Biol. Cybern. 84, 57–62 (2001)

Biological Cybernetics

Silicon synaptic depression

Christoph Rasche[1], **Richard H. R. Hahnloser**[2]

[1] Institute of Neuroinformatics, ETHZ/UNIZH, Winterthurerstr. 190, 8057 Zürich, Switzerland
[2] Department of Brain and Cognitive Sciences, MIT E25-210, 45 Carleton Street, Cambridge, MA 02139, USA

Received: 10 February 2000 / Accepted in revised form: 16 June 2000

Abstract. The recent quantitative description of activity-dependent depression in the synaptic transmission between cortical neurons has lead to many interesting suggestions of possible computational implications. Based on a simple biological model, we have constructed an analog circuit that emulates the properties of short-term depressing synapses. The circuit comprises only seven transistors and two capacitors per synapse, and is able to reproduce computational features of depressing synapses such as the 1/F law, the detection of long intervals of presynaptic silence and the sensitivity to redistribution of presynaptic firing rates. It provides a useful basis for implementing neural networks with dynamical synapses.

1 Introduction

Neurons are highly dynamic devices. They show adaptation at all functional levels – soma, dendrite and synapse – and on different time scales, from milliseconds to hours to days (Koch 1999). To explore the effects of adapting neurons in a network, computer simulations (e.g. Bower and Beeman 1994; Hines 1989) are helpful but slow if the network is large and if the model neuron contains many (adapting) time constants. Neuromorphic engineers therefore take the approach of emulating neurons in analog complementary metal-oxide semiconductor (CMOS) circuits, designed in very large scale integrated (VLSI) technology (Mead 1989; Douglas et al. 1995; Watson 1997; Elias and Northmore 1999). Thereby one gains speed: silicon neuromorphic circuits emulate the neural computations in real-time, independent of the size of the neural network. Various adaptive neuronal mechanisms have already been transformed into silicon analog circuits (Mahowald and Douglas 1991; Shin and Koch 1999; Simoni and DeWeerth 1999) and here we report an adaptive silicon synapse.

A striking feature of synaptic transmission between neocortical pyramidal neurons is the adaptation of amplitudes of excitatory postsynaptic potentials (EPSPs) to the short-term history of presynaptic activity (Abbott et al. 1997; Markram and Tsodyks 1996). Particularly interesting is synaptic short-term depression, which is the decrease of synaptic strength, measured by the amplitude of EPSPs, in response to ongoing presynaptic activity. Phenomenologically, synaptic depression can be described by a dynamical EPSP amplitude A that is reduced by a depressive factor $d < 1$ immediately after a presynaptic spike and that recovers towards some resting value between presynaptic spikes (Abbott et al. 1997). The larger the interval between two consecutive presynaptic spikes, the longer is the recovery time of the synapse, and the larger is the amplitude elicited by the second spike. Several suggestions have been made about the computational use of depressing synapses, as dynamical cortical gain control (Abbott et al. 1997; Nelson and Turrigiano 1998) or reading out neuronal synchrony (Senn et al. 1998) (for a recent review, see Maass and Zador 1999). Here we present an analog electronic circuit implementation of short-term depressing synapses and demonstrate their computational impact onto a postsynaptic target neuron.

Similar to previous neuromorphic spiking neural networks (Douglas et al. 1995; Deiss et al. 1999; Elias and Northmore 1999), we represent spikes, that are sent between neurons, as (digital) pulses. Our synaptic circuit transforms a presynaptic pulse into an equivalent of an excitatory postsynaptic current (EPSC). The synaptic current is then dumped onto circuitry simulating the passive RC-like behavior of a neuron generating an EPSP (Mahowald 1992; Elias and Northmore 1999). In previous work, synapses were modeled as static elements (Rasche and Douglas 1999). That is, the amplitudes of elicited EPSPs were constant and independent of the presynaptic stimulation frequency. Here we extend this static synapse using a simple circuit that modes the

Correspondence to: C. Rasche
(e-mail: rasche@klab.caltech.edu)

dynamics of synaptic depression. We call this extended circuit the depressing synapse.

