AN ADAPTIVE SILICON SYNAPSE

Elisabetta Chicca, Giacomo Indiveri and Rodney Douglas

Institute of Neuroinformatics, University of Zurich and ETH Zurich
Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
chicca@ini.phys.ethz.ch, giacomo@ini.phys.ethz.ch

ABSTRACT

We present an analog circuit for implementing models of synapses with short–term adaptation, derive analytical solutions for spiking input signals, and present experimental results measured from a chip fabricated using a standard 1.5μm CMOS technology. The circuit is suitable for integration in large arrays of integrate–and–fire neurons and consequently for evaluating computational roles of short–term adaptation at the network level.

1. INTRODUCTION

Silicon synapses are circuits, typically used in VLSI networks of integrate–and–fire neurons [1, 2, 3, 4, 5], that implement models of biological synapses. Recent developments in the neuroscience community evidence how biological synapses are not simple interfacing elements for transmitting signals across neurons, but play an important computational role in biological neural networks [6]. One of the peculiar properties of biological synapses is their ability to exhibit *short–term plasticity*: the dynamic modulation of synaptic strength by the timing of the input stimulation [7]. Although there has been significant progress in understanding the mechanisms underlying short–term synaptic plasticity, its functional relevance in neural circuits remains relatively obscure. Several hypotheses have been proposed for the computational roles of short–term synaptic plasticity [8], including local gain control [9], stimulus specific adaptation [10], and nonlinear temporal summation [11]. Practical constraints on computational resources restrict the testing of these hypotheses to relatively small simulated networks, simple input signals, or long simulation times. Silicon implementations of synapses that exhibit short-term plasticity are suitable for evaluating the computational roles of synaptic adaptation in large networks of spiking neurons, using complex stimuli and in real–time [12, 13]. We propose an analog circuit for implementing such a type of synapse, derive analytical solutions for pulse input signals (spikes) extending the analysis presented in [14], and present experimental data measured from a prototype chip fabricated using a standard 1.5μm CMOS technology.

2. THE SYNAPTIC CIRCUIT

The circuit has been designed to be used in multi–chip systems interfaced using a communication protocol based on the *Address Event Representation* (AER) [15]. This communication protocol allows the implementation of any arbitrary connectivity among spiking neurons within and across the chip boundaries [16, 17]. A common architecture for an AER VLSI network of spiking neurons is a one–dimensional array of integrate–and–fire circuits arranged in a column, with several afferent synaptic circuits for each neuron (see Fig. 1). Each synapse in the synaptic array is identified by its row and column address. When an AER input request signal arrives, the row and column addresses are encoded and the corresponding synapse is stimulated.

The core of the synaptic circuit consists of two *Current Mirror Integrators* (CMI) [1] that integrate the input spikes and produce a positive (*facilitating*) and a negative (*depressing*) current with different dynamics. Each time an input spike arrives, the facilitating CMI adds a set amount of charge onto its integrating capacitor, modifying the facilitating current accordingly. At the same time, a different amount of charge is summed onto the capacitor of the depressing CMI, modifying the depressing current. The facilitating and depressing current are then subtracted to generate the net output current.

The proposed AER synaptic circuit is shown in Fig. 2. The facilitating CMI is implemented by transistors $M4$-$M5$ and by capacitor C_f. The depressing CMI is implemented by $M12$ and $M14$ and by C_d. Both CMIs are interconnected to circuital blocks that implement the AER interface (transistors $M1$–$M2$ and $M9$–$M10$, all acting as digital switches).

The AER part of the synaptic circuit operates as follows: upon the arrival of a request signal, the row and column encoders set the voltages V_{Qy} and V_{Qx} of the synapse identified by the row and column addresses high (see Fig. 1). Transistor $M9$ then generates the AER acknowledge signal transmitted back to the sender chip.

