Programmable Neuromorphic Circuits for Spike-based Neural Dynamics

Mostafa Rahimi Azghadi[*†], Saber Moradi[†], and Giacomo Indiveri[†]
[*]School of Electrical and Electronic Engineering, The University of Adelaide, Australia.
Email: mostafa@eleceng.adelaide.edu.au
[†]Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland.
Email: [mostafa|saber|giacomo]@ini.phys.ethz.ch

Abstract—Hardware implementations of spiking neural networks offer promising solutions for a wide set of tasks, ranging from autonomous robotics to brain machine interfaces. We propose a set of programmable hybrid analog/digital neuromorphic circuits than can be used to build compact low-power neural processing systems. In particular, we present both CMOS and hybrid memristor/CMOS synaptic circuits that have programmable synaptic weights and exhibit biologically plausible response properties. For the CMOS circuits, we present experimental results demonstrating that they operate correctly over a wide range input frequencies; for the hybrid memristor/CMOS circuits we present circuit simulation results validating their expected response properties.

I. Introduction

Artificial spike-based neural networks offer a promising paradigm for a new generation of brain-inspired computational models. A wide range of theoretical and computational models have already been proposed for both basic neuroscience research [1], [2] and practical applications [3], [4]. Neuromorphic Very Large Scale Integration (VLSI) circuits represent an ideal technology for implementing these types of networks using hybrid analog/digital design techniques, and for building devices that have a very high potential in a wide range of applications [5]. In particular, the main advantage of implementing these spiking neural networks in neuromorphic VLSI technology is their compactness and low power consumption which are critical features when implementing a large scale neural architectures [6].

Synapses are an essential component of spiking neural networks. They represent at the same time the site of memory (as they store the network's synaptic weight values), and play a fundamental role in computation (as they implement crucial temporal and non-linear dynamics). In these networks the synaptic weight is directly associated with the activity of pre-synaptic and post-synaptic neurons [1]. Different types of learning algorithms have been proposed, to update the synaptic weight depending on pre- and post-synaptic activity [7]–[9]. The different learning strategies have a profound effect on the post-synaptic neuron functionality and on the spiking-neural network behavior [10]. Implementing such types of synapses and learning mechanisms in compact electronic systems is crucial, for developing efficient large-scale spiking neural networks and brain-inspired computing technologies.

In this paper we present a set of programmable neuromorphic synapse circuits that have been fabricated using standard Complementary Metal Oxide Semiconductor (CMOS) VLSI process [11] and that can support different weight-update mechanisms and learning strategies. Furthermore we propose novel hybrid synapse circuits that apply the same principles used in the programmable CMOS synapses with nano-scale memristive devices [12]–[15].

Specifically, we present a set of experimental results measured from the fabricated synapse circuits demonstrating how they can be calibrated to a specific targeted behavior, and show simulation results for hybrid memristor/CMOS devices demonstrating how they would produce results analogous to the ones obtained with conventional CMOS technology.

The paper is organized as follows: Section II briefly describes the architecture of the neuromorphic chip that implements the programmable CMOS synapse circuits. The experimental results measured from the chip synapse circuits, are shown in Section III. Section IV shows the simulated response properties of the hybrid CMOS/memristor synapse circuit. Finally section V is the concluding remarks of this paper.

II. A spiking multi-neuron CMOS device

The photo micro-graph of the chip comprising the programmable CMOS synapse circuits is illustrated in Fig. 1a. The chip implements a spiking neural network of 32 adaptive exponential Integrate-and-Fire (I&F) neuron circuits [16] with dynamic synapse circuits. All analog circuits on the chip have programmable bias parameters that can be set with an on-chip 32-bit temperature-compensated programmable bias generator [17]. In addition there are 32×32 5-bit digital Static Random Access Memory (SRAM) cells with asynchronous interfacing circuits for storing the weight values of the neural network synapses. The weights can be altered by off-chip algorithms that implement different types of synaptic plasticity rules, such as Spike-Timing Dependent Plasticity (STDP) [18]–[20]. The updated weights can be then written back on-chip, to the synaptic weight SRAM memory cells, to change the post-synaptic neuron behavior, for solving cognitive tasks like pattern recognition and classification.

