Neural Nets WIRN11
B. Apolloni et al. (Eds.)
IOS Press, 2011

doi:10.3233/978-1-60750-972-1-305

A VLSI neuromorphic device for implementing spike-based neural networks

Giacomo INDIVERI [a,1] and Elisabetta CHICCA [a,b]
[a] *Institute of Neuroinformatics, University of Zurich and ETH Zurich*
Winterthurerstrasse 190, Zurich, Switzerland
[b] *Cognitive Interaction Center of Excellence (CITEC), University of Bielefeld*
Universitätstrasse 21-23, Bielefeld, Germany

Abstract. We present a neuromorphic VLSI device which comprises hybrid analog/digital circuits for implementing networks of spiking neurons. Each neuron integrates input currents from a row of multiple analog synaptic circuit. The synapses integrate incoming spikes, and produce output currents which have temporal dynamics analogous to those of biological post synaptic currents. The VLSI device can be used to implement real-time models of cortical networks, as well as real-time learning and classification tasks. We describe the chip architecture and the analog circuits used to implement the neurons and synapses. We describe the functionality of these circuits and present experimental results demonstrating the network level functionality.

Keywords. Neuromorphic circuits, Integrate-and-Fire (I&F) neuron, synapse, Winner-Take-All (WTA), Address-Event Representation (AER), spike-based plasticity, STDP, learning

Introduction

With the technological advancements in both conventional computing architectures and custom Very Large Scale Integration (VLSI) implementations, spiking neural networks have been gaining renewed interest in recent years [9,16,20,28,29,34]. Hardware implementations of spiking neural networks can be useful tools for basic research investigations (*e.g.*, by computational neuroscientists), and for exploring the implementation of alternative classes of brain-inspired general-purpose computational architectures. Examples of devices recently proposed to build hardware networks of spiking neurons range from reconfigurable arrays of Integrate-and-Fire (I&F) neuron models [11,12,27,33], to learning architectures implementing detailed models of spike-based synaptic plasticity [2,3,10,26,27,36,39].

Within this context we propose a neuromorphic VLSI device that comprises an array of I&F silicon neurons and a matrix of synapse circuits that exhibit biologically realistic synaptic dynamics and implement a spike-driven learning mechanism. The silicon

[1] Corresponding Author; E-mail: giacomo@ini.phys.ethz.ch.

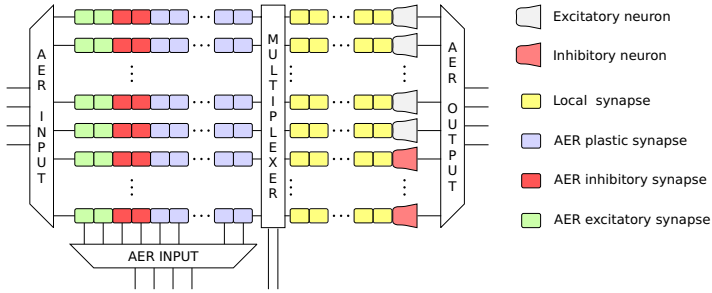

Figure 1. IFSLWTA chip diagram: the chip comprises a linear array of 128 neuron circuits, each connected to a row of afferent synapse circuits. Each neuron is connected to 40 synapses, subdivided into four classes. Local synapses can be used to implement WTA architectures, and the last four neurons can be configured as inhibitory neurons (see text for details).

neurons and synapses inside the chip are implemented using low-power analog circuits. The neurons are connected in a soft Winner-Take-All (WTA) arrangement via a set of local synaptic circuits, but comprise also externally addressable excitatory, inhibitory, and plastic synapses. The spike-based plasticity circuits are also analog, and implement a learning algorithms based on the work proposed in [8,21]. Conversely, the peripheral circuits used to transmit spikes into and out of the chip use self-clocked asynchronous digital logic. The spiking neuron circuits, the dynamic synapses and the learning circuits implemented in this device allow us to design spike-based neural networks and architectures for signal processing, recognition and classification. The local hardwired connectivity allows us to implement models of cooperative-competitive (soft WTA) networks, while the real-time asynchronous digital communication infrastructure allows us to design arbitrary network topologies. In addition the digital protocol used allows us to interface multiple chips together and to build large-scale complex neural systems. The neural systems built using this approach can process sensory signals, adapt to the statistics of their inputs, learn and classify complex sequences of spatio-temporal patterns, and eventually interact with the user and the environment in real-time.

In the next sections we describe the architecture of the VLSI device, of the neuron and synapse circuits, of the learning circuits, and present experimental results measured from both single circuits inside the chip, and at the system level.

