



## VLSI circuits implementing computational models of neocortical circuits

Jayawan H.B. Wijekoon, Piotr Dudek\*

The University of Manchester, United Kingdom

### ARTICLE INFO

#### Article history:

Received 20 September 2011

Received in revised form 27 January 2012

Accepted 31 January 2012

#### Keywords:

Mixed signal VLSI

Neuromorphic

Silicon neuron

Spiking and bursting

Silicon synapse

Neocortex

Neural circuits

Computing architecture

### ABSTRACT

This paper overviews the design and implementation of three neuromorphic integrated circuits developed for the COLAMN (“Novel Computing Architecture for Cognitive Systems based on the Laminar Microcircuitry of the Neocortex”) project. The circuits are implemented in a standard 0.35 μm CMOS technology and include spiking and bursting neuron models, and synapses with short-term (facilitating/depressing) and long-term (STDP and dopamine-modulated STDP) dynamics. They enable execution of complex nonlinear models in accelerated-time, as compared with biology, and with low power consumption. The neural dynamics are implemented using analogue circuit techniques, with digital asynchronous event-based input and output. The circuits provide configurable hardware blocks that can be used to simulate a variety of neural networks. The paper presents experimental results obtained from the fabricated devices, and discusses the advantages and disadvantages of the analogue circuit approach to computational neural modelling.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

The simulation of computational models of neural systems, at various levels of abstraction, has become an established method of scientific investigation in neuroscience (Destexhe and Crunelli, 2008). The hardware platforms used to carry out these simulations range from traditional desktop computers, providing adequate computer power for moderate-scale experiments, through small computer clusters, possibly harnessing the computer power of commodity parallel processors in modern graphics processor units (GPUs) (Fidjeland and Shanahan, 2010; Brumby et al., 2010), to large purpose-built computer clusters (Markram, 2006; Ananthanarayanan et al., 2009), providing the performance required for larger-scale network simulations. The size of the network, and the complexity of the computational models of the individual cells (e.g. simplified integrate and fire point neurons versus biophysically accurate multi-compartmental conductance-based neuron models with elaborate dendritic morphologies), determine the simulation time as well as memory/storage requirements on general-purpose devices. The performance of present day computers is often a limiting factor in the progress of computational neuroscience research (e.g. simulating a relatively small cortical network for several seconds of biological time may take several hours or days, and/or require a large and power-hungry supercomputing facility (Indiveri et al., 2011)).

For this reason, several special-purpose digital computer architectures have been proposed that sacrifice some flexibility – for example, through optimising the performance of the execution for a particular computational model (Jin et al., 2010), or providing specialised hardware implementing a specific neuron model (Schoenauer et al., 2002; Merolla et al., 2011) – in order to provide increased system performance. These systems aim to improve the processing speed, or some other system implementation parameter (e.g. lower power consumption, smaller physical size, etc.).

An even more radical approach is to construct application-specific analogue Very Large Scale of Integration (VLSI) integrated circuits (ICs), implementing directly in silicon the required computational model of a neural system, exploiting the analogies between the physics of ion transport in neural cells and the physical properties of microelectronic devices (Mead, 1990; Mahowald and Douglas, 1991; Farquhar and Hasler, 2005), or using analogue circuits to implement the equivalent model equations (Indiveri, 2003; Wijekoon and Dudek, 2008a,b; Millner et al., 2010). This idea has been known as “neuromorphic engineering” (Mead, 1989), although the term has been lately applied to many specialised digital architectures as well (Sharp et al., 2011; Merolla et al., 2011). In practice, most of the neuromorphic systems are mixed-mode implementations that combine the analogue circuit implementations of model equations and digital event-based read-out (Boahen, 2000) and spike routing sub-system (Lin et al., 2006; Vogelstein et al., 2007; Serrano-Gotarredona et al., 2009; Schemmel et al., 2010).

While the tailor-made mixed signal VLSI circuits are less flexible than digital computers, they can offer much higher processing

\* Corresponding author.

E-mail address: [p.dudek@manchester.ac.uk](mailto:p.dudek@manchester.ac.uk) (P. Dudek).

speeds and orders-of-magnitude improvements in the energy consumption per operation (Indiveri et al., 2011). This technique enables real-time (or faster) implementation of neural system simulations. The neuromorphic approach also emphasises the practical technological applications of the neuroscientific knowledge. It is envisaged that the understanding of the biological brains will lead to the construction of future brain-like computer systems, and that the overall architecture and principles of operation, as well as the detailed circuitry of these future computing devices, could be closely modelled on those of biological brains. The focus of neuromorphic circuit design is often on low-power, low-cost and compact system implementations, and it is hoped that the techniques developed through the implementations of neural systems in hardware could be directly applicable to the design of intelligent machines, such as autonomous robots, or future “cognitive supercomputers”.

Since the pioneering work on neuromorphic circuits in the late 1980s (Mead, 1990), a number of CMOS implementations of ‘silicon neurons’ (Mahowald and Douglas, 1991; Linares-Barranco et al., 1991; Schultz and Jabri, 1995; Patel and DeWeerth, 1997; Simoni and DeWeerth, 1999; Indiveri, 2003; Nakada et al., 2005; Rangan et al., 2010; van Schaik et al., 2010; Indiveri et al., 2011) and ‘silicon synapses’ (Hafliger et al., 1997; Bofill-i Petit and Murray, 2004; Indiveri et al., 2006; Koickal et al., 2007; Tanaka et al., 2007) have been presented. Recently, a number of systems have been proposed (Arthur and Boahen, 2004; Vogelstein et al., 2007; Merolla et al., 2007, 2011; Giulioni et al., 2008; Schemmel et al., 2010; Sharp et al., 2011) that attempt to facilitate the implementation of large-scale hardware neural networks, through the integration of thousands of silicon neurons and synapses in a single microelectronic IC.

In this paper we present neuromorphic circuits developed in the course of the COLAMN project (<http://colamn.plymouth.ac.uk/column-project/>). Section 2 outlines the design principles, and presents models and integrated circuit implementations. Section 3 presents experimental results obtained from the fabricated devices. Section 4 discusses the results, and the advantages and the limitations of the presented circuits, and the neuromorphic approach in general, to large-scale neural modelling.

## 2. Methods

### 2.1. Design principles

The established practice, when constructing large scale micro-electronic devices, is to use, as much as possible, simple, regularly connected, repeatable circuit elements (e.g. basic logic gates or memory cells in a traditional logic design). This makes it easier to design, fabricate, test, and use the device. However, when looking at biological brains it is evident that specialised structures are built with complex, heterogeneous neural elements and with elaborate inter-neuron connectivity. Although the main constituents of these specialised structures are neurons and synapses, distinct individual neurons or synapses show diverse non-linear responses to the same inputs.

A trade-off exists between the biological plausibility of the implemented model and the complexity (and hence the size and power requirements) of the physical circuit implementation. Detailed analogue VLSI circuit implementations of complex non-linear computational models are possible, but consume exceedingly large silicon area and thus allow only a few neurons to be implemented on a single silicon IC (Saighi et al., 2011). Similarly, detailed multi-compartmental models of dendritic morphologies have been demonstrated, but allow only a very limited number of neurons to be implemented on a single device (Wang and Liu, 2011).

Due to the small size of the network that can be feasibly implemented in hardware, these approaches are generally not providing any advantage over more conventional implementations based on a simulation on a general-purpose digital processor.

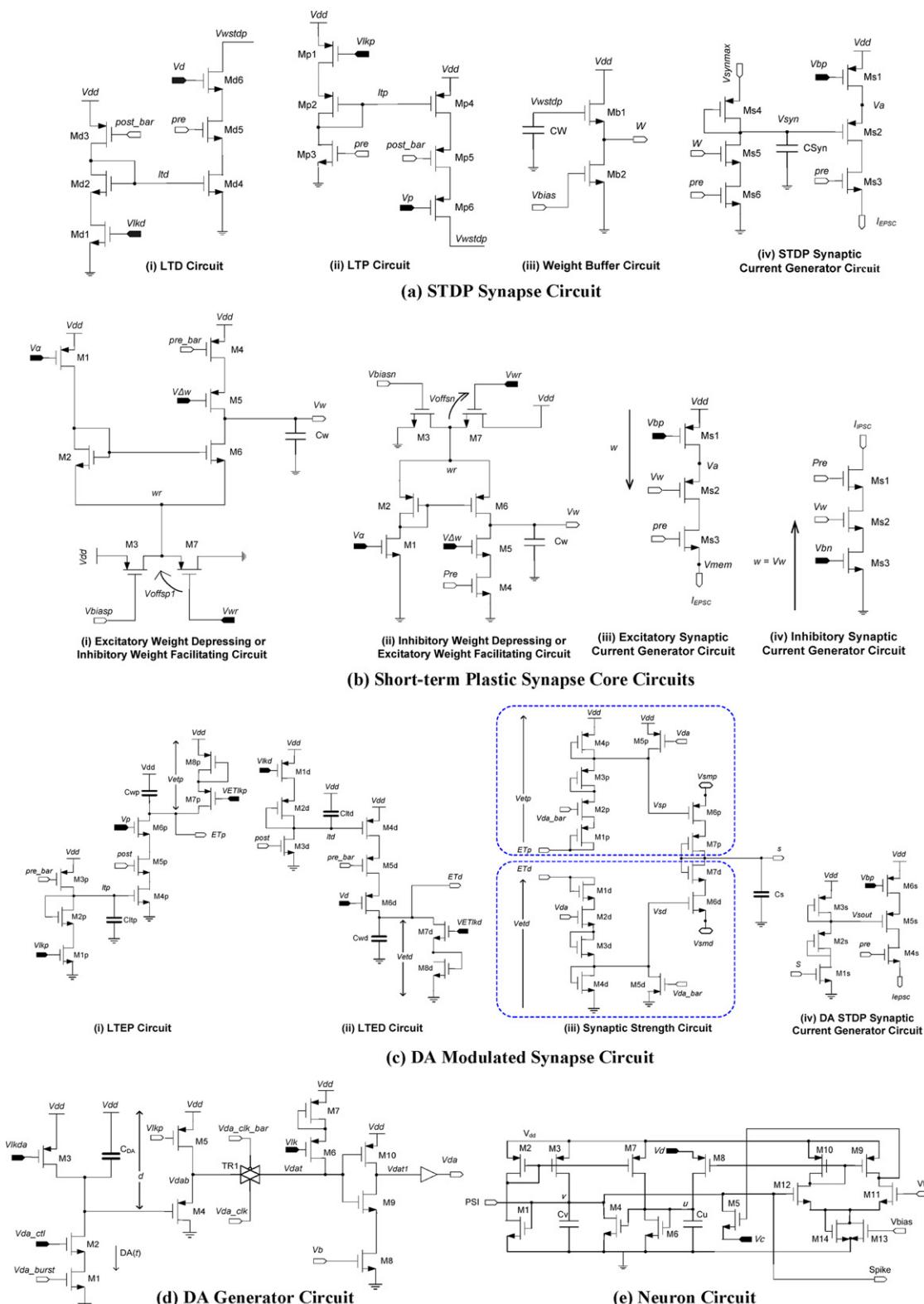
On the other hand, using simpler models such as integrate and fire point neurons, fixed weight synapses ignoring synaptic dynamics, and simplifying assumptions about the connectivity, allows larger networks – up to several thousands of neurons – to be implemented on a single IC (Vogelstein et al., 2007; Lin et al., 2006), promising a large speed-up over a digital simulation approach. However, the increased network size is obtained at the cost of severely reduced biological plausibility of the model. Such systems can be used to study computational properties of more abstract spiking neural networks, but the applicability of such hardware as a tool for neuroscience research is limited.