The on-chip synapse circuits integrate incoming spikes and produce Excitatory Post Synaptic Currents (EPSCs) with
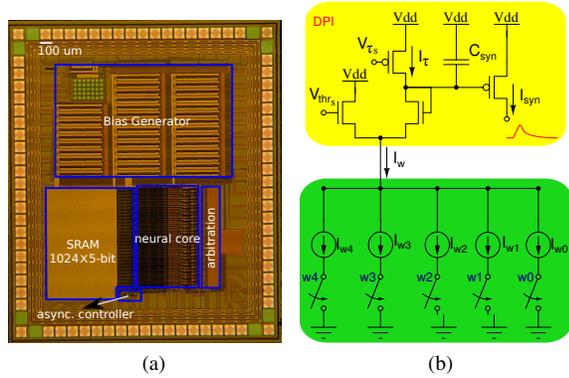
(a)   (b)

Fig. 1: (a) Chip micro-graph. The multi-neuron chip was fabricated using a standard $0.35\,\mu$m CMOS technology and occupies an area of 2.1×2.5 mm^2. The programmable synapses are integrated inside the neural core block. (b) Schematic diagram of the programmable synapse circuit. The top part of the diagram represents a DPI circuit which implements the temporal dynamics. The bottom part of the diagram represents the DAC that converts the SRAM 5-bit weight into a corresponding synaptic current.

amplitudes proportional to their corresponding stored weights. The temporal response properties of the circuit exhibit dynamics that are bio-physically realistic and have biologically plausible time constants [21]. The part of the synapse circuit that produces the slow temporal dynamics is the log-domain Differential Pair Integrator (DPI) filter [22], [23], shown in Fig. 1b. By using the DPI in its linear regime it is possible to time-multiplex the contributions from multiple spiking inputs (e.g., via multiple SRAM cells), thus requiring one single integrating element and saving precious silicon real-estate. This time multiplexing scheme, and circuits implemented on the chip of Fig. 1a have been fully characterized in [11], while the description of the DPI synapse dynamics has been presented in [22].

Here we characterize the response properties of the synapse circuits as a function of input spike frequency and of programmable weight values, and evaluate their linear characteristics and dynamic range properties.

III. PROGRAMMABLE CMOS SYNAPSE RESPONSE PROPERTIES

During the experiments carried out in this section we employed one single post-synaptic neuron and one single input synapse. We characterized the response properties of the combined synapse-neuron circuit by sending input spikes to the synapse, and measuring output spikes from the neuron. The synapse circuit integrates input spikes to produce an output current (the synapse EPSC) that has a mean steady-state amplitude which depends on both the input spike train frequency and its synaptic weight. The silicon neuron integrates this current and produces an output spike train with a mean firing rate that is linear with the current [16] (see
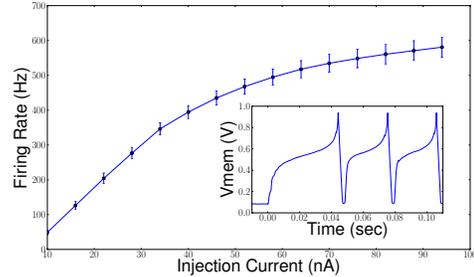

Fig. 2: Silicon neuron response properties. Mean firing rates of all neurons on the chip, as a function of input current. The figure inset shows the membrane potential of a single neuron.

Fig. 2). Analogous to software models of biological neurons and synapses, the hardware neuron and synapse circuits have multiple parameters that control the way they function. In order to optimize the use of the neuron and synapse circuits in various applications like computation, brain machine interface, pattern recognition etc., it is necessary to tune their analog bias parameters to a set of specific values that results in the required features and expected behavior of the neuromorphic chip. In Fig. 3 we show how it is possible to optimize the circuit biases for a specific range of pre-synaptic frequencies, so that the combined synapse-neuron circuits respond linearly to their afferent synaptic inputs, for all possible weight values that can be programmed. Under this condition, the neuron is able to show identical (gain=1) or linearly higher (gain>1) or lower (gain<1) post-synaptic output frequency, compared to afferent pre-synaptic input frequency. Here gain is defined as the fraction of post-synaptic to pre-synaptic firing rates in the highest synaptic weight setting (w=31). In Fig. 3a we show the circuit response properties using a set of bias parameters optimized for biologically plausible firing rates. Figure 3b shows analogous results, but with parameters optimized for very high firing rates (e.g., for applications that need to process incoming data quickly, and for neuromorphic systems that do not need to interact with the environment). When there are synaptic inputs with various firing rates, the neuron and synapses should be tuned to act linearly for the whole possible input firing range of frequencies. In the calibration of the bias values in our circuit, the main sets of parameters that were tuned are those related to the five synaptic currents $I_{w0} - I_{w4}$ depicted in Fig. 1b. Those parameters were tuned in a way to establish a linear relationship between the 32-state digital synaptic weight values and the neuron post-synaptic firing rates.