1. System description

1.1. The I&F Stop-Learning Winner-Take-All (IFSLWTA) chip

The I&F Stop-Learning Winner-Take-All (IFSLWTA) chip comprises an array of 128 silicon neurons and 5120 synaptic circuits in total. The neurons are arranged in 128 rows, each containing 40 synapses (see Fig. 1). The chip was produced using a standard 0.35μm CMOS technology and it occupies an area of 10 mm^2. Within each row, the synapse circuits are subdivided into four main groups: the first group has 8 synapses, locally connected (hardwired); the second group has 28 synapses, externally addressable, with spike-based learning circuits; a third group has 2 inhibitory synapses with fixed (tunable) weights and externally addressable; and the fourth group has 2 externally ad-

dressable excitatory synapses, with fixed weight and short-term depression properties. The last four neurons in the array can be used as inhibitory neurons, by activating the corresponding local synapses. The local hardwired connectivity is used to implement soft WTA topologies. Indeed, soft WTA networks of spiking neurons typically consist of a group of interacting neurons which compete with each other in response to an input stimulus. The neurons with highest response suppress all other neurons to win the competition. Competition is achieved through a recurrent pattern of connectivity involving both excitatory and inhibitory connections. Cooperation between neighboring neurons is mediated by excitatory connections, while suppression is mediated by inhibitory neurons that project to all other excitatory neuron. The detailed network connectivity implemented in the I&F Stop-Learning Winner-Take-All (`IFSLWTA`) is described in Section 1.3.1.

A multiplexer circuit placed between the externally addressable synapses and the neurons allows the user to re-route the synaptic inputs. In the default state (*i.e.*, with each control signal set to ground) each row of synapses is connected to the corresponding neuron in the same row. However, with different control signal settings, multiple synapse rows can be merged together. For example, by merging pairs of rows, the chip will have 64 neurons, with 64 input synapses each (the remaining 64 neurons remain disconnected, without synaptic inputs); by merging four rows at a time, there will be 32 neurons, each with 128 inputs; by merging 8 rows, there will be 16 neurons, each with 256 inputs, and so forth, until there is one neuron with 4096 inputs. This multiplexing scheme offers therefore additional flexibility that allows users to choose the number of perceptrons and the number of input synapses per perceptron, in their application.

The Input/Output (I/O) peripheral circuits implement an asynchronous communication protocol based on the Address Event Representation (AER) [7,14,32]. In AER, each neuron on a sending device is assigned an address. When the neuron produces a spike its address is instantaneously put on an asynchronous digital bus. Event 'collisions' (cases in which sending nodes attempt to transmit their addresses at exactly the same time) are managed by an on-chip arbitration circuit, within the "AER output" block of Fig. 1. Systems containing more than two AER chips can be constructed by implementing special purpose off-chip arbitration schemes, *e.g.*, using Field Programmable Gate Array (FPGA) devices [11,19].

As the neuron and synapse circuits can be biased to produce biologically realistic time constants, the chip can be used in behaving systems to process and classify sensory signals generated by other AER neuromorphic sensors, and produce motor outputs in real-time.

1.2. The synapse and neuron circuits

Input spike patterns are provided to the synapses via the asynchronous AER input blocks (see Fig. 1). Each synapse circuit generates an output current proportional to the mean rate of its input spike train, modulated by its synaptic weight. The I&F neurons integrate the sum of the input currents produced by the synapses and generate output spike trains with mean firing rates proportional to their total input current.

All synapses of all four groups of Fig. 1 use the same type of output circuit: the Differential Pair Integrator (DPI) circuit [4]. This circuit, shown in Fig. 2a is a log-domain linear temporal filter which implements biologically realistic synaptic dynam-