Our goal in this work was to develop circuits that can be used to simulate a variety of cortical networks, with the ability to model the heterogeneity of basic cortical neuron cell classes and responses, and basic synaptic dynamics and plasticity mechanisms. To provide a performance advantage over digital implementations, compact circuits allowing thousands of basic cells to be integrated on a single silicon IC were required. At the same time, a degree of generality was desirable, so that the model parameters could be adjusted and tuned, providing a flexible and versatile tool for simulating cortical circuits.

Based on the above motivation, we have decided to use some established generic computational neural models (Abbott et al., 1997; Bi and Poo, 1998; Tsodyks, 2002; Izhikevich, 2003) as a guide to arrive at qualitatively similar circuit-based models that mimic the neural dynamics with the most compact physical implementations. The models we have developed are therefore less complex than the more detailed biophysical models used in many studies, but at the same time more biologically plausible than those provided by many other spiking neural network VLSI devices.

The majority of neuromorphic devices described in the literature operate in biological real-time (i.e. one second of neural activity on a chip corresponds to one second of the neuronal activity of the biological systems). This is partly motivated by the desire to provide compatibility with biological systems (LeMasson et al., 2002; Vogelstein et al., 2007), and real-time sensory data (Liu et al., 2010; Lichtsteiner and Delbrück, 2005; Zaghloul and Boahen, 2004, 2006), and partly so that the requirements on the I/O and communication sub-system are reduced. To leverage the inherently higher speeds of modern microelectronic devices, we design our circuits to operate in accelerated-time (approximately three orders of magnitude faster than biology, i.e. one second of biological time corresponds to approximately 1 ms on a chip). A similar approach has been used by Schemmel et al. (2010). This approach provides considerable computational speed up, especially important in the case where the experiment requires extensive parameter sweeps, or multiple runs needed for statistical analysis. At the same time, increasing the operational speed can improve the total energy efficiency of the system (provided the static currents dominate the power consumption, as is usually the case).

In our prototype implementations, we have used a bulk-silicon CMOS (Complementary Metal-Oxide-Semiconductor) technology, which follows the industry-standard silicon device fabrication process. The 0.35 μm feature size is several generations behind the state-of-the-art in the microelectronic industry, however, it has the advantage of a proven and stable fabrication process, low-cost prototyping, and suitability for analogue designs. While the scaling of the proposed designs to smaller device dimensions is not straightforward, the presented circuits provide some indication of what would be feasible in other fabrication technologies.



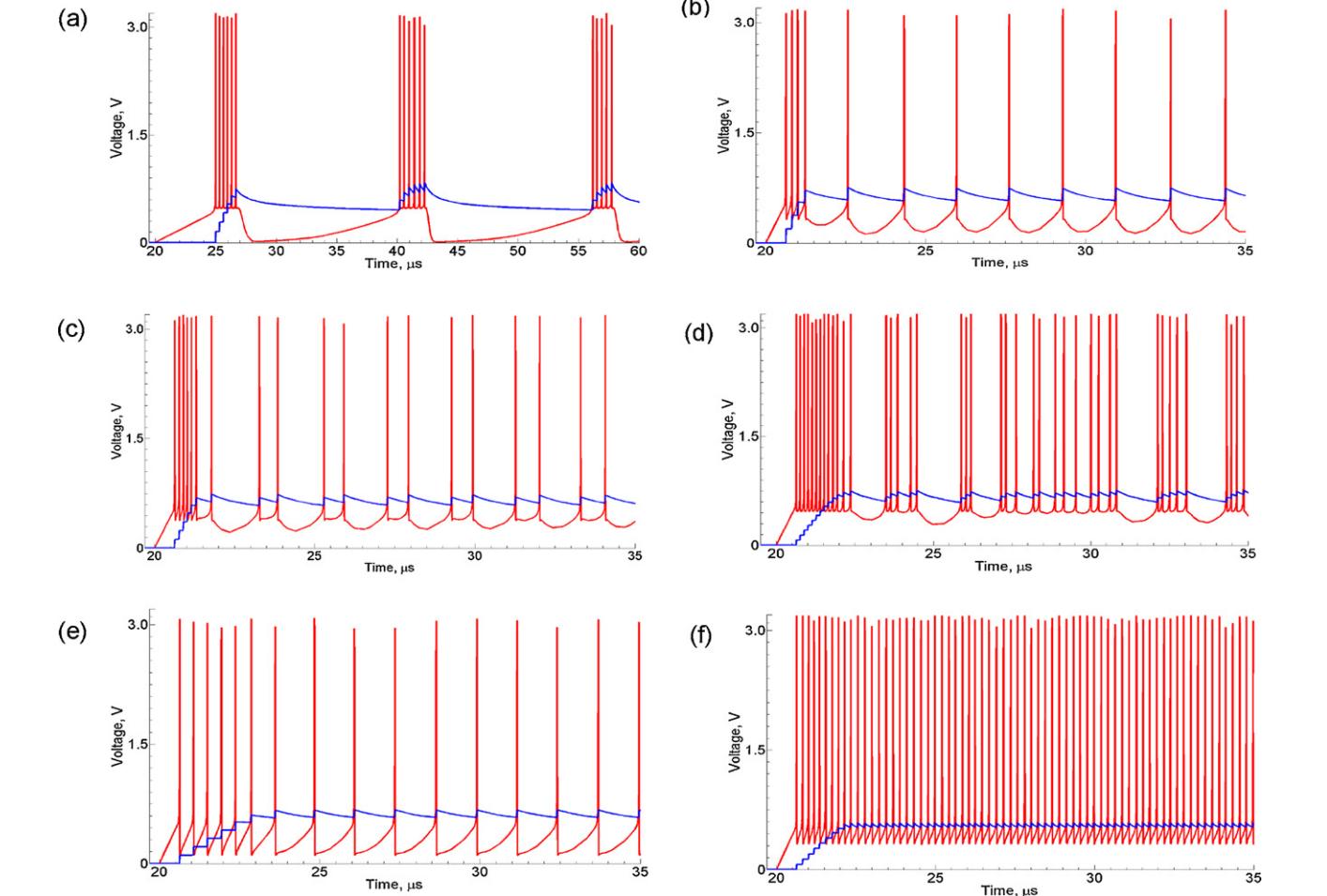
**Fig. 1.** Transistor-level schematics of the synapse and neuron circuits. (a) Sub-circuits of the STDP synapse: long-term depression (LTD) and long-term potentiation (LTP) circuits generate the STDP weight, weight buffer circuit buffers the weight of the synapse and the STDP synaptic current generator circuit generates the weight dependent excitatory post-synaptic current. (b) Sub-circuits of the short term plastic synapse: (i) and (ii) implement depressing or facilitating synaptic dynamics; (iii) and (iv) generate weight dependent excitatory and inhibitory synaptic currents. Two of these four sub-circuits are used to create a synapse: (i) and (iii), (i) and (iv), (ii) and (iii), (ii) and (iv), respectively, form excitatory depressing, inhibitory facilitating, excitatory facilitating, and an inhibitory depressing synapses (some synapse types use additional level shifters). (c) Sub-circuits of the DA modulated synapse: long-term eligibility potentiation (LTEP) and long-term eligibility depression (LTED) circuits generate the eligibility trace in a similar way to the STDP circuit, but with weight leakage; the synaptic strength circuit generates the synaptic strength using the DA signal and the potentiation and depression parts of the eligibility trace; the DA STDP current generator circuit generates the synaptic strength dependent excitatory post-synaptic current. (d) The DA generator circuit that produces global DA signal ( $V_{da}$ ) to many synapses, where the pulse width of the DA signal represents the extracellular DA level. (e) Cortical neuron circuit from Wijekoon and Dudek (2008a,b). Note: suffix “.bar” indicates signal inversion, e.g., inverted signal  $pre$  is labelled as  $pre\_bar$ .