2 Methods

In Fig. 1 the static synapse is marked by the three grey blocks 'input', 'strength' and 'EPSC modelling'. The presynaptic pulse, arriving in the 'input' block, triggers a current pulse, I_{WGT}, in the 'strength' block, which is then converted into an approximation of a real EPSC, I_{EPSC}, in the 'EPSC modelling' block (see Rasche and Douglas 1999, for more details).

The depressing synapse includes the block 'dynamic weight', which modulates the synaptic strength by changing the amplitude of the current pulses I_{WGT} according to the synaptic weight voltage V_{w}. This voltage is held by a capacitor C_{w} and is changed by a switch and a diode. Every time a presynaptic pulse arrives on the synapse, the switch removes a fixed amount of charge from the capacitor and so decreases V_{w} by a fixed voltage drop ΔV_{w}. The exact value of ΔV_{w} is determined by the parameter WD ('weight depression'). Between presynaptic spikes, V_{w} recovers (increases) logarithmically in time towards a resting value given by the source voltage RW of the diode, called the 'resting weight'. In other words, the diode and the capacitor have a similar role as a RC circuit with a battery set to the resting weight RW. We thus have three parameters in our 'dynamic weight' block: WD, RW and C_{w}. The capacitance C_{w} is fixed for a given process and can only be modified by a new design. WD and RW are voltages that need adjustment.

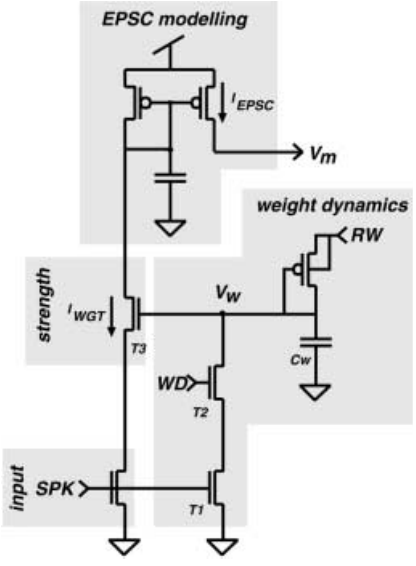

Fig. 1. Circuitry for synaptic depression. A simple static synapse is modeled by the three grey blocks labeled 'input', 'strength' and 'EPSC modelling'. The fourth block, 'weight dynamics', models the time dependence of the synaptic strength, set by the 'synaptic weight' voltage V_{w}. It consists of a switch and a diode. The switch (formed by transistors T1 and T2) is driven by a presynaptic pulse, SPK, on the gate of T1. The gate voltage WD ('weight depression') of T2 causes a voltage drop ΔV_{w} after a presynaptic pulse. Between presynaptic pulses, a diode-connected p-type transistor increases V_{w} towards a resting value determined by the diode source RW ('resting weight')

We have constructed a spiking silicon neuron with four such depressing synapses. In approximation to a real neuron, the spiking silicon neuron consists of a leakage, a sodium and a potassium conductance, each modelled by a follower integrator. The output voltage of the follower integrator representing the leaky membrane corresponds to the membrane potential onto which EPSCs are sourced (see Rasche and Douglas, 2000, for more details). A chip of 2.2 mm^2 was fabricated using standard 1.2 μm CMOS technology. Our circuitry uses only a small fraction of this area. Transistor sizes are generally 5 μm by 5 μm, and C_{w} is about 0.2 pF. All our results are drawn from this chip.

Our mathematical results are best expressed by defining the exponential function $f(V) = \mathrm{e}^{\kappa V/U_T}$, where V is a voltage, $\kappa \simeq 0.7$ is the gate efficiency constant and $U_T = 25$ mV is the thermal voltage of a CMOS transistor. For example, the current I_{WGT} flowing through the transistor T3 (in saturation) is proportional to $f(V_{\mathrm{w}})$.

If the capacitance in the EPSC modelling block is small, then $I_{\mathrm{EPSC}} = I_{\mathrm{WGT}}$ (the capacitance is small if for most step currents I_{WGT}, the gate voltage of the current mirror reaches a steady value in much less than the duration $\Delta t = 1$ ms of a presynaptic pulse). In this case, if the follower integrator modelling the leaky membrane is in the linear (RC) regime, then we find proportionality between the EPSP amplitude A and $f(V_{\mathrm{w}})$:

$$A \propto f(V_{\mathrm{w}}) \ . \tag{1}$$

If the follower integrator does not operate in the linear but in the saturated regime, which is often the case for our circuit, then proportionality holds as well in the approximation that the bias current of the follower integrator modelling the leaky membrane is small compared to I_{EPSC}.

