LTP (i.e., through "learning"). This has drastic consequences for our view of learning in biological neural systems. Since synapses are history-dependent it does not suffice to consider training examples that consist of input- and output vectors of numbers. Instead a neural system learns to map certain input functions of time to given output functions of time. Hence a training example consists of a pair of *functions* or *time series*. Furthermore the learning algorithm itself has to specify not only rules for changing the scaling parameter w of a synapse, but also for changing the internal synaptic parameters that regulate the dynamic behavior of the synapse. Experimental evidence that Hebbian learning ("pairing") in biological neural system changes the *dynamic* behavior of a synapse, rather than just the average amplitude of its responses, was provided by [Markram and Tsodyks, 1996]. Their data (see Figure 12.2) suggest that learning may *redistribute* the strength of synaptic responses among the spikes in a train, rather than changing its average response. The particular change in synaptic dynamics that is shown in Figure 12.2 makes the postsynaptic neuron more sensitive to transients, i.e., to rapid changes in the firing rate of the presynaptic neuron. In this way synaptic plasticity may change the sensitivity of a neuron to different neural codes. In this examples it increases the sensitivity for temporal coding, while at the same time decreasing its limiting frequency above which it is unable to distinguish between different presynaptic firing rates. [Liaw and Berger, 1996] have shown that interesting changes of the dynamic properties of a neural circuit with dynamic synapses can already be achieved by just changing the scaling factors of excitatory and inhibitory connections, without changing the hidden parameters that control the dynamics of the synapses themselves.

Preliminary results from [Zador et al., 1998] show that by applying gradient descent learning also to the hidden parameters of a synapse, a neural circuit can in principle "learn" to realize quite general given operators that map input time series to given output time series. We employ here the mean field model for dynamic synapses (described at the end of section 12.3) in order to be able to consider arbitrary bounded time series as network-input and -output. This approach is related to the previous work by [Back and Tsoi, 1991], who have carried out gradient descent for the hidden parameters of linear filters that replace synapses in their model.

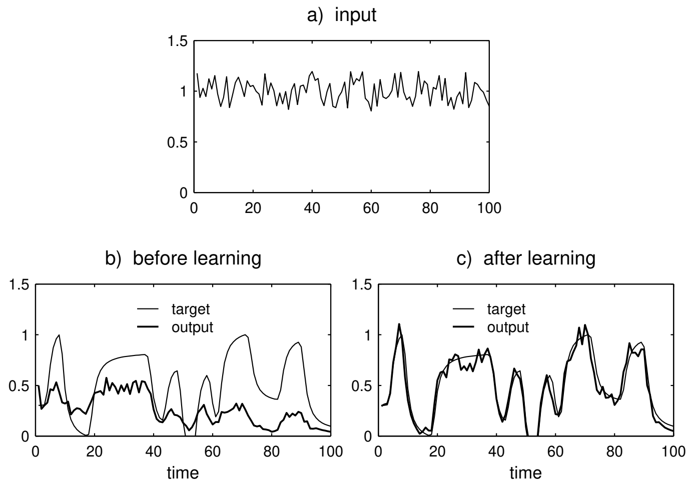

Figure 12.9. Result of gradient descent training for the hidden synaptic parameters in a feedforward neural net with one hidden layer consisting of 12 neurons. The common input time series to the synapses of these 12 neurons is shown in panel a). The target time series and the actual sum of the output time series of the 12 neurons before training are shown in panel b). Panel c) shows the same target time series together with the sum of the output time series of the 12 neurons after applying gradient descent to the 5 hidden parameters $C_0, V_0, \tau_C, \tau_V, \alpha$ of each synapse according to section 12.3. From [Zador et al., 1998].

12.6 Conclusions

We have shown that synapses in biological neural systems play a rather complex role in the context of computing on spike trains. This implies that one is likely to lose a substantial amount of computational power if one models biological networks of spiking neurons by artificial pulsed neural nets that employ the same type of static synapses that are familiar from traditional neural network models.

It has been shown that dynamic synapses of the type that we have described in this chapter can easily be simulated by electronic circuits [Fuchs, 1998], and hence can in principle be integrated into neural VLSI. In addition all analog techniques for storing a weight in neural VLSI automatically create "dynamic synapses", since the value of the stored weight tends to drift. An exciting challenge would be to find computational uses for such inherent synaptic dynamics, which so far has only been viewed as a defect. Finally we have indicated that the notion of learning and the nature of learning algorithms changes if one takes into account that synapses have to be modelled as history-dependent dynamical systems with several hidden parameters that control their dynamics, rather than as static scalar variables ("synaptic weights"). Hence if one wants to mimic adaptive mechanisms of biological neural systems in artificial pulsed neural nets, one is forced go beyond the traditional ideas from neural network theory and look for new types of learning algorithms.

References

[Abbott et al., 1997] Abbott, L., Varela, J., Sen, K., and S.B., N. (1997). Synaptic depression and cortical gain control. *Science*, 275:220–4.