## 2.2. Circuits and models

The basic circuits presented in this paper include a generic cortical neuron circuit and six different synaptic circuits. Two of the six synaptic circuits mimic long term synaptic dynamics, which emulate Spike-Time Dependent Plasticity (STDP), and dopamine modulated STDP(DA-STDP). The remaining four synaptic circuits include short-term synaptic dynamics: excitatory depressing, inhibitory facilitating, inhibitory depressing, and excitatory facilitating synapses. These circuits are designed to implement qualitative behaviours of phenomenological neural models while optimising circuit area and power consumption. Although the resulting circuit models are simple (from a hardware perspective) and functionally close to their counterpart computational neuroscience models, they are elaborate in their mathematical formulation. Approximated mathematical equations of the models are presented in this section to provide an indication of the level of abstraction represented by the circuits. The circuit diagrams are presented for the illustration purposes, however, detailed circuit descriptions and transistor-level analysis are beyond the scope of this paper (for detailed circuit descriptions see (Wijekoon, 2011)).

### 2.2.1. Neuron circuit

A cortical network consists of many types of neurons (Connors and Gutnick, 1990; Nowak et al., 2003; Toledo-Rodriguez et al., 2003; Markram et al., 2004; Ascoli et al., 2008). Different neuron



**Fig. 2.** Simulation of the neuron circuit, showing a plot of membrane voltage,  $V$ , (red trace) and slow variable,  $U$ , (blue trace) against time, obtained in response to a  $0.1 \mu A$  step current injection into the soma (start of the current injection is at  $20 \mu s$ ), for different neuron tuning parameters,  $V_c$  and  $V_d$ , resulting in different characteristic spike patterns: (a) chattering, (b)–(d) intrinsically bursting, (e) regular spiking, and (f) fast spiking. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

types exhibit distinct responses to the same set of input stimuli (Nowak et al., 2003; Ascoli et al., 2008), and this heterogeneity is an important feature of cortical networks. Therefore, we incorporate the diverse neuron responses, similar to the biological neuron responses, in the VLSI implementation. Our silicon neuron circuit (Wijekoon and Dudek, 2008a,b), shown in Fig. 1e, is inspired by the computational model proposed by Izhikevich (2003) and motivated by the desire to achieve a single compact generic circuit that can easily be tuneable to a desired cortical neuron type. The neuron model has two state variables: membrane potential ( $V$ ) and membrane recovery ( $U$ ), whose evolution can be approximated by the following equations:

$$\frac{dV}{dt} = \begin{cases} k_1 V^2 - k_2 V - k_3 U^2 + k_4 U + k_5 + k_6 I & \text{when } V \geq U - k_7 \\ -k_8 V^2 - k_9 V + k_{10} U V + k_{11} + k_{12} I & \text{otherwise} \end{cases} \quad (1)$$

$$\frac{dU}{dt} = k_{13} V^2 - k_{14} V - k_{15} U^2 + k_{16} U + k_{17} \quad (2)$$

with a reset after spike;

$$\text{If } V > V_{th} \text{ then } \begin{cases} V \leftarrow c \\ U \leftarrow U + d \end{cases} \quad (3)$$

where  $I$  is the post-synaptic input current,  $k_1$  to  $k_{17}$  are constants which depend on the neuron circuit parameters and the process parameters of the CMOS technology;  $V_{th}$  sets the spike detection threshold;  $c$  and  $d$  are tuning parameters of the neuron that can

be set by applying two external voltages  $V_c$  and  $V_d$  (for a detailed description of the model see Wijekoon and Dudek, 2008a,b).

This silicon neuron circuit produces biologically plausible action potentials (see Fig. 2) and is capable of mimicking spiking and bursting firing patterns observed in cortical neurons. The circuit can be easily configured to produce regular spiking, fast spiking, chattering, intrinsically bursting, and other complex activity patterns. The circuit is compact and low-power – it consumes about 8 pJ per spike, i.e. orders of magnitude less than what is possible using digital technology to emulate the spiking/bursting behaviour (Wijekoon and Dudek, 2008a,b).

### 2.2.2. Long-term synaptic dynamic circuits

The STDP plasticity rule has been proposed as the neuronal mechanism for learning and memory (Tsodyks, 2002; Morrison et al., 2008), and the dopamine (DA) modulated STDP plasticity rule is believed to be the mechanism for reinforcement learning in a cortical network (Izhikevich, 2007). According to this rule, the synaptic modification due to STDP is modulated by the level of extracellular DA concentration. That is, the extracellular DA level regulates the long-term potentiation (LTP) and long-term depression (LTD) of the synaptic weight (Fellous and Suri, 2003; Izhikevich, 2007).

The implemented STDP circuit is shown in Fig. 1a. The LTD and LTP circuits in Fig. 1a compute the STDP( ) and the capacitor  $C_W$  stores the weight of the synapse. This accelerated-time STDP circuit is based on the biological-time STDP circuit proposed by Indiveri et al. (2006). It produces synaptic weight changes based on the timing of pre-synaptic and post-synaptic spikes, following the STDP curve (typical shape of the STDP curve, using piecewise linear approximation, is shown in Fig. 3a). The amount of weight change due to LTP and LTD and the duration of the effect of the LTP and LTD can be configured independently using externally controllable tuning

voltages  $V_p$ ,  $V_d$ ,  $V_{lkp}$ , and  $V_{lkd}$  respectively.

$\text{STDP}(\Delta t_{pp})$

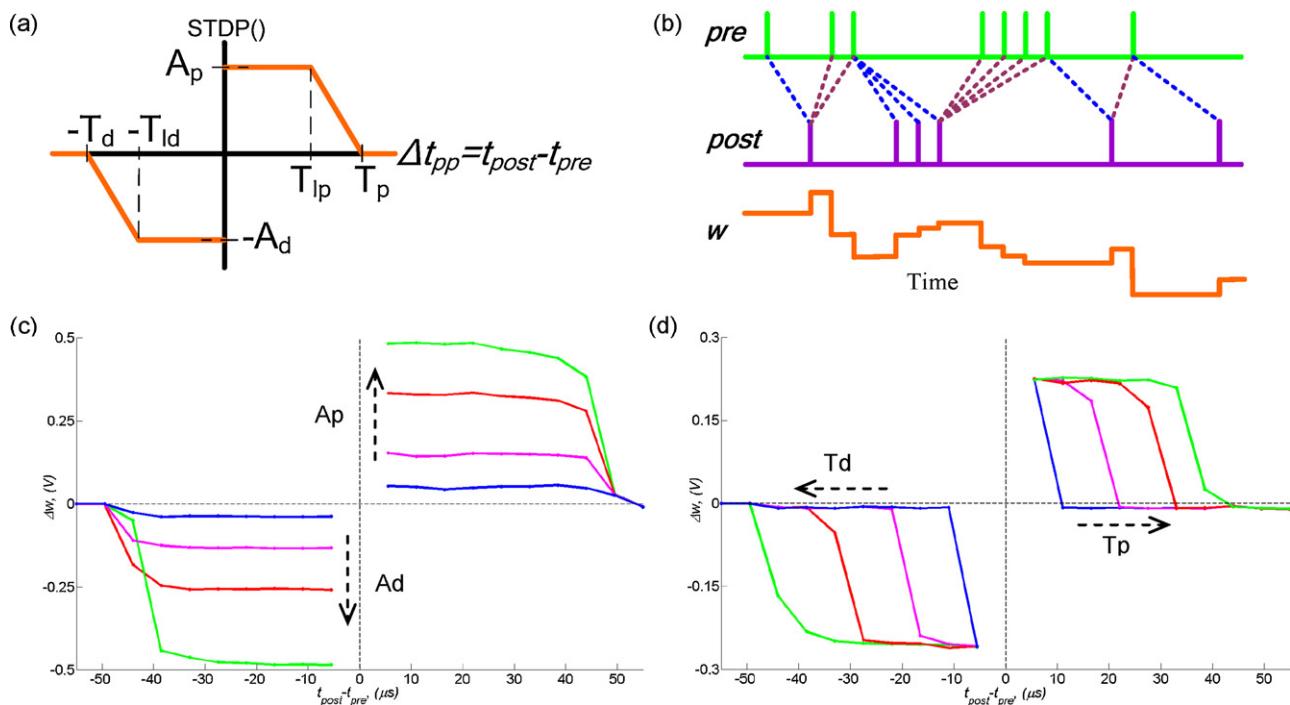
$$= \begin{cases} \left( \frac{A_d \times (\Delta t_{pp} + T_{ld})}{(T_d - T_{ld})} \right) & \text{when } -T_d < \Delta t_{pp} \leq -T_{ld} \\ -A_d & \text{when } -T_{ld} < \Delta t_{pp} < 0 \\ A_p & \text{when } 0 < \Delta t_{pp} \leq T_{lp} \\ \left( \frac{A_p \times (\Delta t_{pp} - T_{lp})}{(T_p - T_{lp})} \right) & \text{when } T_{lp} < \Delta t_{pp} < T_p \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where the values of  $A_p$ ,  $A_d$ ,  $T_p$ ,  $T_d$ ,  $T_{lp}$  and  $T_{ld}$  can be set to desired constants using the tuning voltages  $V_p$ ,  $V_d$ ,  $V_{lkp}$  and  $V_{lkd}$ ;  $T_p$  and  $T_d$  are time windows of long-term potentiation and long-term depression, these can be adjusted in the circuit to any value between 1  $\mu\text{s}$  and 100  $\mu\text{s}$ ;  $T_{lp}$  and  $T_{ld}$  determine the linear region of the STDP curve and depend on the circuit parameters, the process parameters, and operational region of the transistors;  $t_{post}$  and  $t_{pre}$  are the times of the post-synaptic spike event and pre-synaptic spike event respectively,  $\Delta t_{pp} = t_{post} - t_{pre}$ .