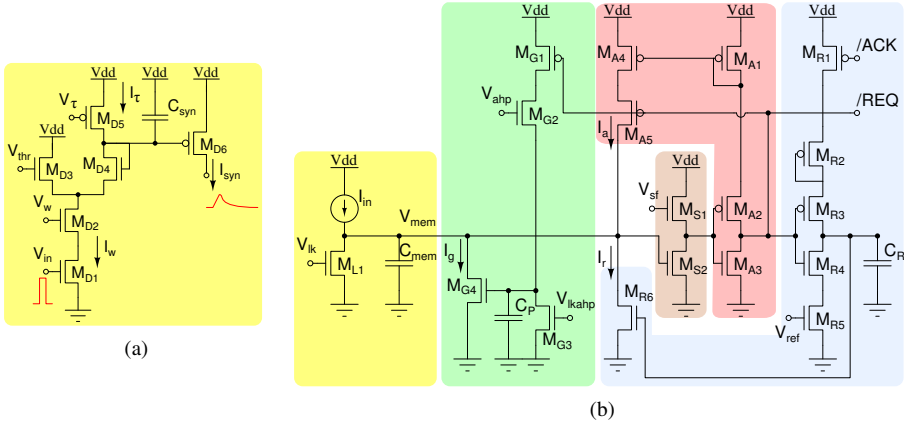

Figure 2. (a) Schematic diagram of a DPI synapse: input voltage pulses are integrated, and output currents are produced with biologically realistic dynamics. The synaptic weight can be modulated by changing the V_w and V_{thr} biases, while the circuit's time constant is set with the V_τ bias. (b) Schematic diagram of a leaky integrate-and-fire neuron. The input current I_{in} is integrated onto the neuron's membrane capacitor C_{mem} until the spiking threshold is reached. At that point a spike is emitted, the membrane capacitor is reset to zero, and the input current starts to be integrated again. The MOSFET M_{L1} implements the "leak" module. The "spiking threshold" module (MOSFETs M_{S1-2}) modulates the voltage at which the neuron spikes. The "adaptation" module (M_{G1-4}) subtracts a firing rate dependent current from the input node. The amplitude of this current increases with each output spike and decreases exponentially with time. The "refractory period" module (M_{R1-6}) sets a maximum firing rate for the neuron. The "positive feedback" module (M_{A1-5})is activated when the neuron begins to spike, and is used to reduce the transition period in which the inverters switch polarity, dramatically reducing power consumption. The neuron's properties can be set by changing the circuit's bias parameters (V_{lk}, V_{ahp}, V_{lkahp}, V_{sf}, and V_{ref}).

ics, following the model proposed in [15]. It supports a wide range of synaptic properties, ranging from short-term depression to conductance-based Excitatory Post Synaptic Current (EPSC) generation.

The local excitatory synapses of the device (the ones belonging to the fourth group, in Fig. 1) are implemented directly with the DPI circuit shown in Fig. 2a. The externally addressable excitatory synapses (the ones in the first group of synapses in Fig. 1) are implemented with an instance of the Differential Pair Integrator (DPI) interfaced to a circuit that implements *short-term depression* of the synaptic weight, as described in [4]. Both the local and the externally addressable inhibitory synapses are implemented with a complementary version of the circuit shown in Fig. 2a (*e.g.*, the differential pair is made with pFETs, the output transistor is an nFET, *etc.*) The plastic synapses comprise both the DPI of Fig. 2a as well as the spike-driven learning circuits described in Section 1.3.

The neuron blocks of Fig. 1 comprise both an I&F "soma" circuit, responsible for the spike generation, as well as additional spike-based learning circuits. Figure 2b shows the soma I&F circuit diagram. This circuit has been first proposed and characterized in [27]. It can be subdivided into five main functional blocks: an input/leak block (composed by the nFET M_{L1} and the membrane capacitor C_{mem}); a spiking threshold modulation block (M_{S1-2}); a spike generation block (M_{A1-5}); a reset and refractory period block (M_{R1-6}); and a spike-frequency adaptation block (M_{G1-4}). If the input current I_{in} is larger than the leak current set by V_{lk} of M_{L1}, then the voltage V_{mem} on the neuron's membrane capacitor C_{mem} increases linearly until it approaches the spiking threshold. At that point

the spike generation block M_{A1-5} starts to inject an additional current which increases V_{mem} even more. This positive feedback quickly brings V_{mem} above the spiking threshold and makes the inverter M_{A2-3} switch. As a consequence the output signal $/REQ$ goes from V_{dd} to zero, signaling the occurrence of a spike. Once the AER I/O circuits acknowledge the neuron's request to transmit the spike, the neuron is reset to zero and is clamped to the zero potential for a *refractory period* set by V_{ref}. Once M_{R6} is switched off again, the integration cycle repeats. The spiking threshold modulation block M_{S1-2} modulates the voltage at which the neuron spikes. The spike-frequency adaptation block M_{G1-4} subtracts a firing rate dependent current from the input node. The amplitude of this current increases with each output spike and decreases exponentially with time. This is a negative feedback effect that makes the neuron act as a high-pass filter. Conversely, the positive feedback module M_{A1-5} has the effect of speeding-up the action potential generation, thus reducing the inverter switching period and the circuit's power consumption substantially. All the circuit's biases (V_{lk}, V_{ahp}, V_{lkahp}, V_{sf}, and V_{ref}) are subthreshold voltages that can be used to change the neuron's firing properties.

Figure 3 shows experimental results taken by sending spikes to the AER synapses and measuring the membrane potential and output spikes from the afferent neuron. Specifically, the data of Fig. 3a shows how the neuron integrates input spikes sent to one of its excitatory synapses at a rate of 100 Hz . After the membrane potential is reset, the neuron is kept silent for a refractory period set by V_{ref} of Fig. 2b. The data of Fig. 3b shows the effect of an inhibitory synapse on the membrane potential of a neuron receiving a constant current. The inhibitory synapse comprises also a conductance-based circuit of the type described in [4]. In this experiment the conductance resting potential was set at 0.4V. As the membrane potential rises above this threshold, the inhibitory synapse starts to be effective, and each input spike, represented by the vertical dashed lines, induces a downward jump. The data of Fig. 3c shows the effect of the input excitatory synapse when the short-term depression circuits are activated. In this case, when the synapse is stimulated with a series of spikes, the synaptic weight decreases with each spike. The data shows the effect of short-term depression for three different values of depression rate. Finally, Fig. 3d shows the neuron's mean firing rate (averaged across all the neurons on the chip) in response to spike trains of increasing frequency sent to the first non-plastic synapse of each neuron. The standard deviation showed in the shaded region around the mean response shows the intrinsic degree of inhomogeneities (mismatch) present in the chip, due to the fabrication process [30].