[Allen and Stevens, 1994] Allen, C. and Stevens, C. (1994). An evaluation of causes for unreliability of synaptic transmission. *Proc. Natl. Acad. Sci.*, USA, 91:10380–3.

[Atwood et al., 1978] Atwood, H., Govind, C., and I., K. (1978). Non homogeneous excitatory synapses of a crab stomach muscle. *Journal of Neurobiology*, 9:17–28.

[Back and Tsoi, 1991] 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.

[Bolshakov and Siegelbaum, 1995] Bolshakov, V. and Siegelbaum, S. A. (1995). Regulation of hippocampal transmitter release during development and long-term potentiation. *Science*, 269:1730–4.

[Brown et al., 1976] Brown, T., Perkel, D., and Feldman, M. (1976). Evoked neurotransmitter release: statistical effects of non uniformity and non stationarity. *Proc. Natl. Acad. Sci.*, USA, 73:2913–2917.

[Castro-Alamancos and Connors, 1997] Castro-Alamancos, M. and Connors, B. (1997). Distinct forms of short-term plasticity at excitatory synapses of hippocampus and neocortex. *Proc. Natl. Acad. Sci.*, USA, 94:4161–4166.

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

[Fisher et al., 1997] Fisher, S., Fischer, T., and Carew, T. (1997). Multiple overlapping processes underlying short-term synaptic enhancement. *Trends in Neuroscience*, 20:170–7.

[Fuchs, 1998] Fuchs, H. (1998). Neural networks with dynamic synapses in analog electronics. In preparation.

[Hessler et al., 1993] Hessler, N., Shirke, A., and R., M. (1993). The probability of transmitter release at a mammalian central synapse. *Nature*, 366:569–572.

[Katz, 1966] Katz, B. (1966). *Nerve, Muscle, and Synapse*. New York, McGraw-Hill.

[Liaw and Berger, 1996] Liaw, J.-S. and Berger, T. (1996). Dynamic synapse: A new concept of neural representation and computation. *Hippocampus*, 6:591–600.

[Maass and Zador, 1998] Maass, W. and Zador, A. (1998). Dynamic stochastic synapses as computational units, *Advances in Neural Information Processing Systems*, vol. 10, 194–200. Journal version appears in *Neural Computation*.

[Magleby, 1987] Magleby, K. (1987). Short term synaptic plasticity. In *Synaptic Function*, Edelman, G. M., Gall, W. E., and Cowan, W. M., editors, Wiley, New York.

[Markram and Tsodyks, 1996] Markram, H. and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature*, 382:807–10.

[Markram and Tsodyks, 1997] Markram, H. and Tsodyks, M. (1997). The information content of action potential trains: a synaptic basis. *Proc. of ICANN 97*, 13–23.

[Manabe and Nicoll, 1994] Manabe, T. and Nicoll, R. (1994). Long-term potentiation: evidence against an increase in transmitter release probability in the ca1 region of the hippocampus. *Science*, 265:1888–92.

[Murthy et al., 1997] Murthy, V., Sejnowski, T., and Stevens, C. (1997). Heterogeneous release properties of visualized individual hippocampal synapses. *Neuron*, 18:599–612.

[Principe, 1994] Principe, J. C. (1994). An analysis of the gamma memory in dynamic neural networks. *IEEE Trans. on Neural Networks*, 5(2), 331–337.

[Rosenmund et al., 1993] Rosenmund, C., Clements, J., and Westbrook, G. (1993). Nonuniform probability of glutamate release at a hippocampal synapse. *Science*, 262:754–757.

[Ryan et al., 1996] Ryan, T., Ziv, N., and Smith, S. (1996). Potentiation of evoked vesicle turnover at individually resolved synaptic boutons. *Neuron*, 17:125–34.

[Stratford et al., 1996] Stratford, K. J., Tarczy-Hornoch, K., Martin, K. A. C., Bannister, N. J., and Jack J. J. B. (1996). Excitatory synaptic inputs to spiny stellate cells in cat visual cortex. *Nature*, 382:258–61.

[Tsodyks and Markram, 1997] Tsodyks, M. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci.*, USA, 94:719–23.

[Varela et al. 1997] Varela, J. A., Sen, K., Gibson, J., Fost, J., Abbott, L. F., and Nelson, S. B. (1997). A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *J. Neurosci*, 17:7926–7940.

[Zador, 1998] Zador, A. M. (1998). Synaptic plasticity. In *Biophysics of Computation*, *in press*, Koch, C., editor, MIT Press, Boston.

[Zador and Dobrunz, 1997] Zador, A. and Dobrunz, L. (1997). Dynamic synapses in the cortex. *Neuron*, 19:1–4.

[Zador et al., 1998] Zador, A. M., Maass, W., and Natschläger, T. (1998) Learning in neural networks with dynamic synapses, in preparation.

[Zucker, 1989] Zucker, R. (1989). Short-term synaptic plasticity. *Annual Review of Neuroscience*, 12:13–31.