The weight of the STDP synapse,  $w$ , evolves as in the equation given below:

$$\frac{dw}{dt} = \frac{-(w - w_{rest})}{\tau_l(w)} + \text{STDP}(\Delta t_{pp})\delta(t - t_{pre/post}) \quad (5)$$

where  $\delta(t)$  is the Dirac-delta function that provides a step-increase or -decrease of  $w$  for immediate pairings of pre- and post-synaptic neuron firing times;  $w_{rest}$  is a resting weight dependant on circuit parameters. The circuit holds the synaptic weight using a leaky capacitor,  $\tau_l(w)$  determines the speed of leakage (its average value corresponds to a time constant of around 40 ms) which limits the



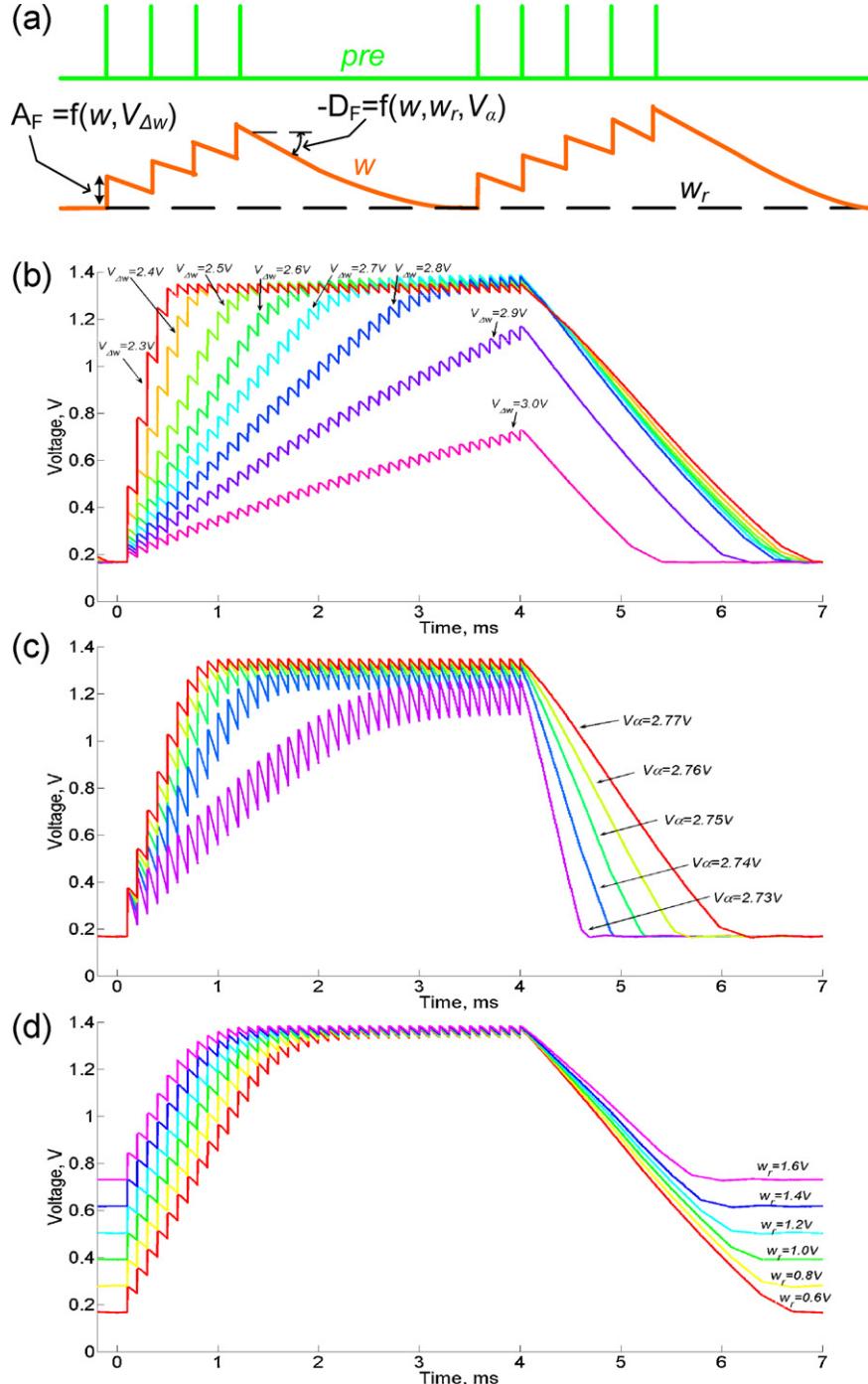
**Fig. 3.** (a) Piece-wise linear approximation of the STDP curve implemented on chip; parameters are explained in Eq. (4). (b) Illustration of the implemented spike pairing scheme (Morrison et al., 2007, 2008): each post-synaptic spike is paired with the last preceding pre-synaptic spike (blue dotted line) and each pre-synaptic spike is paired with the last preceding post-synaptic spike (brown dotted line); the blue dotted line indicates the pairings that contribute towards the potentiation of a synapse, and brown dotted lines indicate the pairings that contribute towards the depression. (c) Simulated STDP curves for different values of  $A_p$  and  $A_d$ , obtained by varying tuning parameters  $V_p$  and  $V_d$ ; (d) simulated STDP curves for different values of long-term potentiation ( $T_p$ ) and depression ( $T_d$ ) time window, obtained by varying tuning parameters  $V_{lkp}$  and  $V_{lkd}$ . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

operational time of an experiment. All times are scaled by  $10^{-3}$  with respect to the biological time.

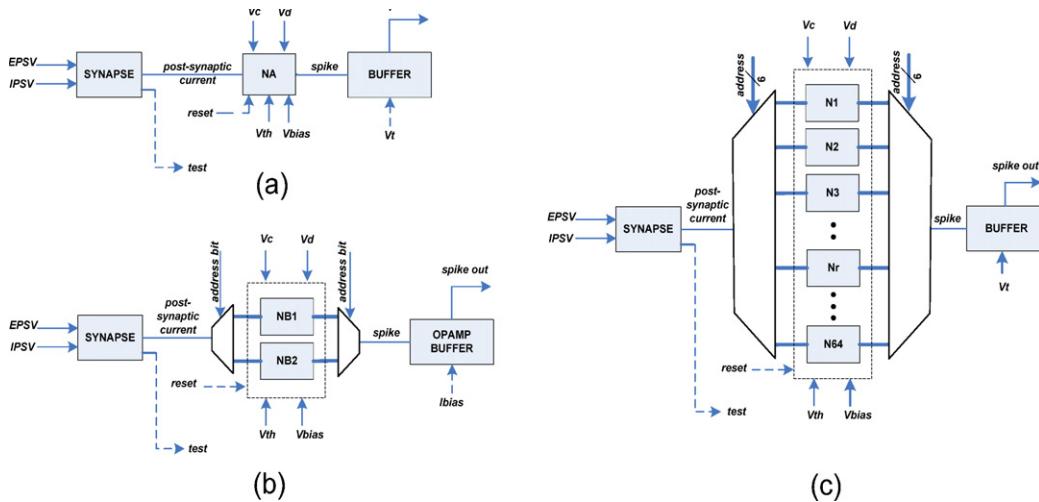
The dopamine modulated synaptic dynamic circuit (Wijekoon and Dudek, 2011) shown in Fig. 1c and d is implemented based on the computational model proposed by Izhikevich (2007). This circuit has been designed by extending the STDP synapse circuit to facilitate the synaptic weight change based on eligibility traces and the dopamine concentration signal. Similar to STDP synapse, the changes in an eligibility trace due to LTP or LTD, and the time windows of the LTP or LTD can be configured independently using

the externally controllable voltages  $V_p$ ,  $V_d$ ,  $V_{kp}$ , and  $V_{kd}$ , respectively. The DA concentration signal can be generated globally using an external voltage bias or using a burst of spikes (Wijekoon and Dudek, 2011). According to the approximated silicon DA-STDP synapse model, the strength of the STDP synapse,  $s$ , evolves as in the equations given below:

$$\frac{dc}{dt} = \frac{-c}{\tau_{ET}(c)} + \text{STDP}(\Delta t_{pp})\delta(t - t_{\text{pre/post}}) \quad (6)$$



**Fig. 4.** (a) Illustration of the weight  $w$  (shown in orange) of the inhibitory facilitating synapse (IFS), when an input spike train  $pre$  (shown in green) is provided; parameters are explained in Eq. (9). (b)–(d) Simulated response of the IFS circuit to a 10 kHz burst of pre-synaptic input spikes, lasting for 4 ms and followed by a silent period. The plots show changes in the synaptic weight,  $w$ , for different values of: (b) degree of facilitation  $A_F$  (controlled using the tuning voltage  $V_{\Delta w}$ ), (c) speed of recovery,  $D_F$  (mainly controlled using the tuning voltage  $V_\alpha$ ), and (d) resting weight of the synapse,  $w_r$ .