Notice that if the capacitance in the EPSC modelling block is not small in the above sense, then proportionality between A and $f(V_{\mathrm{w}})$ does not hold. Calculations show that in this case the relationship between A and $f(V_{\mathrm{w}})$ is approximatively exponential. However, for the following, we assume that (1) holds true for our circuit.

The resting or initial amplitude A_1 of the depressing synapse is determined by RW. Initially, the synapse is fully recovered and so $V_{\mathrm{w}} \simeq \mathrm{RW}$, which leads to $A_1 \propto f(\mathrm{RW})$. From (1) we also get that the amplitude after a presynaptic spike is reduced by a constant factor: $A_1 = f(V_{\mathrm{w}}) \to f(V_{\mathrm{w}} - \Delta V_{\mathrm{w}}) = f(-\Delta V_{\mathrm{w}})A_1 = dA_1$. The depressive factor d is thus given by

$$d = f(-\Delta V_{\mathrm{w}}) < 1 \ . \tag{2}$$

Assuming that the transistor T2 is always in saturation, which is the case for large values of RW and not too high presynaptic frequencies F, we find that the voltage drop ΔV_{w} induced by a presynaptic pulse is approximatively given by:

$$\Delta V_{\mathrm{w}} \simeq \frac{I_{\mathrm{T2}}\Delta t}{C_{\mathrm{w}}} \propto \frac{\Delta t}{C_{\mathrm{w}}} f(\mathrm{WD}) \ , \tag{3}$$

where I_{T2} denotes the current flowing through transistor T2. From (2) and (3) we see that the parameter WD determines the factor d in a very sensitive manner (by a double exponential). In conclusion, similar to the simple model described in Sect. 1, because ΔV_w is independent of V_w, we get that d is independent of the momentary EPSP amplitude.

3 Results

In Fig. 2 we show the responses of the membrane potential V_m and the synaptic weight V_w to two periodic presynaptic trains of pulses at different frequencies. The amplitudes A_1 of the initial EPSPs are equal for both trains. Subsequent amplitudes are depressed, due to a decrease of V_w following each spike. Depression continues until a steady amplitude A_s is reached, which occurs when the voltage drop ΔV_w after a pulse equals the recovery between pulses. The steady amplitude A_s for the train of higher frequency (Fig. 2a) is smaller than that of the train of lower frequency (Fig. 2b).

We have measured the steady amplitudes A_s for a range of stimulation frequencies (Fig. 3a). We have also calculated A_s by considering that the weight voltage V_w and $f(V_w)$ are both periodic in time with period $T = 1/F$. Between spikes, $f(V_w)$ recovers linearly in time from its lower value $f(V_s)$:

$$f(V_w) - f(V_s) \propto \frac{T}{C_w} \quad . \tag{4}$$

For this calculation we have assumed that the diode remains saturated, which turns out to be a good approximation for the parameters in the caption of Fig. 3a if the frequency of the presynaptic train is larger than 10 Hz. Evaluating (4) at time $t = T$, using $f(V_w) = f(V_s) + f(\Delta V_w)$, we obtain:

$$A_s \propto f(V_s) \propto \frac{t}{C_w(f(\Delta V_w) - 1)} \propto 1/F \quad . \tag{5}$$

This equation expresses the steady EPSP amplitude A_s as a function of the fixed capacitance C_w and the weight depression WD (via Eq. 3). Most importantly, we see that A_s is inversely proportional to the presynaptic frequency F, which faithfully models the experimental finding (see Sect. 4).

Figure 3b shows the effect that the weight depression has on the initial depression. When WD is large, a large voltage drop ΔV_w results (a small d), which in turn causes a fast depression of the first few amplitudes.