In addition to the types of synaptic dynamics described in Fig. 3, the group of plastic synapses of Fig. 1 comprise also weight update circuits that implement long-term changes in their synaptic weight voltages, as described in the following section.

1.3. The spike-based learning circuits

The spike-based learning algorithm implemented in this device is based on the model described in [8]. This algorithm can be used to implement supervised learning protocols, and train neurons to act as perceptrons or binary classifiers. Input patterns are encoded as sets of spike trains with different mean frequencies, while the neuron's output firing rate represents the binary classifier output. The learning circuits that implement this algorithm can be subdivided into two main blocks: a spike-triggered weight-update module with bistable weights, present in each plastic synapse, and a post-synaptic stop-learning

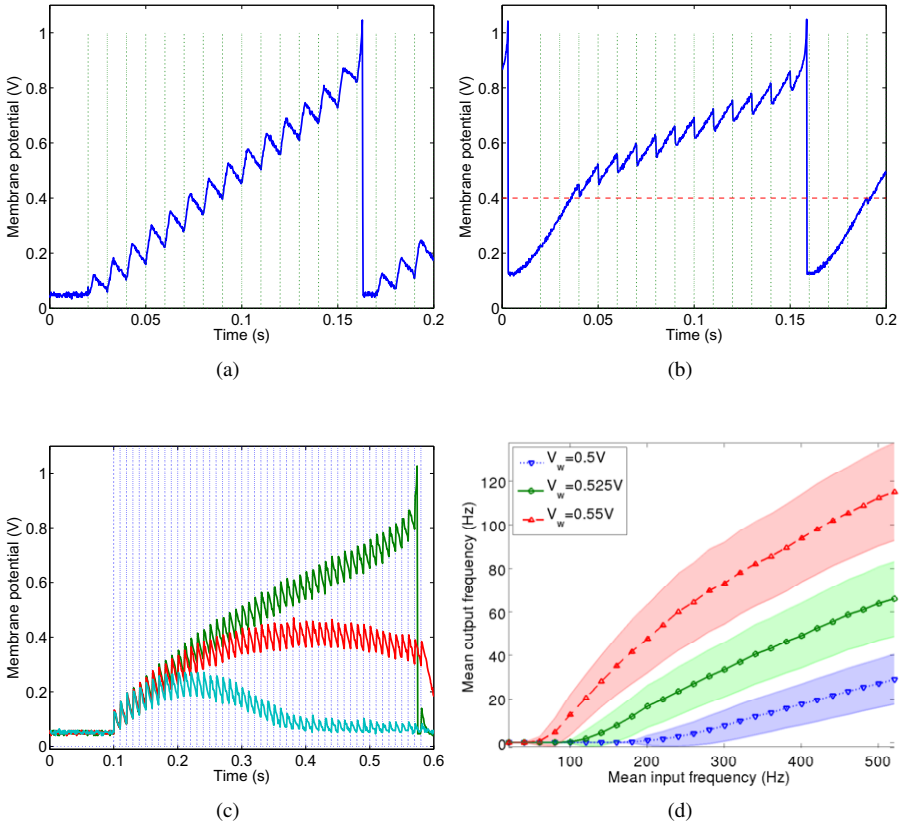

Figure 3. (a) Membrane potential of I&F neuron in response to a 100 Hz input spike train sent to an excitatory synapse; (b) Neuron response to a constant injection current and an inhibitory conductance-based synapse, stimulated by a 100 Hz spike train; (c) Neuron membrane potential recorded when stimulating a short-term depressing synapse with a 100 Hz spike train, for different values of adaptation rate; (d) Mean output frequency of all neurons on the chip, in response to regular pre-synaptic spike-trains of increasing frequency sent to the excitatory synapse, for three different values of the synaptic weight. The shaded areas represent the standard deviation ranges.

control module, present in the neuron's soma. The 'stop-learning' circuits implement the characteristic feature of this algorithm, which stops updating weights if the output neuron has a very high or very low output firing rate, indicating the fact that the dot product between the input vector and the learned synaptic weights is either close to one (pattern recognized as belonging to the trained class) or close to zero (pattern not in trained class).

The post-synaptic stop-learning control circuits are shown in Fig. 4b. These circuits produce two global signals V_{UP} and V_{DN}, shared among all synapses belonging to the same dendritic tree, to enable positive and/or negative weight updates respectively. Post-synaptic spikes produced by the I&F neuron are integrated by a DPI circuit ($M_{D1} - M_{D5}$ of Fig. 4b). The DPI produces the signal V_{Ca} which is related to the Calcium concentration in real neurons, and represents the recent spiking activity of the neuron. This signal is compared to three different thresholds (V_{thk1}, V_{thk2}, and V_{thk3}) by three corresponding current-mode winner-take-all circuits [31]. In parallel, the neuron's membrane potential