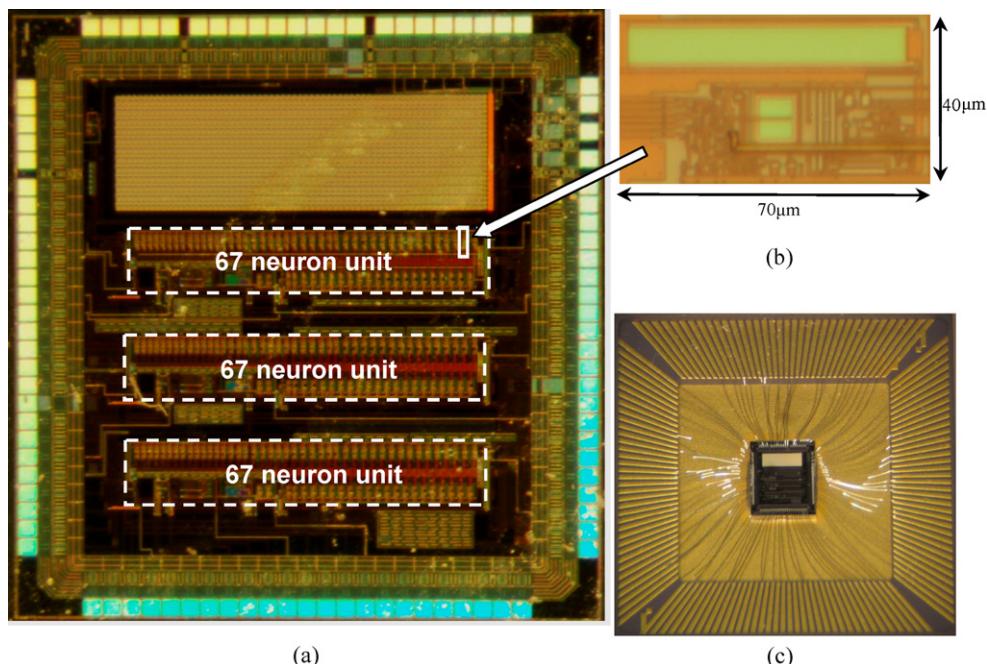
**Fig. 5.** Block diagram of a 67 neuron cell unit of the cortical neuron chip: (a) single neuron test circuit: a neuron circuit (NA) is connected to a synapse, and the buffered membrane voltage of a neuron (*spike out*) can be observed externally; (b) two neurons test circuit: post-synaptic current is generated for one of the neurons (NB1 and NB2), using a global synapse; outputs from the neurons are multiplexed using an *address bit*, and one neuron's output can be observed at a given time; tuning parameters are shared by the neurons. (c) 64 neurons with different circuit parameters (various capacitor and transistor sizes); one synapse and two neuron tuning parameters are shared by all the neurons; outputs from the neurons are multiplexed using a 6 bit *address*, and the output of one neuron can be observed at a given time. Note: the voltages EPSV and IPSV are used to regulate the excitatory and inhibitory post-synaptic current to the neurons respectively;  $V_c$  and  $V_d$  are tuning parameters of the neuron; other inputs include bias voltages, reset, and address bits use to multiplex the observable outputs; *test* is an internal signal of the synapse which is used to measure the synaptic currents.

$$\frac{ds}{dt} = f(cd) \quad (7)$$

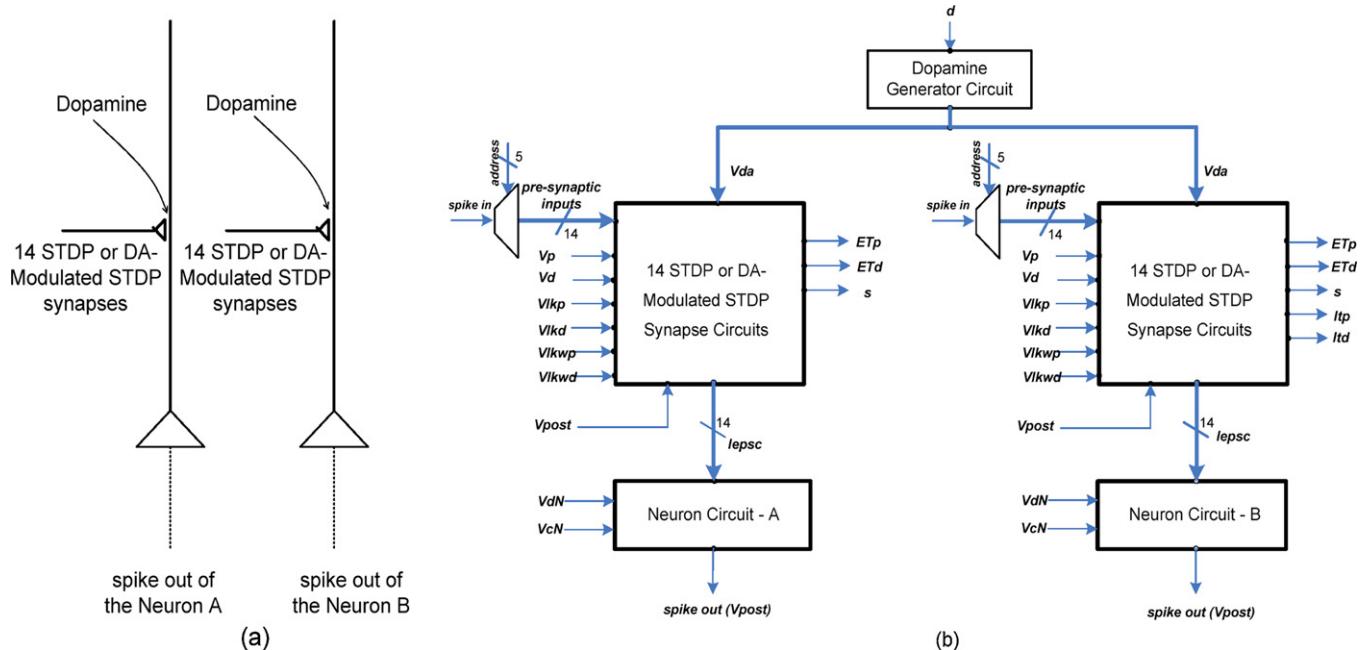
$$\frac{dd}{dt} = \frac{-d}{\tau_{DA}(d)} + DA(t) \quad (8)$$

In the above equations  $c$  is the synaptic eligibility trace;  $\delta(t)$  is the Dirac-delta function that provides a step-increase or -decrease of  $c$  for every pre- and post-synaptic neuron firing time (see Fig. 3c); the function  $STDP()$  describes the spike-timing-dependent change of the eligibility (as in Eq. (4)); a sub-circuit in Fig. 1d, implements  $f(cd)$  to generate the approximate product of  $c$  and the

$d$  that represents the extracellular concentration of DA.  $DA(t)$  is the amount of the DA released due to the activities of the dopaminergic neuron (the spike output of the dopaminergic neuron is connected to  $V_{da\_burst}$  node of the DA generator circuit shown in Fig. 1d; hence the activities of the dopaminergic neuron increases the dopamine level; the strength of this increase can be adjusted using externally controllable voltage  $V_{da\_ctl}$ ). Leakage parameters  $\tau_{ET}=f(c,V_{ETlkp},V_{ETlkd})$  and  $\tau_{DA}=f(d,V_{lkda})$  can be tuned using externally controllable voltages  $V_{ETlkp}$ ,  $V_{ETlkd}$  and  $V_{lkda}$ . Typical values of  $\tau_{ET}$  and  $\tau_{DA}$  are 1 ms and 0.2 ms respectively. In the case of DA modulated synapse, the use of capacitors to hold the memory traces does not limit the



**Fig. 6.** Photograph of the Cortical Neuron IC: (a) the full IC layout showing 202 neurons, these include 3 blocks of 67 neuron cell units with different circuit parameters and an isolated neuron, (b) a microphotograph of a single neuron, including the output buffer and control circuit, and (c) picture of the IC wire-bonded inside a package.



**Fig. 7.** The architecture of the DA modulated synapse chip: (a) synapse and neuron composition, (b) block diagram of the circuit showing tuning parameters and observable outputs;  $d$  is the externally controllable extracellular DA level;  $V_p$ ,  $V_d$ ,  $V_{lkp}$ ,  $V_{lkd}$ ,  $V_{lkwp}$  and  $V_{lkwd}$  are tuning parameters of a synapse;  $ET_p$ ,  $ET_d$ ,  $s$ ,  $ltp$  and  $ltc$  are internal signals of a synapse;  $V_{cN}$  and  $V_{dN}$  are the tuning parameters of a neuron.

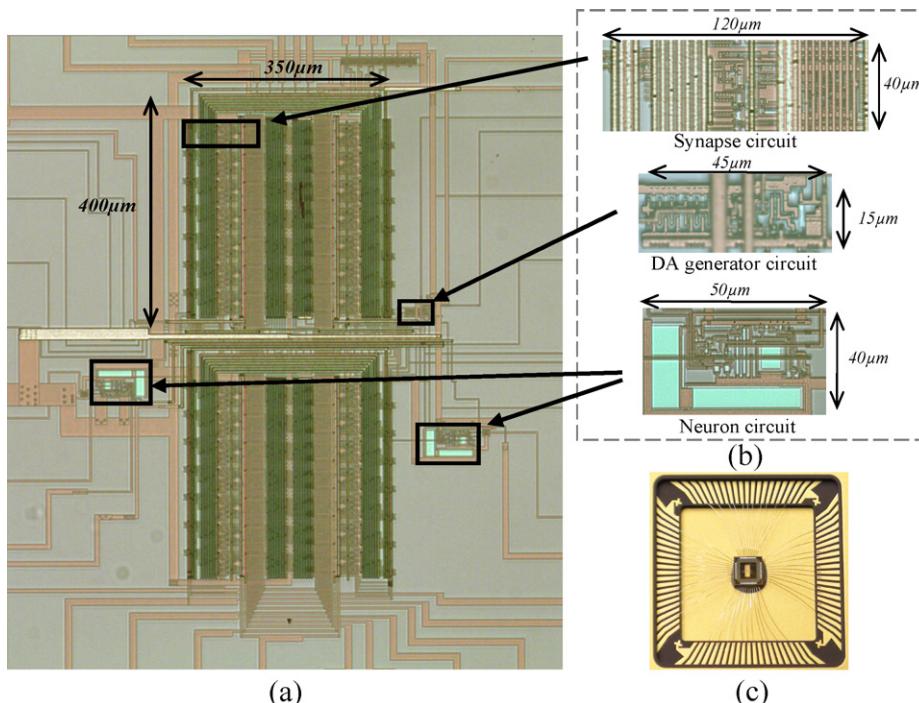
operational time of an experiment as the leakage of the eligibility trace is a crucial feature in the DA modulated synapse (Izhikevich, 2007).