Only rarely do real neurons fire with a regular train of action potentials such as in Fig. 2. Very often, the spike trains of neurons can be described by Poisson distributions (Softky and Koch 1993) – the interspike intervals are exponentially distributed. Next we used a stimulus of a Poisson-distributed train of pulses of mean frequency λ. Figure 4 compares the integrating properties of a depressing synapse to that of a static synapse. In order to establish a reference for the comparison, the amplitude of EPSPs elicited by the static synapse was adjusted to be equal to the steady amplitude A_s of the depressing synapse that was elicited by a regular spike train of frequency λ. The firing threshold of the postsynaptic neuron was set to a level such that stimulation of a recovered depressing synapse was able to trigger postsynaptic spikes. Therefore, the depressing synapse triggered a spike in the postsynaptic neuron when the interval to the previous presynaptic pulse was large. On the other hand, the static synapse triggered spikes when the interval to the previous pulse was small, as consecutive EPSPs added up and so allowed the membrane potential to cross firing threshold.

Figure 5 shows the effect of redistributing presynaptic firing rates. Four synapses are stimulated with

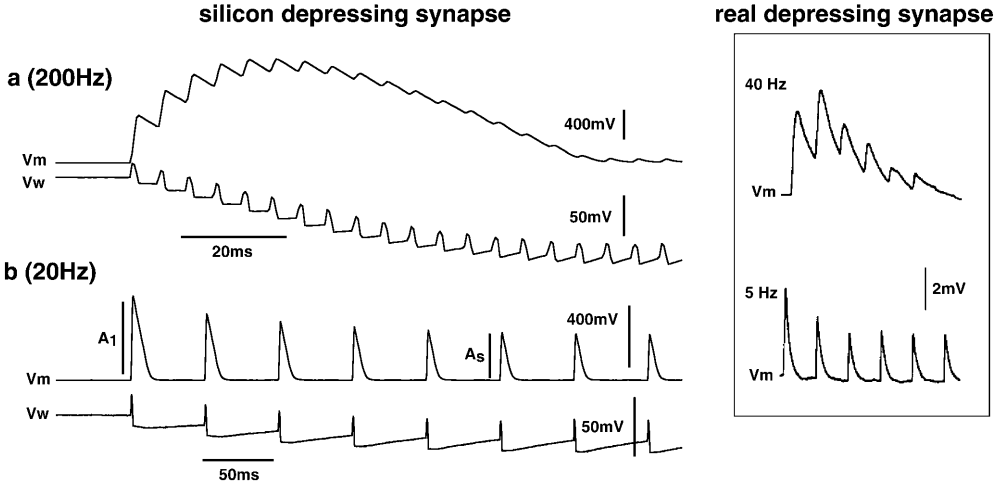

Fig. 2. Depressing EPSPs. This and the following figures show recordings from a fabricated chip. The dynamics of membrane potential V_m and synaptic weight V_w are shown at two different stimulation frequencies: **a** 200 Hz, **b** 20 Hz. In both cases, the initial EPSP amplitude, A_1, is about 500 mV. The steady amplitude, A_s, depends on the presynaptic frequency and is smaller in case **b**. The small 'spikes' on the V_w trace are due to crosstalk from the presynaptic pulse applied to the synapse. Parameter values: RW = 0.757 V, WD = 0.2 V. Resting membrane potential = 2.0 V. On the right hand side, real V_m recordings of a cortical neuron are shown for a high and a low stimulation frequency, taken with permission from (Markram and Tsodyks 1996)