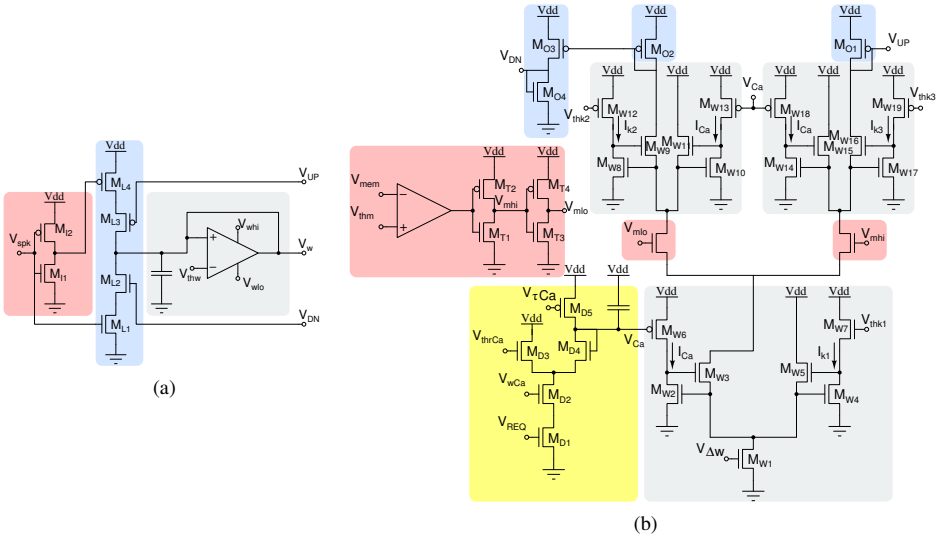

Figure 4. Spike-based learning circuits. (a) Pre-synaptic weight-update module (present at each synapse). (b) Post-synaptic stop-learning control circuits (present at the soma).

V_{mem} is compared to a fixed threshold V_{thm} by a transconductance amplifier. The values of V_{UP} and V_{DN} depend on the output of this amplifier, as well as the Calcium concentration signal V_{Ca} . Specifically if $V_{thk1} < V_{Ca} < V_{thk3}$ and $V_{mem} > V_{mth}$ then increases in synaptic weights ($V_{UP} < V_{dd}$) are enabled. And if $V_{thk1} < V_{Ca} < V_{thk2}$ and $V_{mem} < V_{thm}$, then decreases in synaptic weights ($V_{DN} > 0$) are enabled. Otherwise no changes in the synaptic weights are allowed ($V_{UP} = V_{dd}$, and $V_{DN} = 0$).

The pre-synaptic weight-update module comprises four main blocks: an input stage (see $M_{I1} - M_{I2}$ in Fig. 4a), a spike-triggered weight update circuit ($M_{L1} - M_{L4}$ of Fig. 4a), a bi-stability weight refresh circuit (see transconductance amplifier in Fig. 4a), and a current-mode DPI circuit (not shown). The bi-stability weight refresh circuit is a positive-feedback amplifier with very small "slew-rate" that compares the weight voltage V_w to a set threshold V_{thw}, and slowly drives it toward one of the two rails V_{whi} or V_{wlo}, depending whether $V_w > V_{thw}$ or $V_w < V_{thw}$ respectively. This bistable drive is continuous and its effect is superimposed to the one from the spike-triggered weight update circuit. Upon the arrival of an input address-event, two digital pulses trigger the weight update block and increase or decrease the weight, depending on the values of V_{UP} and V_{DN}: if during a pre-synaptic spike the V_{UP} signal from the post-synaptic stop-learning control module is enabled ($V_{UP} < V_{dd}$), the synapse's weight V_w undergoes an instantaneous increase. Similarly, if during a pre-synaptic spike the V_{DN} signal from the post-synaptic weight control module is high, V_w undergoes an instantaneous decrease. The amplitude of the EPSC produced by the DPI upon the arrival of the pre-synaptic spike is proportional to $V_{\Delta w}$.

In [36] we show how such circuits can be used to carry out classification tasks, and characterize the performance of these types of VLSI learning systems. Figure 5 shows an example of a training protocol that can stochastically induce Long Term Depression (LTD) of the synaptic weight. The stochasticity enables an unbiased reduction in the

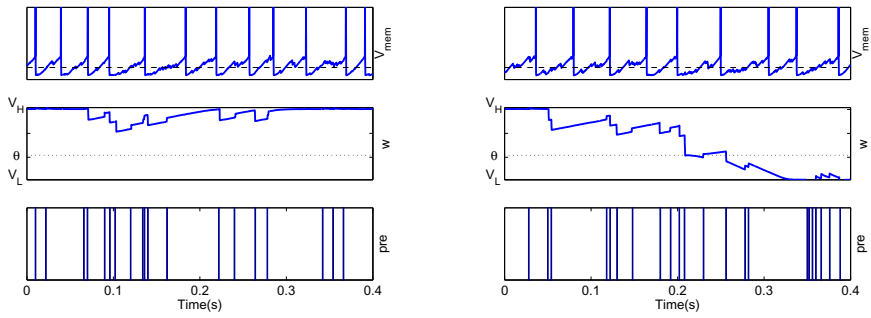

Figure 5. Stochastic transitions in synaptic states. In both figures the non-plastic synapse is stimulated with Poisson distributed spikes, that makes the post-synaptic neuron to fire at an average rate of 30Hz (V_{mem}). The pre-synaptic input (*pre*) is stimulated with Poisson distributed spike trains with a mean firing rate of 60Hz. (a) The updates in the synaptic weight (w) did not produce an LTD transition during the 400ms stimulus presentation. (b) The updates in the synaptic weight produced an LTD transition that remains consolidated. V_H and V_L shows the potentiated and depressed levels respectively. w denotes the synaptic weight, and θ the bi-stability threshold. Adapted from [36].