### 2.2.3. Short-term synaptic dynamics

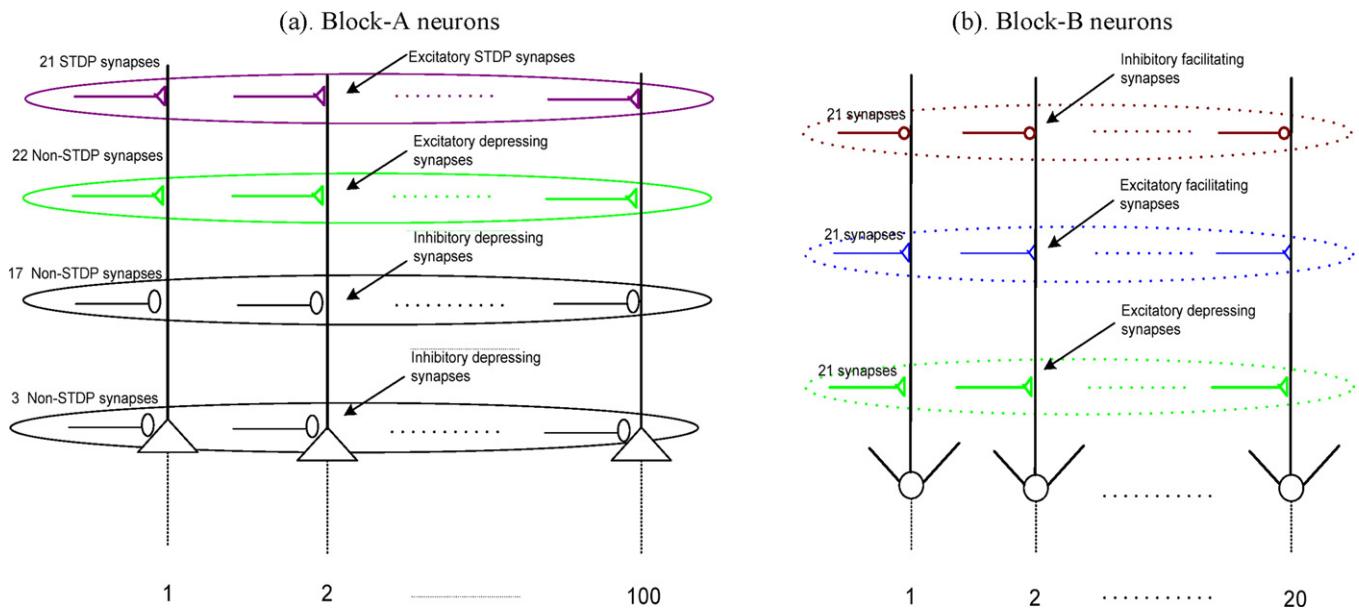
The synaptic facilitating and depressing circuits are designed based on the computational neural model of short-term dynamics proposed by Abbott et al. (1997). The core circuits used to generate inhibitory and excitatory synapses with facilitation or depression

are shown in Fig. 1b. Facilitation dynamics of the model is illustrated in Fig. 4a. According to the approximated silicon synapse circuit model, facilitating and depressing weight of the synapse,  $w$  evolves as in the equations given below.

$$\text{for facilitating synapse } \frac{dw}{dt} = -D_F(w)H(w - w_r) + A_F(w)\delta(t - t_{pre}) \quad (9)$$



**Fig. 8.** Photograph of the DA modulated synapse chip: (a) layout of the chip without the pads, (b) enlarged pictures of a synapse circuit that includes pre-synaptic spike generator and test output circuits, the DA generator circuit, and a neuron circuit, (c) picture of the chip wire-bonded inside a package.



**Fig. 9.** Synapse and neuron composition of the Cortical Neural Layer (CNL) chip: this shows 120 neurons and their input synapse types.

$$\text{for depressing synapse } \frac{dw}{dt} = D_D(w)H(w_r - w) - A_D(w)\delta(t - t_{pre}) \quad (10)$$

where  $t_{pre}$  is the arrival time of the pre-synaptic spike event;  $H(w)$  is unit step function,  $\delta(t)$  is Dirac-delta function; the functions  $A_F = f(w, V_{\Delta w})$  and  $D_F = f(w, V_{\Delta w})$  decide the weight dependent strength of facilitation and depression respectively; these strengths can be tuned using the externally controllable voltage  $V_{\Delta w}$  of the synapse circuit;  $w_r$  is resting voltage of the synaptic weight and its value can be set using the externally controllable voltages  $V_{wr}$ . The decay functions  $D_F = f(w, w_r, V_\alpha)$  and  $D_D = f(w, w_r, V_\alpha)$  in the equations above determine the recovery speed of the facilitated and the depressed weight,  $w$ , back to their resting weight  $w_r$ ; the recovery time of facilitation and depression can be tuned using the externally controllable voltage  $V_\alpha$  of the synapse circuit.  $D_F$  and  $D_D$  act as a linear or exponential decays depending on the parameter values,  $w_r$  and  $V_\alpha$ .

In summary, the resting weight of the synapse, strength of facilitation or depression, and the recovery time can be configured independently using externally control tuning voltages (as shown in Fig. 4). The facilitating or depressing dynamics circuits can connect to an inhibitory or excitatory circuit to generate inhibitory or excitatory post synaptic current respectively. This post synaptic current can be scaled using an externally adjustable bias voltage.

### 2.3. Integrated circuit implementations

Different combinations of the neural circuits presented above have been fabricated in three integrated circuits, in a standard  $0.35\text{ }\mu\text{m}$  CMOS technology. The prototype ICs serve as a proof-of-concept and illustrate how analogue circuitry can be used to implement complex functionality with minimum power consumption. The fabricated ICs include the “Cortical Neuron” IC, the “DA Modulated Synapses” IC, and the “Cortical Neural Layer” (CNL) IC. The former two ICs are designed to test the function of the basic neural elements, and the CNL IC is designed to facilitate cortical network emulations.

#### 2.3.1. Cortical Neuron IC

The first prototype integrated circuit, the Cortical Neuron IC, contains 202 neuron cells, with varied circuit parameters

(transistor sizes and capacitances). It was fabricated to test the functionality of the neuron circuit, and experimentally obtain the best combinations of circuit parameters, so that the neuron circuit is capable of reproducing the widest range of neural firing patterns, using two tuning parameters.

The 202 neurons are grouped into three of 67 neuron cells and one isolated neuron cell. The block diagram of the 67 neuron cell unit is shown in Fig. 5. The cells are individually accessible and do not form any network. In addition to the neuron cells, the IC contains multiplexers, buffers and simple synaptic circuitry to generate excitatory and inhibitory postsynaptic currents. The different neurons are provided with three different types of output buffers to feed the membrane potential signal to the output pads. The circuit also contains a multiplexing unit that selects one neuron at a time. Some cells are designed with an additional external membrane potential resetting circuit. The size of the IC is approximately  $3\text{ mm} \times 2\text{ mm}$ , and it has 84 pins. A photograph of the IC layout with an enlarged individual neuron, as well as a photograph of the IC wire bonded to a package is shown in Fig. 6.

#### 2.3.2. DA modulated synapses IC

This prototype test integrated circuit contains twenty eight STDP/DA-STDP synapses with a global DA generator circuit, and two cortical neuron circuits. The block diagram of the IC is shown in Fig. 7. The STDP/DA-STDP synapse can be configured to work as a STDP synapse or as a DA-modulated STDP synapse. Fig. 8 shows the layouts of the cells and the photographs of the fabricated IC. Although the actual circuit area is approximately  $0.5\text{ mm} \times 0.8\text{ mm}$ , a large IC area ( $1.8\text{ mm} \times 1.8\text{ mm}$ ) is required to accommodate 44 pins needed by the circuits. This prototype IC is fabricated to test the functionality of the STDP synapse circuit and DA-Modulated STDP synapse circuit along with the neuron circuit. Accordingly, some of the internal states of all the synapses can be observed externally and, at any given time, the internal states of two synapses can be observed along with the inverted spike outputs of both neurons, so that the STDP or DA-STDP traces can be obtained and the functionality of the synapse can be tested and calibrated.

#### 2.3.3. Cortical Neural Layer (CNL) IC

The CNL IC implements neural circuits with a similar cell composition to the neocortex. The CNL IC models 120 cortical neurons