IV. HYBRID CMOS/MEMRISTOR SYNAPSE CIRCUIT

Memristors are nano-electronic devices characterized by a variable resistance which changes dynamically as the device is being used [12], [13]. Unlike active CMOS circuits, memristors can store their internal state without dissipating power. Given their extremely compact size and their extremely low-power characteristics, memristive devices are ideal elements
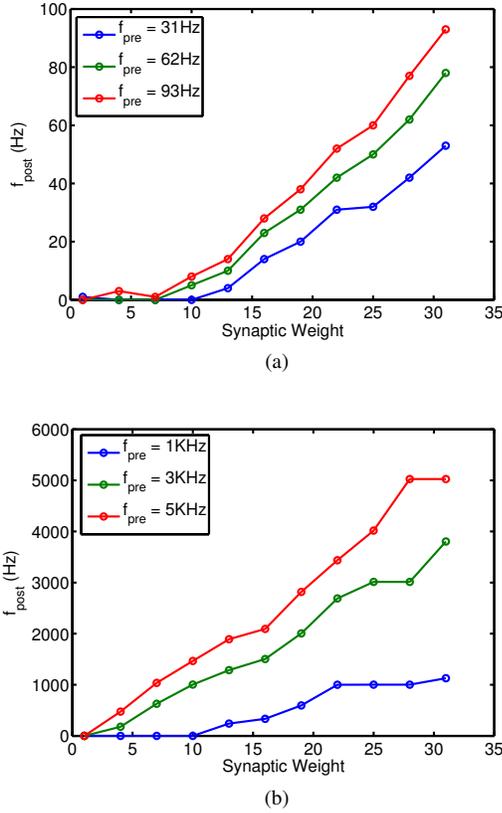
(a)

(b)

Fig. 3: Optimizing the analog bias parameters to determine a linear relationship between pre-synaptic input and post-synaptic output spiking activities. (a) Regular spike trains with pre-synaptic input frequencies of 31, 62, and 93 Hz are applied to the synapse-neuron circuit, for different synaptic weight settings. (b) Analogous experiment to the one carried out in (a), but with pre-synaptic input frequencies of 1, 3, and 5 KHz.

for implementing neuromorphic synapse circuits [15]. By designing hybrid nano-electronic CMOS circuits it is possible to obtain dense integration of low-power, long-term synaptic weight storage elements interfaced to circuits that emulate detailed synaptic biophysics and implement relevant computational properties of neural systems. Here we propose a novel CMOS/memristor synapse circuit analogous to the one depicted in Fig. 1b, but in which compact low-power memristive devices replace the tunable constant current sources and the programmable synaptic weight digital switches.

The hybrid CMOS/memristor circuit is depicted in Fig. 4. It shows a possible implementation of an array of N synaptic weight elements with independent inputs, that integrate their input spikes via a shared DPI circuit. The V_w bias voltage is common to all synaptic inputs, and sets the maximum current that can be produced by each memristor upon the arrival of an input spike. At each synaptic input, the memristor conductance modulates the current being produced by the synapse very much like conductance changes in real synapses affect the EPSCs they produce. Larger memristor conductances
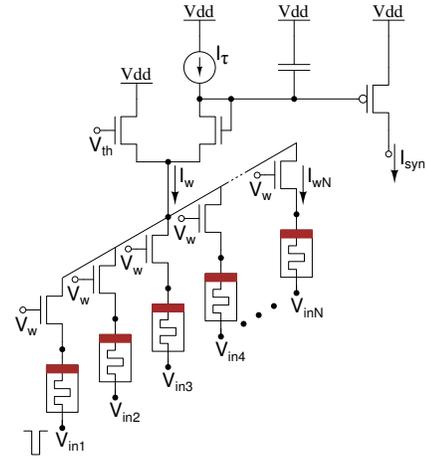

Fig. 4: Neuromorphic memristive synapse circuit diagram. The circuit integrates output currents from an array of memristive elements. The inputs and synaptic weights are independent, but the temporal dynamics are shared.

correspond to larger synaptic weights.