When the synapse is addressed, transistors $M1$–$M2$ and $M10$ are active and the currents set by the bias voltages

0-7803-7761-3/03/$17.00 ©2003 IEEE

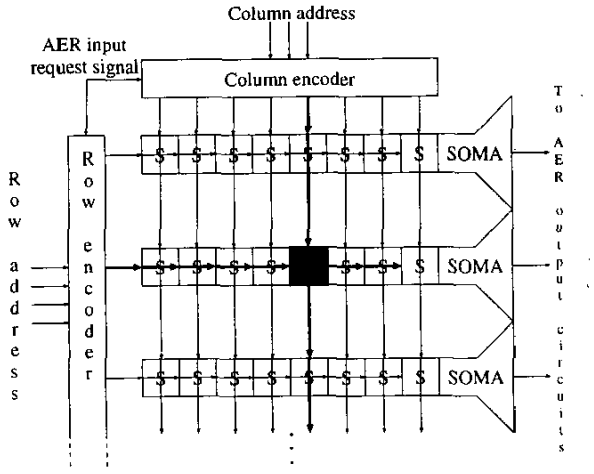

Figure 1: Schematic diagram of the chip architecture. Every square marked with an S represents a silicon synapse. When an AER request signal arrives, the column and row encoders generate an input pulse on the addressed synapse (filled square). The output spikes are sent to the on–chip AER output circuits which generate address–event signals.

Vw_f and Vw_d flow through transistors $M3$ and $M11$ respectively. The capacitors C_f and C_d integrate these currents, changing the voltages across them. The bias voltages $V\tau_f$ and $V\tau_d$ are used to change the temporal dynamics of the facilitating and the depressing parts of the synapse. The facilitating current I_f has an exponential dependence on the voltage across capacitor C_f and on the voltage $V\tau_f$, similarly, the depressing current I_d changes exponentially with the voltage across capacitor C_d and with $V\tau_d$. The synaptic output current I_{out} is the sum of the output currents of the two CMIs (when they are not limited by the cascode transistors $M6$ and $M13$).

To derive the dynamics of the output current I_{out} in response to a train of spikes, we need to evaluate the response of the two CMIs. The CMI has been analytically characterized in [1, 14] where an explicit solution is given only for the steady state condition. We derived a more general explicit solution that does not require the steady state assumption. If we consider an input train of spikes with fixed duration (Δt) for which $t_i = t_0, t_1, t_2, ...$ (with $t_0 < t_1 < t_2 < ...$) are the times at which the pulses occur, we can write the facilitating current as:

$$I_f(t) = \frac{1}{e^{-\frac{I_{in}}{Q_T}(t-t_i)}\left(\frac{1}{I_f(t_i)} - \frac{1}{AI_{in}}\right) + \frac{1}{AI_{in}}} \quad (1)$$

$$I_f(t) = \frac{1}{\frac{1}{AQ_T}(t - t_i - \Delta t) + \frac{1}{I_f(t_i+\Delta t)}} \quad (2)$$

where I_{in} is the current flowing through transistor $M4$ upon

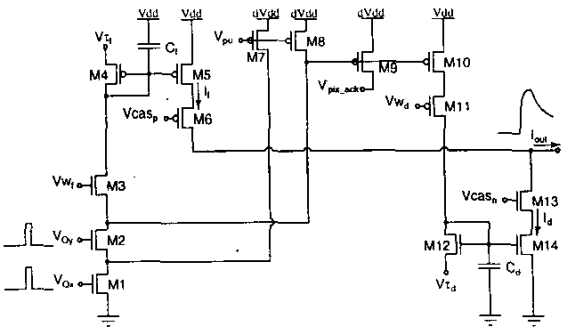

Figure 2: Schematic diagram of the adaptive synapse. Transistors $M3$–$M6$ and capacitor C_f implement the facilitating block of the synaptic circuit. Transistors $M11$–$M14$ and capacitor C_d implement the depressing block. Transistors $M7$–$M8$ prevent pump charge effects by transistors $M1$ and $M2$ [17].