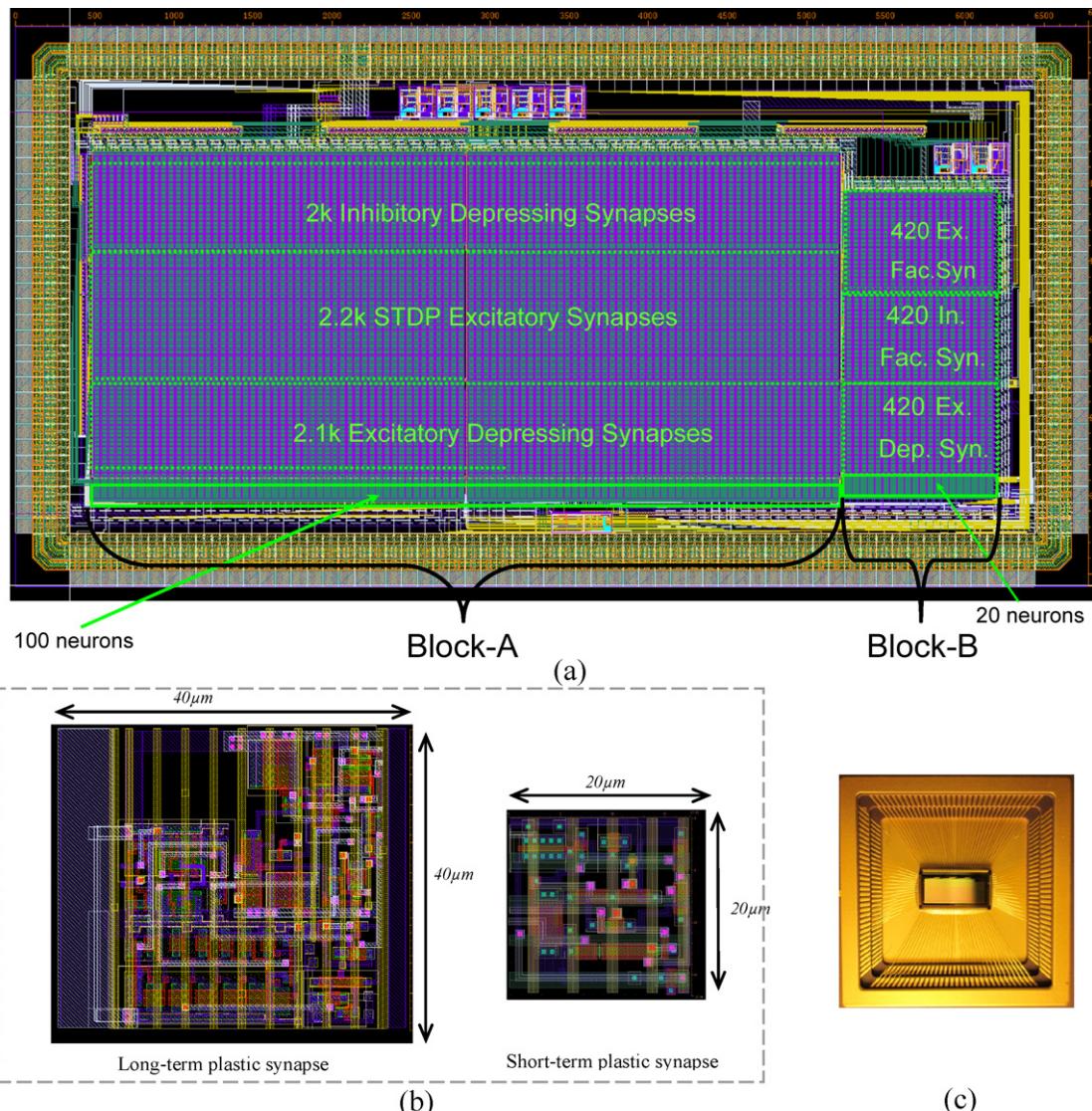
and 7560 synapses. The IC comprises generic neuron and synapse circuits with configurable neuronal connections. The neurons of the IC can be configured to different known types of neurons. The IC is also equipped with different short-term and long-term dynamics synapse circuits that include inhibitory, excitatory, facilitating and depressing and STDP dynamics.

The neural elements occupy two separate blocks: Block-A and Block-B. Block-A consists of 100 neurons and 6300 synapses. Each of the neurons in this block receives inputs from 43 excitatory depressing synapses (21 STDP and 22 Non-STDP excitatory depressing synapses) and 20 inhibitory depressing non-STDP synapses (3 somatic and 17 distal inhibitory synapses; these two types are distinguished by the range of available synaptic strengths). Block-B consists of 20 neurons and 1260 synapses. Each of the Block-B neurons receives input from 63 non-STDP synapses. The 63 synapses comprise an equal number of excitatory facilitating, inhibitory facilitating, and excitatory depressing synapses. A diagram of synapse and neuron on the IC is shown in Fig. 9.

Most of the neuron outputs are available in parallel from the IC pins. Some neuron outputs are accessible serially, and a few of the neuron outputs are internally wired into the synaptic input array. The pre-synaptic spike inputs can be provided externally by

addressing the synapses using the address bus of the IC. The internal states of the selected synapses can be calibrated and/or observed externally. The parameters of the neurons are set (in groups) using 60 external bias voltages. The size of the IC is 24 mm<sup>2</sup> (6.78 mm by 3.58 mm), and it has 180 pins. Fig. 10 shows the layout of the IC showing the physical location of the synapses and neurons, and a photograph of the fabricated IC.

The CNL IC can be configured to have a heterogeneous neuron and synapse type combination, such that it could closely represent the neuron and synapse type composition of the cortical layer. By combining many ICs together, a six-layered VLSI cortical network could be built to resemble a small-scale network of the neocortex. Although multi-compartmental dendrites are not modelled, the neuron model can express a richness of behaviour that can represent the variations in dendritic morphology and ion channel distribution, in addition to the neural dynamics on the cell body. Further, some effect on the membrane integration due to the distance of the synapse from the soma can be modelled using the dendritic delays and tuning the adjustable strength of the post-synaptic current. Axonal and dendritic delays can be programmed, if required, in the spike-routing network (not discussed here).



**Fig. 10.** The Cortical Neural Layer chip: (a) layout of the CNL chip that contains 7560 synapses, 120 neurons, and auxiliary circuits; (b) layouts of the short- and long-term plastic synapse circuits; (c) picture of the chip wire-bonded inside a package. Syn., synapse; Ex., excitatory; In., inhibitory; Fac., facilitating; Dep., depressing.

### 3. Results

#### 3.1. Neuron

Selected experimental waveforms obtained from the Cortical Neuron IC are shown in Fig. 11. These include fast spiking, low-threshold spiking, regular spiking, intrinsically bursting, and chattering spike patterns when a  $0.1\text{ }\mu\text{A}$  step current is injected into the soma, for different neuron tuning parameters,  $V_c$  and  $V_d$ . For more experimental results and their corresponding tuning parameters see Wijekoon and Dudek (2006, 2008a,b). As shown in Table 1, all the main types of firing patterns can be obtained from the neuron circuit for any process variations that are inherent in the fabrication technology.

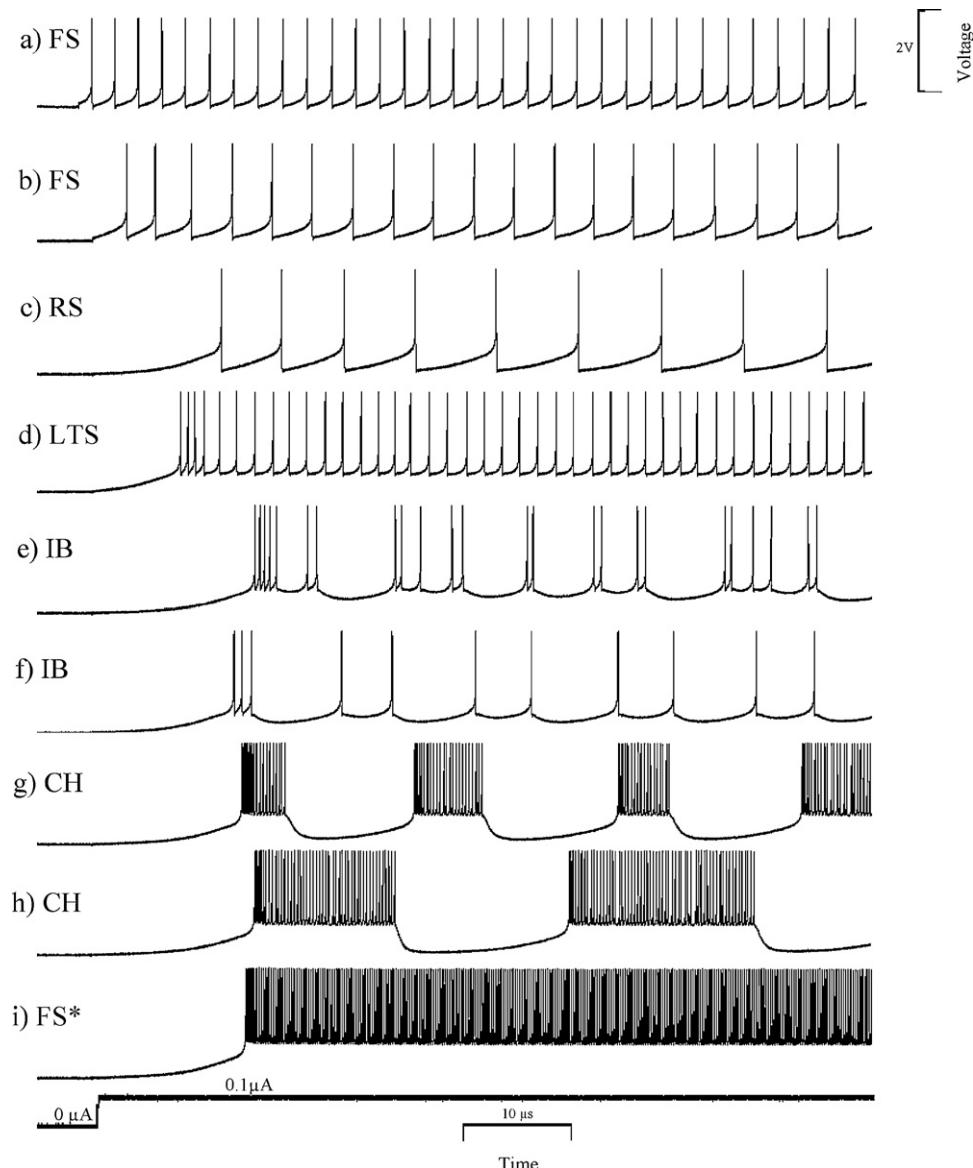
#### 3.2. Long-term synapse

Long-term potentiating and depressing effects of the STDP (LTP – when post-synaptic spike follows the pre-synaptic spike, and LTD