number of synaptic modifications and plays a key role in memory formation [21]: this mechanism allows the learning to slow down the process of synaptic modification, and increases the network's the memory [22]. In our hardware implementation, the stochasticity is obtained thanks to the Poisson nature of the spikes used as input signals. Despite the mean firing rates of input and output patterns are the same, the synaptic weight can change (see Fig. 5b) or not (see Fig. 5a) depending on the specific timing of the input and output spikes.

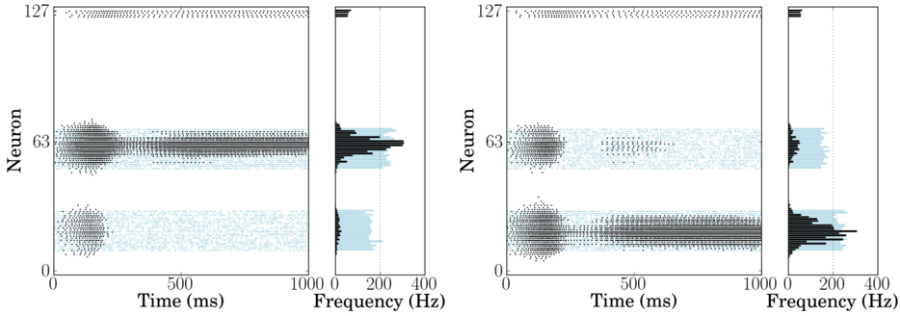

Figure 6. Selective amplification experiment. The network is stimulated in two regions, one centered around unit 20 and the other around unit 60, with Poisson spike trains of mean firing rate 180 Hz and 240 Hz. The figures show the networks response to these inputs (black) and their respective steady state firing rates on the right panels (calculated for > 500 ms). Neurons 124 to 127 are the 4 inhibitory neurons of the soft WTA network. In the right and left panel the input amplitudes are swapped. The mismatch compensated results show smooth activity profiles, and are invariant to input swapping, suggesting that the mismatch in the local weights has been partially compensated. Adapted from [38].

1.3.1. The soft WTA network

Soft WTA networks are believed to play a central role in neural processing of the neocortex [6,18]. Computational models of these types of networks emulate the cortical pattern of connectivity and explore their computational properties, studying their role in processing sensory inputs and in generating behavioral outputs. Competition and cooperation in soft WTA networks make the output of individual neurons depend on the activity of all other neurons in the network, in addition to their own inputs. As a result, these networks perform both common linear operations and complex non-linear operations. The linear operations include analog gain (linear amplification of the feed-forward input, mediated by the recurrent excitation and/or common mode input), and locus invariance [25]. The non-linear operations include non-linear selection or soft WTA behavior [1,13,24], signal restoration [13,17], and multi-stability [1,24].

The array of excitatory neurons in the `IFSLWTA` chip implements a one dimensional soft WTA network in which each neuron is connected to their first, second and third nearest neighbors via excitatory synapses. The last four neurons in the array act as global inhibitory neurons which receive input from all the excitatory neurons and return inhibition to all of them. The network behaves as a WTA when two localized stimuli are presented: it selects and amplifies one of the stimuli while suppressing its response to the other. Figure 6 shows the activity of `IFSLWTA` chip performing selective amplification. The network is stimulated in two regions, one centered around unit 20 and the other around unit 60, with Poisson spike trains of mean firing rate 180 Hz and 240 Hz. The most active neurons cooperatively amplify their activity through lateral excitation and efficiently drive the global inhibitory neuron to suppress the activity of other neurons. For strong lateral excitation, amplification is observed for the neurons receiving the input with highest mean frequency and suppression of neurons stimulated by trains with lower mean frequencies occur. A synaptic scaling method was applied to compensate for the device mismatch present in the VLSI neural network (see [38] for details). The mismatch compensated results show smooth activity profiles, and are invariant to input swapping (see left and right panel in Fig. 6), suggesting that the mismatch in the local weights has been partially compensated.

This architecture and similar variants have been used to model response properties of cortical neurons (*e.g.*, [5,23,24,25]).

2. Conclusion

In this paper we proposed a neuromorphic VLSI device able to directly implement soft WTA networks of spiking neurons via local hardwired synaptic connections, as well as networks of arbitrary topology via the AER communication circuits. The chip comprises also a reconfigurable array of plastic synapses, with on-chip learning capabilities, that enable users to explore spike-based learning tasks and implement perceptron learning schemes for classification and recognition tasks.

We described the analog circuits used to implement synapses, neurons and learning circuits, and demonstrated their functionality with experimental results measured from the chip.