presentation of an input pulse, $Q_T = C_f U_T/\kappa$, U_T is the thermal voltage, κ is the subthreshold slope factor [18], and A is an exponential amplification factor related to the bias voltage $V\tau_f$:

$$A = e^{\frac{V\tau_f - V_{dd}}{U_T}}$$

Equation 1 holds during the spike ($t_i < t \leq t_i + \Delta t$) and eq. 2 holds between two spikes ($t_i + \Delta t < t \leq t_{i+1}$). If the interval between spikes is constant and equal to T, a steady state is reached and the current oscillates between the two values:

$$I_f(t_i) = \frac{1-e^{-\frac{I_{in}}{Q_T}\Delta t}}{\frac{T}{AQ_T} + \frac{1}{AI_{in}}\left(1-e^{-\frac{I_{in}}{Q_T}\Delta t}\right)} \quad (3)$$

$$I_f(t_i + \Delta t) = \frac{1}{\frac{T}{AQ_T} + \frac{1}{AI_{in}}\left(e^{\frac{I_{in}}{Q_T}\Delta t} - 1\right)} \quad (4)$$

The current I_d through the n-type CMI is characterized by equivalent relationships, and the output current of the synaptic circuit is $I_{out} = I_f - I_d$, provided that the CMIs' output current is not limited by the cascode transistors $M6$ and $M13$.

3. EXPERIMENTAL RESULTS

We assembled several instances of the circuit shown in Fig. 2 to *leaky* integrate–and–fire neurons of the type described in [19], together with AER interfacing circuits (as in Fig. 1). The resulting AER neural network was fabricated using a standard 1.5μm CMOS technology. The response of the synaptic circuit was tested by stimulating it with periodic trains of pulses at different frequencies and by measuring

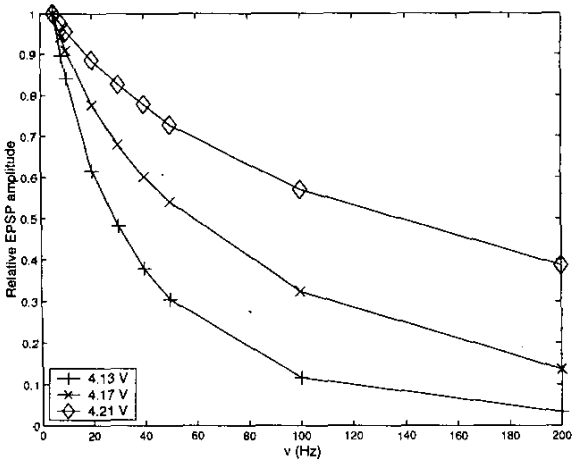

Figure 3: Normalized steady amplitude of EPSP as a function of presynaptic frequency (ν) for three different values of Vw_d (see lower left inset in graph). Each trace is normalized with respect to its maximum EPSP amplitude (656 mV, 748 mV, and 808 mV respectively). The data were measured with $V\tau_f = 4.89\ V$, $Vw_d = 4.19\ V$ and $V\tau_d = 0.1\ V$.

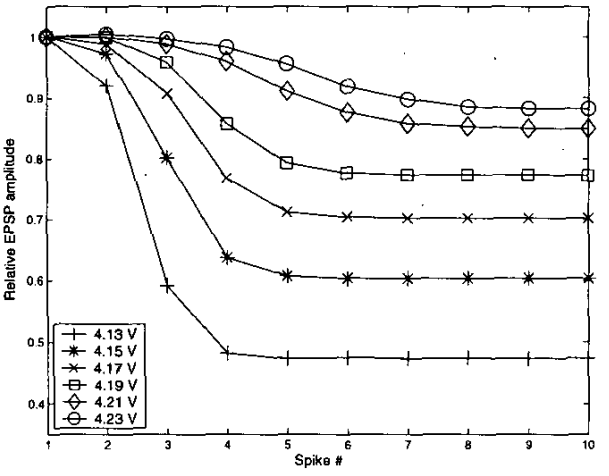

Figure 4: Normalized EPSP amplitude in response to the first ten pulses of a 20Hz train of spikes for three different values of Vw_d (see lower left inset in graph). Each trace is normalized with respect to its maximum EPSP amplitude (approximately 830 mV for all traces). The data were measured with all other parameters set to the same values as reported in Fig. 3.