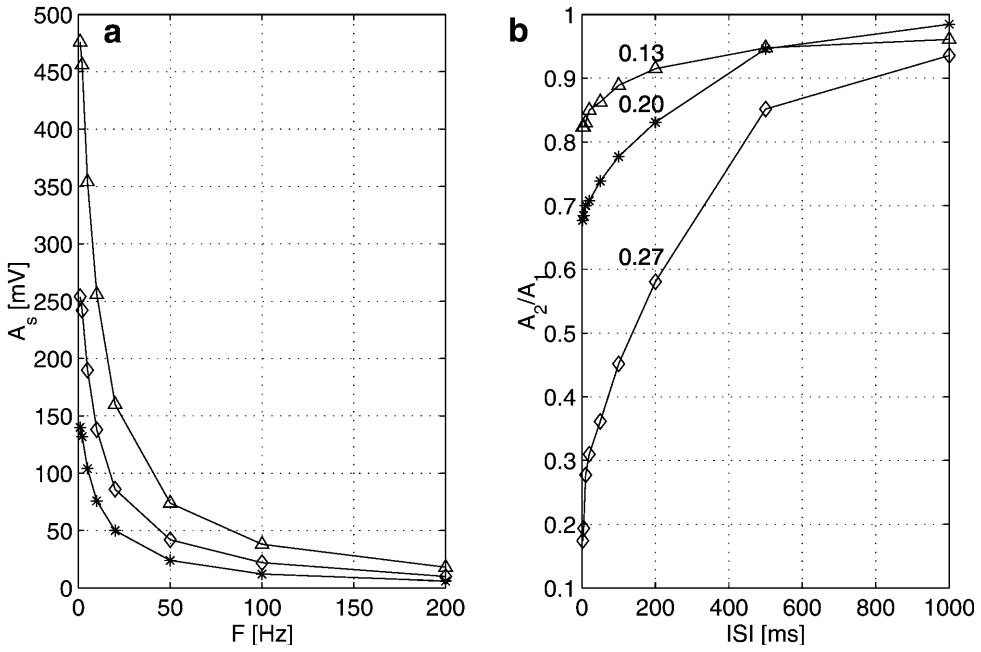

Fig. 3. Frequency and interspike-interval (ISI) dependence of EPSPs. **a** Steady amplitude A_s as a function of presynaptic frequency F: A_s was measured for frequencies from 1 to 500 Hz and for three different initial amplitudes A_1 set by RW (A_1 = 476 mV, RW = 0.81 V; A_1 = 254 mV, RW = 0.79 V; A_1 = 140 mV, RW = 0.77 V). The larger A_1 is, the larger is A_s. WD = 0.2 V. **b** The effect of different weight depressions WD on the amplitudes of consecutive pulses (WD = 0.13 V, 0.20 V, 0.27 V). A_2/A_1 is the ratio of the second and the initial EPSP amplitudes, shown as a function of their time interval. For large ISIs (i.e. >500 ms) the curves converge to a ratio of unity (hardly any depression). The larger the weight depression WD, the smaller is the amplitude ratio (the faster is the depression)

Poisson-distributed pulses, two at a high frequency and two at a low frequency. At time zero the high and low frequencies are interchanged so that the sum of all presynaptic firing rates remains constant. Before this rate redistribution, the less stimulated synapses have recovered. Shortly after the rate redistribution, these