The soft WTA properties, the possibility to reconfigure the network architecture offered by the AER infrastructure and the learning capabilities of the `IFSLWTA` chip were

used to implement standard classification tasks [36], state-dependent computation [37] and models of auditory processing [35].

We are now developing a software infrastructure [40] to provide higher level access to the hardware and allow researchers to use it in conjunction with simulation tools and in a wide range of application domains.

Acknowledgments

This work was supported by the EU ERC grant 257219 (neuroP), the EU ICT FP7 grant 231168 (SCANDLE), and by the Swiss National Science Foundation grant 119973 (SoundRec). We would like to thank Emre Neftci and Fabio Stefanini for carrying out some of the data measurements.

References

[1] S. Amari and M.A. Arbib. Competition and cooperation in neural nets. In J. Metzler, editor, *Systems Neuroscience*, pages 119–165. Academic Press, 1977.

[2] J. Arthur and K. Boahen. Learning in silicon: Timing is everything. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems 18*. MIT Press, Cambridge, MA, 2006.

[3] D. Badoni, M. Giulioni, V. Dante, and P. Del Giudice. An aVLSI recurrent network of spiking neurons with reconfigurable and plastic synapses. In *International Symposium on Circuits and Systems, ISCAS 2006*, pages 1227–1230. IEEE, May 2006.

[4] C. Bartolozzi and G. Indiveri. Synaptic dynamics in analog VLSI. *Neural Computation*, 19(10):2581–2603, Oct 2007.

[5] R. Ben-Yishai, R. Lev Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences of the USA*, 92(9):3844–3848, April 1995.

[6] T. Binzegger, R.J. Douglas, and K. Martin. A quantitative map of the circuit of cat primary visual cortex. *Journal of Neuroscience*, 24(39):8441–53, 2004.

[7] K.A. Boahen. Point-to-point connectivity between neuromorphic chips using address-events. *IEEE Transactions on Circuits and Systems II*, 47(5):416–34, 2000.

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

[9] R. Brette, M. Rudolph, T. Carnevale, M. Hines, D. Beeman, J.M. Bower, M. Diesmann, A. Morrison, P.H. Harris Jr. F.C. Goodman, M. Zirpe, T. Natschläger, D. Pecevski, B. Ermentrout, M. Djurfeldt, A. Lansner, O. Rochel, T. Vieville, E. Muller, A.P. Davison, S. El Boustani, and A. Destexhe. Simulation of networks of spiking neurons: A review of tools and strategies. *Journal of Computational Neuroscience*, 23(3):349–398, December 2007.

[10] E. Chicca, D. Badoni, V. Dante, M. D'Andreagiovanni, G. Salina, L. Carota, S. Fusi, and P. Del Giudice. A VLSI recurrent network of integrate–and–fire neurons connected by plastic synapses with long term memory. *IEEE Transactions on Neural Networks*, 14(5):1297–1307, September 2003.

[11] E. Chicca, A.M. Whatley, P. Lichtsteiner, V. Dante, T. Delbruck, P. Del Giudice, R.J. Douglas, and G. Indiveri. A multi-chip pulse-based neuromorphic infrastructure and its application to a model of orientation selectivity. *IEEE Transactions on Circuits and Systems I*, 5(54):981–993, 2007.

[12] T.Y. W. Choi, P.A. Merolla, J.V. Arthur, K.A. Boahen, and B.E. Shi. Neuromorphic implementation of orientation hypercolumns. *IEEE Transactions on Circuits and Systems I*, 52(6):1049–1060, 2005.

[13] P. Dayan and L.F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. MIT Press, 2001.

[14] S.R. Deiss, R.J. Douglas, and A.M. Whatley. A pulse-coded communications infrastructure for neuromorphic systems. In W. Maass and C.M. Bishop, editors, *Pulsed Neural Networks*, chapter 6, pages 157–78. MIT Press, 1998.

[15] A. Destexhe, Z.F. Mainen, and T.J. Sejnowski. *Methods in Neuronal Modelling, from ions to networks*, chapter Kinetic Models of Synaptic Transmission, pages 1–25. The MIT Press, Cambridge, Massachussets, 1998.

[16] M. Djurfeldt, M. Lundqvist, C. Johansson, M. Rehn, Ö. Ekeberg, and A. Lansner. Brain-scale simulation of the neocortex on the ibm blue gene/l supercomputer. *IBM J. Res. Dev.*, 52(1/2):31–41, 2008.

[17] R.J. Douglas, M.A. Mahowald, and K.A.C. Martin. Hybrid analog-digital architectures for neuromorphic systems. In *Proc. IEEE World Congress on Computational Intelligence*, volume 3, pages 1848–1853. IEEE, 1994.

[18] R.J. Douglas and K.A.C. Martin. Neural circuits of the neocortex. *Annual Review of Neuroscience*, 27:419–51, 2004.

[19] D.B. Fasnacht, A.M. Whatley, and G. Indiveri. A serial communication infrastructure for multi-chip address event system. In *International Symposium on Circuits and Systems, ISCAS 2008*, pages 648–651. IEEE, May 2008.