Figure 5 shows the results of SPICE simulations of the circuit in Fig. 4. The simulations were done with 180 nm CMOS process models. The V_{thr} bias voltage was set to 0V and the I_τ current (implemented by a p-type Metal Oxide Semiconductor Field Effect Transistor (MOSFET)) was set to 80 pA. The data was obtained by simulating only one input memristive branch, and sweeping the memristor impedance from 7 MΩ to 1 MΩ, with the V_w voltage bias set to 600 mV. In these simulations we set the memristor in its High-Resistive State (HRS), and assumed we could modulate the value of the resistance to obtain four distinct analog states [24], [25]. The circuit supports also the operation of the memristor as a binary device working in either state (HRS or Low-Resistive State (LRS)). This bi-stable mode of using the memristor is compatible with biologically plausible learning mechanisms, such as those proposed in [8], and implemented in [26]. The proposed circuit shows only the elements required for a "non-plastic" operation, in which the synapse is stimulated to produce an EPSC with an amplitude set by the conductance of the memristor, but without necessarily changing the conductance value in the process. Additional circuit elements would be required for learning mechanisms that change the value of the memristor's conductance. These additional circuit elements would be quite compact, as the large part of the learning circuits could be implemented at the Input/Output (I/O) periphery of the synaptic array, for example with pulse-shaping circuits and architectures analogous to the ones described in [27], or with circuits that check the state of the neuron and of it's recent spiking history, such as those proposed in [23].

V. CONCLUSIONS

We presented two different hardware implementations of programmable synapse circuits, in CMOS and nano-electronic technologies. In the standard CMOS design, the synaptic
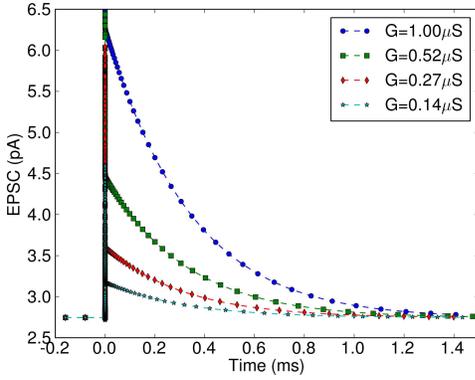
Fig. 5: SPICE simulation results of the hybrid CMOS/memristor synapse. Output current I_{syn} in response to a pre-synaptic input spike, for 4 different values of memristor conductance.

weights are stored in digital SRAM circuits, while in the hybrid CMOS/memristor design they are represented by the conductance of the memristor devices. Due to their small size, memristor devices are expected to provide higher density compared to CMOS memory circuits. We showed in this work how existing spiking neural network designs are compatible with hybrid CMOS/memristor designs, by providing both experimental results and circuit simulations. Furthermore, we focused our efforts in designing circuits that are biophysically realistic and can provide biologically plausible time-constants if required. In addition to reproducing the biologically plausible temporal dynamics, both designs have programmable synaptic weights, an important property of biological synapses and a crucial requirement for learning in artificial neural networks. We showed how the proposed neuron-synapse circuits can be tuned to respond appropriately for different ranges of output firing rates, which makes the design usable in different types of applications. This, together with the circuit compactness and low-power consumption features makes the proposed designs suitable candidates for implementing large scale spiking neural networks for real-world applications. More generally, the circuits proposed represent a useful building block for research on learning in hardware spiking neural networks, built using both standard CMOS technologies and more advanced nano-technologies.

VI. Acknowledgments

References

[1] R. Kempter, W. Gerstner, and J. van Hemmen, "Hebbian learning and spiking neurons," *Physical Review E*, vol. 59, no. 4, pp. 4498–4514, 1999.

[2] W. Gerstner and W. Kistler, *Spiking Neuron Models. Single Neurons, Populations, Plasticity*. Cambridge University Press, 2002.

[3] A. Belatreche, L. P. Maguire, and M. McGinnity, "Advances in design and application of spiking neural networks," *Soft Computing*, vol. 11, no. 3, pp. 239–248, Jan 2006.

[4] P. Rowcliffe and J. Feng, "Training spiking neuronal networks with applications in engineering tasks," *Neural Networks, IEEE Transactions on*, vol. 19, no. 9, pp. 1626–1640, Sept. 2008.