the change in the voltage across the input capacitor of the *leaky* integrate–and–fire neuron connected to it. In biology this change is commonly referred to as an *excitatory post synaptic potential* (EPSP). The four voltages Vw_f, $V\tau_f$, Vw_d and $V\tau_d$ (see Fig. 2) are used to shape the EPSP dynamics. Specifically, they were set in a way to make the facilitating current reach its steady state before the depressing one. In this way the circuit's overall behavior models a *depressing* synapse. The input spike was varied from 5 to 200Hz. We measured the steady–state amplitude of the corresponding EPSP in response to these stimuli (see Fig. 3). We observed a decreasing steady EPSP amplitude for increasing frequency, as reported for biological synapses [9]. In Fig. 3 we show how this function can be modified by changing, for example, the bias voltage Vw_d of the depressing CMI weight. The contribution of each input spike to the dynamic modulation of the synaptic strength is shown in Fig. 4, where the response to a 20Hz train of spikes is plotted for six different values of the depressing weight ($Vw_d = 4.13\ V, 4.15\ V, 4.17\ V, 4.19\ V, 4.21\ V$ and $4.23\ V$). Both the EPSP measured at the steady state and the rate at which the steady state is reached are sensitive to this parameter.

3.1. Device mismatch and noise

The CMIs operate in the subthreshold, or weak inversion domain [18], and they are the parts of the circuit where the sensitivity to device mismatch is highest. To characterize the inhomogeneities in the total output currents of several synaptic circuits and to evaluate the effect of increasing transistor size, we implemented a chip with two arrays of 32 identical synapses. In the first array the transistors implementing the CMIs are $8\lambda \times 8\lambda$, in the second array they are $16\lambda \times 16\lambda$, where λ is half of the minimum gate width ($\lambda = 0.8\mu m$ in our prototype chip). The relative variation of the steady state output current among the 32 copies of the circuit is about 17% for the array with $8\lambda \times 8\lambda$ transistors and about 7% for the array with $16\lambda \times 16\lambda$ transistors. This result is good compared to the typical variation in transistors subthreshold current [20]. Furthermore, it shows that it is possible to significantly improve the reliability of the circuit by increasing the size of the transistors in the CMIs.

At the single synapse level there are no critical sources of noise. The systematic offset that could be, in theory, introduced by the facilitating CMI would be negligible compared to the explicit leak current of the integrate–and–fire neuron connected to the synapse.

4. CONCLUSIONS

We presented a silicon adaptive synapse consisting of 14 transistors and two capacitors, including the AER interfacing circuitry. Four bias voltages are used to shape the dynamics of the synaptic circuit's output current. We reported experimental data acquired from a VLSI implementation of the circuit (1.5μm standard CMOS technology). The out-

put current of the circuit adapts to the input and reaches a steady-state if the input frequency is constant (see Fig. 4). The data show that the steady EPSP amplitude produced by the output synaptic current decreases as the input frequency increases (see Fig. 3), in agreement with neurophysiological experimental results. We plan to design large arrays of integrate–and–fire neurons interconnected by our adaptive synapse and to implement multi–chip systems interfaced using the AER protocol, to explore the computational roles of short–term synaptic plasticity at the network level.

5. REFERENCES

[1] K. A. Boahen. *Retinomorphic Vision Systems: Reverse Engineering the Vertebrate Retina*. Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1997.