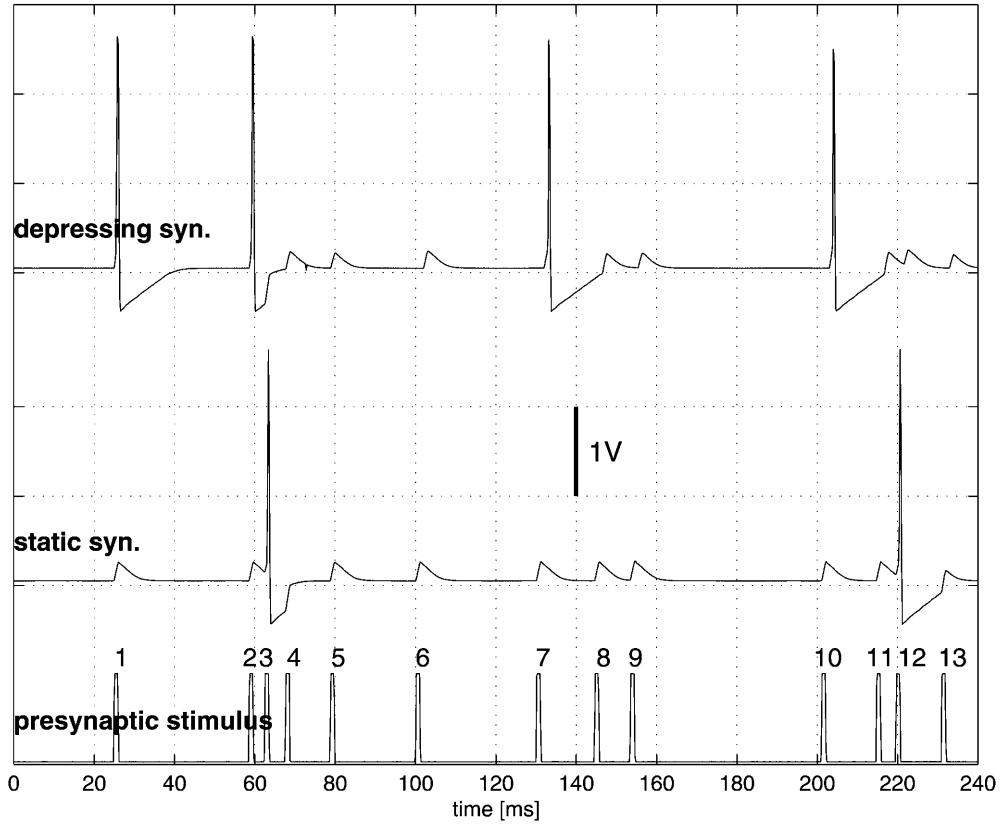

Fig. 4. Integrating properties of the depressing synapse. Response to a Poisson-distributed train of pulses (50 Hz, 'presynaptic stimulus') by a depressing synapse (WD = 0.2 V, RW = 0.773 V) and a static synapse (V_w is fixed by using a follower). The depressing synapse triggers spikes when there is a sufficiently large interval between pulses (e.g. pulses 1, 2, 7 and 10). The static synapse triggers a spike only when two consecutive pulses are close to each other (e.g. pulses 2, 3 and 11, 12). Parameters: spiking threshold = 2.230 V