[20] R. Freidman. Reverse engineering the brain. *Biomedical Computation Review*, 5(2):10–17, 2009.

[21] S. Fusi, M. Annunziato, D. Badoni, A. Salamon, and D.J. Amit. Spike–driven synaptic plasticity: theory, simulation, VLSI implementation. *Neural Computation*, 12:2227–58, 2000.

[22] S. Fusi and W. Senn. Eluding oblivion with smart stochastic selection of synaptic updates. *Chaos, An Interdisciplinary Journal of Nonlinear Science*, 16(026112):1–11, 2006.

[23] R. Hahnloser, R.J. Douglas, M.A. Mahowald, and K. Hepp. Feedback interactions between neuronal pointers and maps for attentional processing. *Nature Neuroscience*, 2:746–752, 1999.

[24] R. Hahnloser, R. Sarpeshkar, M.A. Mahowald, R.J. Douglas, and S. Seung. Digital selection and analog amplification co-exist in an electronic circuit inspired by neocortex. *Nature*, 405(6789):947–951, 2000.

[25] D. Hansel and H. Sompolinsky. *Methods in Neuronal Modeling*, chapter Modeling Feature Selectivity in Local Cortical Circuits, pages 499–567. MIT Press, Cambridge, Massachusetts, 1998.

[26] A. Bofill i Petit and A.F. Murray. Learning temporal correlations in biologically-inspired aVLSI. In *International Symposium on Circuits and Systems, ISCAS 2003*, volume V, pages 817–820. IEEE, 2003.

[27] G. Indiveri, E. Chicca, and R.J. Douglas. A VLSI array of low-power spiking neurons and bistable synapses with spike–timing dependent plasticity. *IEEE Transactions on Neural Networks*, 17(1):211–221, Jan 2006.

[28] G. Indiveri, B. Linares-Barranco, T.J. Hamilton, A. van Schaik, R. Etienne-Cummings, T. Delbruck, S.-C. Liu, P. Dudek, P. Häfliger, S. Renaud, J. Schemmel, G. Cauwenberghs, J. Arthur, K. Hynna, F. Folowosele, S. Saighi, T. Serrano-Gotarredona, J. Wijekoon, Y. Wang, and K. Boahen. Neuromorphic silicon neuron circuits. *Frontiers in Neuroscience*, 5:1–23, 2011.

[29] E. Izhikevich and G. Edelman. Large-scale model of mammalian thalamocortical systems. *Proceedings of the National Academy of Science*, 105:3593–3598, 2008.

[30] K.R. Lakshmikumar, R.A. Hadaway, and M.A. Copeland. Characterization and modeling of mismatch in MOS transistors for precision analog design. *IEEE Journal of Solid-State Circuits*, SC-21(6):1057–1066, December 1986.

[31] J. Lazzaro, S. Ryckebusch, M.A. Mahowald, and C.A. Mead. Winner-take-all networks of $O(n)$ complexity. In D.S. Touretzky, editor, *Advances in neural information processing systems*, volume 2, pages 703–711, San Mateo - CA, 1989. Morgan Kaufmann.

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

[33] U. Mallik, M. Clapp, E. Choi, G. Cauwenberghs, and R. Etienne-Cummings. Temporal change threshold detection imager. In *ISSCC Dig. of Tech.Papers*, pages 362–363, San Francisco, 2005.

[34] Henry Markram. The blue brain project. In *ACM/IEEE conference on Supercomputing, SC 2006*, page 53, New York, NY, USA, 2006. IEEE, ACM.

[35] R. Mill, S. Sheik, G. Indiveri, and S. Denham. A model of stimulus-specific adaptation in neuromorphic analog VLSI. *IEEE Transactions on Biomedical Circuits and Systems*, 2011. (submitted).

[36] S. Mitra, S. Fusi, and G. Indiveri. Real-time classification of complex patterns using spike-based learning in neuromorphic VLSI. *IEEE Transactions on Biomedical Circuits and Systems*, 3(1):32–42, Feb. 2009.

[37] E. Neftci, E. Chicca, M. Cook, G. Indiveri, and R.J. Douglas. State-dependent sensory processing in networks of vlsi spiking neurons. In *International Symposium on Circuits and Systems, ISCAS 2010*, pages 2789–2792. IEEE, 2010.

[38] E. Neftci and G. Indiveri. A device mismatch compensation method for VLSI spiking neural networks. In *Biomedical Circuits and Systems Conference BIOCAS 2010*, pages 262–265. IEEE, 2010.

[39] H.K. Riis and P. Hafliger. Spike based learning with weak multi-level static memory. In *International Symposium on Circuits and Systems, ISCAS 2004*, pages 393–396. IEEE, 2004.

[40] S. Sheik, F. Stefanini, E. Neftci, E. Chicca, and G. Indiveri. Systematic configuration and automatic tuning of neuromorphic systems. In *International Symposium on Circuits and Systems, ISCAS 2011*, pages 873–876. IEEE, May 2011.