[2] C. Diorio, P. Hasler, B.A. Minch, and C. Mead. A single-transistor silicon synapse. *IEEE Trans. Electron Devices*, 43(11):1972–1980, 1996.

[3] G. Indiveri, T. Horiuchi, E. Niebur, and R. Douglas. A competitive network of spiking VLSI neurons. In F. Rattay, editor, *World Congress on Neuroinformatics 2001*, ARGESIM Report no. 20, pages 443–455, Vienna, 2001. ARGESIM / ASIM - Verlag.

[4] S.-C. Liu, J. Kramer, G. Indiveri, T. Delbruck, T. Burg, and R. Douglas. Orientation-selective aVLSI spiking neurons. *Neural Networks*, 14(6/7):629–643, 2001. Special Issue on Spiking Neurons in Neuroscience and Technology.

[5] E. Chicca, D. Badoni, V. Dante, M. D'Andreagiovanni, G. Salina, S. Fusi, and P. Del Giudice. A VLSI recurrent network of integrate–and–fire neurons connected by plastic synapses with long term memory. Submitted.

[6] W. Maass and E. D. Sontag. Neural systems as nonlinear filters. *Neural Computation*, 12(8):1743–72, 2000.

[7] R. S. Zucker and W. G. Regehr. Short–term synaptic plasticity. *Annual Reviews Physiology*, 64:355–405, 2002.

[8] A. M. Zador and L. E. Dobrunz. Dynamic synapses in cortex. *Neuron*, 19:1–4, 1997.

[9] L. F. Abbott, J. A. Varela, K. Sen, and S. B. Nelson. Synaptic depression and cortical gain control. *Science*, 275:220–4, 1997.

[10] F. S. Chance and L. F. Abbott. Input-specific adaptation in complex cells through synaptic depression. *Neurocomputing*, 38(40):141–46, 2001.

[11] F. S. Chance, S. B. Nelson, and L. F. Abbott. Synaptic depression and the temporal response characteristics of v1 cells. *The Journal of Neuroscience*, 18(12):4785–99, 1998.

[12] C. Rasche and R. Hahnloser. Silicon synaptic depression. *Biological Cybernetics*, 84:57–62, 2001.

[13] S.-C. Liu. Simple cortical modeling with aVLSI spiking neurons and dynamic synapses. In *ZNZ Symposium*. University of Zurich, Zurich, 2001. http://www.ini.unizh.ch/projects/posters/shih01.pdf.

[14] K. Hynna and K. Boahen. Space–rate coding in an adaptive silicon neuron. *Neural Networks*, 14:645–656, 2001.

[15] M. Mahowald. *An Analog VLSI System for Stereoscopic Vision*. Kluwer, Boston, 1994.

[16] J. Lazzaro, J. Wawrzynek, M. Mahowald, M. Sivilotti, and D. Gillespie. Silicon auditory processors as computer peripherals. *IEEE Trans. on Neural Networks*, 4:523–528, 1993.

[17] K.A. Boahen. Communicating neuronal ensembles between neuromorphic chips. In T. S. Lande, editor, *Neuromorphic Systems Engineering*, pages 229–259. Kluwer Academic, Norwell, MA, 1998.

[18] S.-C. Liu, J. Kramer, G. Indiveri, T. Delbruck, and R. J. Douglas. *Analog VLSI: Circuits and Principles*. MIT Press, 2002.

[19] G. Indiveri. Modeling selective attention using a neuromorphic analog VLSI device. *Neural Computation*, 12(12):2857–2880, Dec 2000.

[20] A. G. Andreou, K. A. Boahen, P. O. Pouliquen, A. Pavasovič, R. E. Jenkins, and K. Strohbehn. Current–mode subthreshold MOS circuits for analog VLSI neural system. *IEEE Transaction on Neural Network*, 2(2):205–13, 1991.