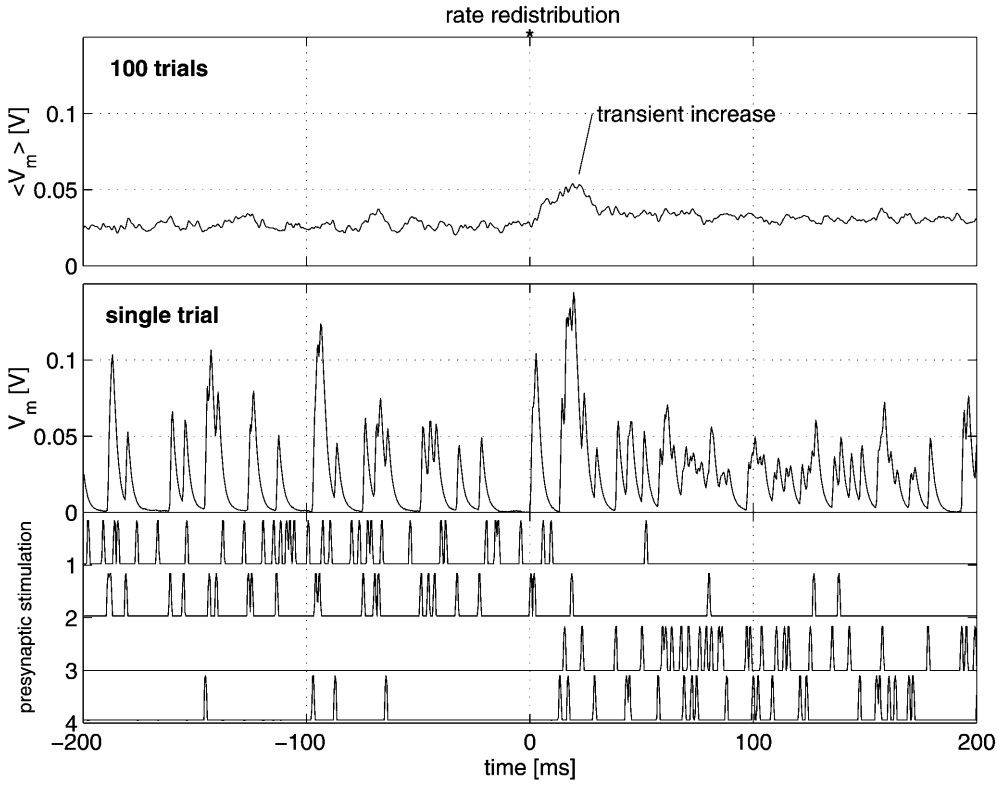

Fig. 5. Effect of presynaptic rate redistribution. Lower panel (presynaptic stimulation): four synapses are stimulated with Poisson-distributed pulse trains, synapses 1 and 2 at 200 Hz and synapses 3 and 4 at 20 Hz. At time zero, the frequencies of the pulses are interchanged. The middle panel (single trial) shows V_m in response to the stimulation shown in the bottom panel. Approximately 10 ms after the rate change, synapses 3 and 4 elicit EPSPs almost simultaneously. Because both synapses were quiet for tens of milliseconds (recovered), the corresponding EPSPs are large. The top panel (100 trials) shows the average membrane potential ($\langle V_m \rangle$) over 100 trials. Immediately after the rate change, there is a transient increase in $\langle V_m \rangle$ that peaks after 10 ms. The mean levels of $\langle V_m \rangle$ before and long after the rate change are slightly different due to mismatch between the synaptic circuits. Parameters: WR = 0.753 V, WD = 0.2 V

synapses cause large additive EPSPs because of the switching from low to high frequency stimulation. Hence, a transient increase in the membrance potential V_m can be seen. Some time after the rate redistribution, the synapses reach an average state of depression and V_m decreases back to its level before the redistribution. In contrast, a neuron with static synapses would not be able to sense this reshuffling of presynaptic firing, because all that a neuron with static synapses is capable of sensing is the sum of presynaptic firing rates.

4 Discussion

Analogous to biology (Markram and Tsodyks 1996; Abbott et al. 1997), EPSP amplitudes of our silicon synapse have a depressing characteristic, and in the stationary case they are inversely proportional on the presynaptic stimulation frequency (Fig. 3a). This law implies that the impact of synaptic currents on the membrane potential tends to be independent of the presynaptic frequency of action potentials, e.g. synapses transmitting high frequencies do not necessarily lead to a stronger depolarization than synapses transmitting low frequencies. Furthermore, if a presynaptic neuron changes its firing from 10 Hz to 20 Hz, this has a similar impact on the postsynaptic membrane potential as a change from 50 Hz to 100 Hz, reminiscent of the Weber-Fechner relation (Abbott et al. 1997). We have not tested this relationship in our silicon circuits, because it relies on a proper RC circuit emulating the leaky membrane of the neuron. The follower integrator we used is merely an approximation to such a RC circuit and is therefore not suitable for demonstrating this law. There exist more realistic RC circuit implementations that might be more appropriate to show this relationship (Elias and Northmore 1999; Rasche 1999).

We have demonstrated several computational functions of our depressing synapse. After a long presynaptic interspike interval, a new spike elicits a large EPSP at a recovered synapse. If the spiking threshold of our silicon neuron is set to a level where only large EPSPs lead to an action potential, then the postsynaptic neuron reports the end of such a silent period, as shown in Fig. 4. Senn et al. (1998) have used this ability to read out the synchrony of spikes between spike trains. A neuron, whose depressing synapses are fed with correlated spike trains, reports synchrony of silent intervals by firing at the end of the synchronous silence.

Silicon depressing synapses are sensitive to the redistribution of presynaptic firing rates (Fig. 5). A

redistribution of presynaptic rates occurs when the input rates assigned to synapses change suddenly, while leaving the overall stimulation to the neuron constant. This computational ability can be beneficial when it comes to reliably detecting a sudden change in sensory input. For example, this might happen during saccadic eye movements in an environment where the luminosity is constant across different directions of the visual field. During a saccadic eye movement, a redistribution of visual inputs takes place, in the form of a retinal translation. In this way, synaptic depression could speed up visual processing immediately after saccades by triggering neurons to respond vigorously to the new input.

Our synaptic circuit can easily be generalized. For example, by replacing the n-type transistor (T3) of the strength block with a p-type transistor (Fig. 1), the circuit works as a facilitating synapse: a decreasing synaptic weight V_w would then increase the synaptic current. A similar model is already implemented (Rasche 1999). As an application, by combining facilitation and depression of different time scales on a single synapse, it might be possible to construct feedforward networks of dynamical synapses that realize arbitrary nonlinear temporal filters, as it has been shown theoretically (Maass and Zador 1999).

A critical feature of synaptic analog VLSI circuits is their size. Because a synapse is the most common circuit in a neuromorphic network, it has to be small to conserve chip area. Our depressing synapse consists of only seven transistors and two capacitors, and is designed to be compact. Another desirable feature of a synaptic circuit is a small number of synaptic parameters, because this reduces the on-chip wiring. Our synapse has only two parameters, RW and WD, and is thus highly suitable for neuromorphic networks (Deiss et al. 1999).