[5] G. Indiveri and T. K. Horiuchi, "Frontiers in neuromorphic engineering," *Frontiers in Neuroscience*, vol. 5, no. 118, 2011.

[6] C. Mead, "Neuromorphic electronic systems," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1629–36, 1990.

[7] L. Abbott and W. Gerstner, "Homeostasis and learning through spike-timing dependent plasticity," 2004.

[8] J. Brader, W. Senn, and S. Fusi, "Learning real world stimuli in a neural network with spike-driven synaptic dynamics," *Neural Computation*, vol. 19, pp. 2881–2912, 2007.

[9] M. Graupner and N. Brunel, "Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location," *Proceedings of the National Academy of Sciences*, 2012.

[10] S. Laughlin and T. Sejnowski, "Communication in neuronal networks," *Science*, vol. 301, no. 5641, pp. 1870–1874, 2003.

[11] S. Moradi and G. Indiveri, "A VLSI network of spiking neurons with an asynchronous static random access memory," in *Biomedical Circuits and Systems Conference (BioCAS), 2011*. IEEE, 2011, pp. 277–280.

[12] L. Chua, "Resistance switching memories are memristors," *Applied Physics A*, vol. 102, no. 4, pp. 765–783, Jan. 2011.

[13] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 459, no. 7250, pp. 1154–1154, June 2009.

[14] T. Prodromakis, C. Toumazou, and L. Chua, "Two centuries of memristors," *Nature Materials*, vol. 11, no. 6, pp. 478–481, 2012.

[15] S. H. Jo, *et al.*, "Nanoscale Memristor Device as Synapse in Neuromorphic Systems," *Nano Letters*, vol. 10, no. 4, pp. 1297–1301, Apr. 2010.

[16] G. Indiveri, *et al.*, "Neuromorphic silicon neuron circuits," *Frontiers in Neuroscience*, vol. 5, pp. 1–23, 2011.

[17] T. Delbruck, R. Berner, P. Lichtsteiner, and C. Dualibe, "32-bit configurable bias current generator with sub-off-current capability," in *International Symposium on Circuits and Systems, (ISCAS), 2010*. IEEE, 2010, pp. 1647–1650.

[18] G. Bi and M. Poo, "Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type," *The Journal of Neuroscience*, vol. 18, no. 24, pp. 10 464–10 472, 1998.

[19] M. R. Azghadi, S. Al-Sarawi, N. Iannella, and D. Abbott, "Design and implementation of BCM rule based on spike-timing dependent plasticity," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2012, pp. 1–7.

[20] ——, "Efficient design of triplet based spike-timing dependent plasticity," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2012, pp. 1–7.

[21] C. Bartolozzi and G. Indiveri, "Synaptic dynamics in analog VLSI," *Neural Computation*, vol. 19, no. 10, pp. 2581–2603, Oct 2007.

[22] C. Bartolozzi, S. Mitra, and G. Indiveri, "An ultra low power current–mode filter for neuromorphic systems and biomedical signal processing," in *Biomedical Circuits and Systems Conference, (BioCAS), 2006*. IEEE, 2006, pp. 130–133.

[23] S. Mitra, G. Indiveri, and R. Etienne-Cummings, "Synthesis of log-domain integrators for silicon synapses with global parametric control," in *International Symposium on Circuits and Systems, (ISCAS), 2010*. IEEE, 2010, pp. 97–100.

[24] T. Prodromakis, B. P. Peh, C. Papavassiliou, and C. Toumazou, "A Versatile Memristor Model With Nonlinear Dopant Kinetics," *Electron Devices, IEEE Transactions on*, vol. 58, no. 9, pp. 3099–3105, 2011.

[25] X. Yang and I.-W. Chen, "Dynamic-load-enabled ultra-low power multiple-state rram devices," *Scientific Reports*, vol. 2, 2012.

[26] S. Mitra, S. Fusi, and G. Indiveri, "Real-time classification of complex patterns using spike-based learning in neuromorphic VLSI," *Biomedical Circuits and Systems, IEEE Transactions on*, vol. 3, no. 1, pp. 32–42, Feb. 2009.

[27] C. Zamarreño-Ramos, *et al.*, "On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex," *Frontiers in neuroscience*, vol. 5, 2011.