Our adaptive synapse regulates the synaptic input to a neuron. But neuronal adaptation can also occur at sites other than synapses. For example, a pyramidal soma can adapt its spiking frequency on a short-term and a long-term basis (Koch 1999). These forms of adaptation represent a regulation of the output of a neuron; they have already been implemented in silicon by various groups (Shin and Koch 1999; Simoni and DeWeerth 1999; Rasche and Douglas 2000). But adaptation could also occur in dendrites (Verschure and König 1999), which is a possibility that silicon models are only starting to imitate (Rasche 1999). It is an open question as to whether it is possible to derive an integrative computational model comprising these various adaptive mechanisms and to test it in a neuromorphic silicon network.

Acknowledgements. C.R. is supported by funding to Rodney Douglas from the Schwerpunktprogramm (SPP) Biotechnologie des Schweizerischen Nationalfonds. R.H. is supported by the Swiss National Science Foundation. We thank Dave Lawrence and Philipp Häfliger for software support. We acknowledge the feedback on an earlier version of the manuscript by our colleagues Thomas Voegtlin, Shih-Chii Liu, Regina Mudra, Tobi Delbrueck and Rodney Douglas.

References

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

Bower JM, Beeman D (1994) The book of GENESIS: Exploring realistic Neural Models with the GEneral NEural SImulation System. Springer, Berlin Heidelberg New York

Deiss S, Douglas R, Whatley A (1999) A pulse-coded communications infrastructure for neuromorphic systems. In: Maass W, Bishop CM (eds) Pulsed Neural Networks. MIT Press, Cambridge, Mass., pp 157–178

Douglas R, Mahowald M, Mead C (1995) Neuromorphic analog VLSI. Ann Rev Neurosci 18: 255–281

Elias JG, Northmore DPM (1999) Building silicon nervous systems with dendritic tree neuromorphs. In: Maass W, Bishop CM (eds) Pulsed Neural Networks. MIT Press, Cambridge, Mass., pp 135–156

Hines M (1989) A program for simulation of nerve equations with branching geometries. Int J Biomed Comput 24: 55–68

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

Maass W, Zador AM (1999) Computing and learning with dynamic synapses. In: Maass W, Bishop CM (eds) Pulsed Neural Networks. MIT Press, Cambridge, Mass., pp 321–336

Mahowald M (1992) Evolving analog VLSI neurons. In: McKenna T, Davis J, Zornetzer S (eds) Single Neuron Computation. Academic Press, Orlando, Fla., pp 413–435

Mahowald M, Douglas R (1991) A silicon neuron. Nature 354: 515–518

Markram H, Tsodyks M (1996) Redistribution of synaptic efficacy between cortical pyramidal neurons. Nature 382: 807–810

Mead CA (1989) Analog VLSI and Neural Systems. Addison-Wesley, Reading, Mass.

Nelson S, Turrigiano G (1998) News and views: Synaptic depression – a key player in the cortical balancing act. Nature Neurosci 1: 539–541

Rasche C (1999) Analog VLSI Circuits for Emulating Computational Features of Pyramidal Cells. PhD thesis, no. 13268, Eidgenössische Technische Hochschule Zürich, Switzerland

Rasche C, Douglas R (1999) Silicon synaptic conductances. J Comput Neurosci 7: 33–39

Rasche C, Douglas R (2000) An improved silicon neuron. Analog Integr Circuits Signal Process, 23: 227–236

Senn W, Segev I, Tsodyks M (1998) Reading neuronal synchrony with depressing synapses. Neural Comput 10: 815–819

Shin J, Koch C (1999) Dynamic range and sensitivity adaptation in a silicon spiking neuron. IEEE Trans Neural Netw 10: 1232–1238

Simoni MF, DeWeerth SP (1999) Adaptation in an a VLSI model of a neuron. IEEE Trans Circuits Sys II 46: 967–970

Softky W, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSP's. Journal of Neuroscience 13: 334–350

Verschure P, König P (1999) On the role of biophysical properties of cortical neurons in binding and segmentation of visual scenes. Neural Comput 11: 1113–1138

Watson A (1997) Neuromorphic engineering. Science 277: 1935