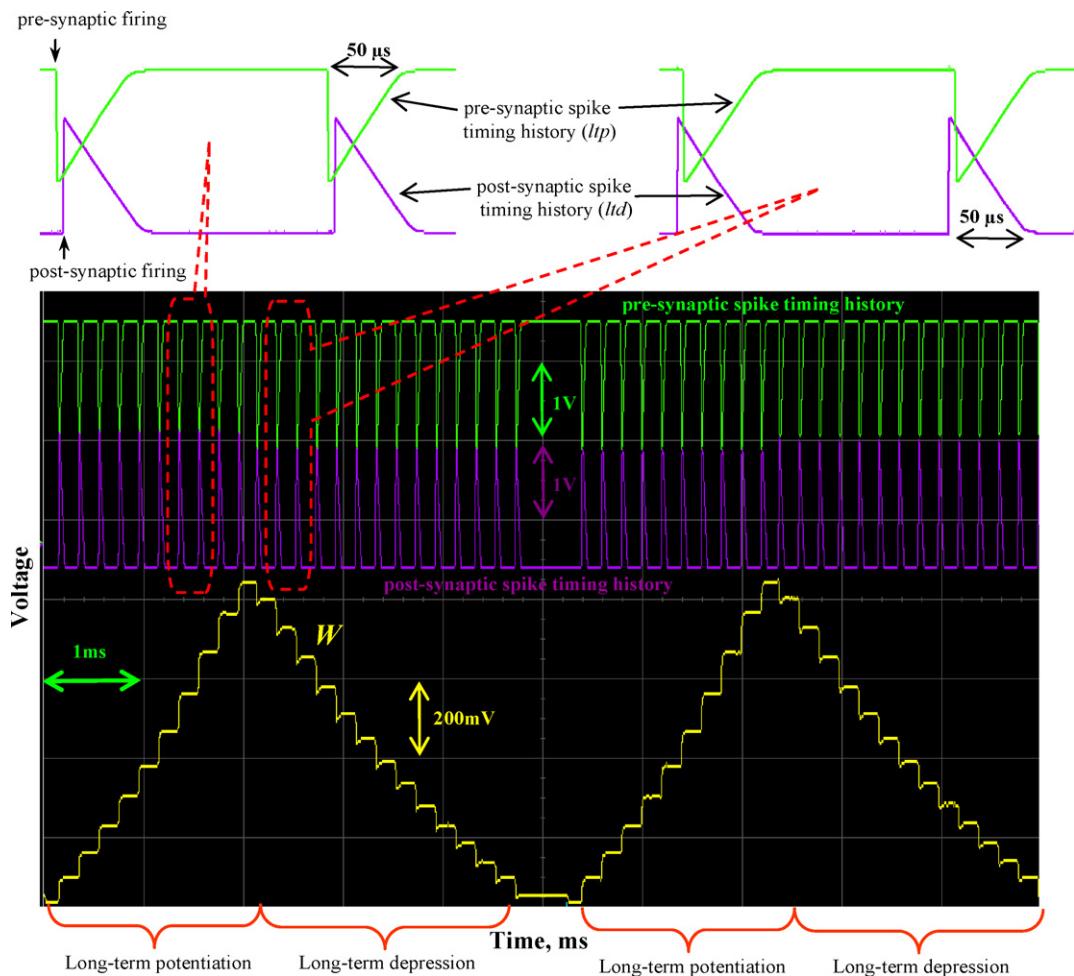
**Table 1**

Values of the neuron tuning parameter,  $V_c$  required to obtain different firing patterns across all process corners of the MOSFETS (corner conditions: WP, worst power; WO, worst one; TM, typical mean; WZ, worst zero; WS, worst speed), when  $V_d = 1.9\text{ V}$ . All the main types of firing patterns can be obtained from the neuron circuit, across process variations. Here both RS1 and RS2 are regular spiking firing patterns.

$V_c$ (V)	Corner analysis (worst case analysis)				
	WP	WO	TM	WZ	WS
0.00	RS2	RS2	RS1	RS1	RS1
0.10	RS2	RS2	RS1	RS1	RS1
0.20	IB	IB	RS2	RS1	RS1
0.30	CH	CH	RS2	RS1	RS1
0.40	CH	CH	IB	RS2	RS2
0.50	FS	FS	CH	RS2	RS2
0.60	FS*	FS*	CH	IB	IB
0.65	FS*	FS*	FS	CH	CH
0.70			FS*	CH	CH
0.75			FS*	FS	FS
0.80				FS*	FS*
0.90				FS*	FS*



**Fig. 11.** Measured waveforms of the membrane voltage of the fabricated VLSI neuron circuit when stimulated with a  $0.1\text{ }\mu\text{A}$  step current (shown at the bottom of the figure) injected into the soma, for different neuron tuning parameters,  $V_c$  and  $V_d$ , resulting in different characteristic spike patterns: (a) and (b) FS, fast spiking, (c) RS, regular spiking, (d) LTS, low threshold spiking, (e) and (f) IB, intrinsically bursting, (g) and (h) CH, chattering, (i) FS\*, very fast spiking.



**Fig. 12.** Experimental waveforms of the fabricated VLSI long-term plastic synapse circuit, demonstrating the long-term potentiating and long-term depressing changes of the synapse weight, for a particular set of circuit tuning parameters. Neuronal spikes trigger pre- and post-synaptic spike timing history traces which linearly decay in time (shown in green and magenta, respectively). These traces correspond to voltages at nodes *ltp* and *ltd* in the circuit in Fig. 1a, and are used to determine time between synaptic events. Yellow trace shows the weight of the synapse: it is increased when a post-synaptic spike (indicated by a step increase in the post-synaptic spike timing history) closely follows a pre-synaptic spike (indicated by a step decrease in the pre-synaptic spike timing history); it is decreased when a pre-synaptic spike follows a post-synaptic spike. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

– when the pre-synaptic spike follows the post-synaptic spike) are experimentally demonstrated in Fig. 12. This includes two iterations of the same stimuli to demonstrate the degree of repeatability of the output responses. The highest value of potentiation per spike pair can be adjusted to a value between 0 and 30% of the maximum weight of the synapse, using the tuning parameter  $V_p$ . The highest value of depression can be adjusted to a value between 0 and 50% of the maximum weight using the parameter  $V_d$ .

Due to the device mismatch (Pelgrom et al., 1989) in the fabricated circuits, the nominally identical synapses, using the same tuning parameters, will produce different responses. This variability is illustrated in Fig. 13, where Monte Carlo simulations are used to analyse the impact of device mismatch on the timing history of pre- and post-synaptic firing events and on the synaptic weight of the STDP synapse circuit.

The STDP curve obtained from the implemented silicon synapse is shown in Fig. 14.

### 3.3. Short-term synapse

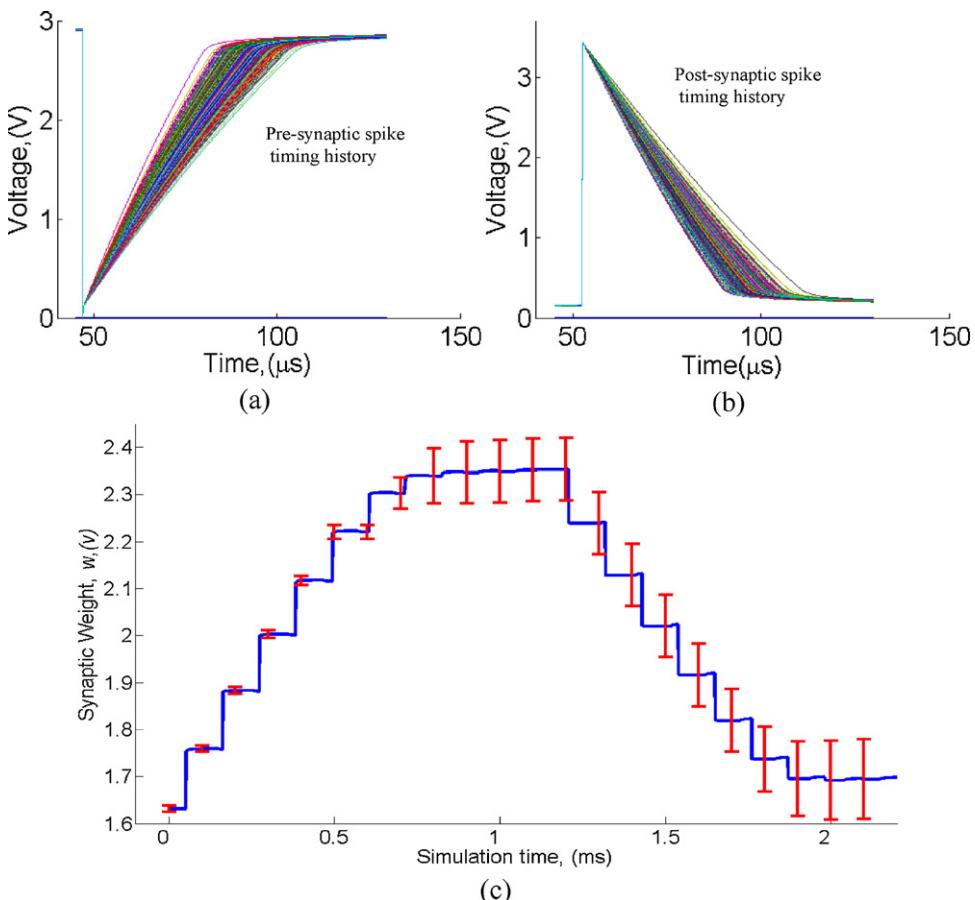
The experimentally obtained waveforms, demonstrating the operation of the excitatory depressing, inhibitory depressing, excitatory facilitating, and inhibitory facilitating synapses are shown in Fig. 15. The results are included for two iterations of the same

stimuli to demonstrate the degree of repeatability of the output responses. The amount of facilitation or depression per spike can be adjusted to a value between 0 and 100% of the maximum possible weight of the synapse, using the tuning parameters  $V_{\Delta w}$  and  $V_\alpha$ . However, resting weight,  $w_r$ , of the synapse can limit the maximum possible facilitation or depression per spike.

The effect of device mismatch is analysed using simulations shown in Fig. 16. All synapses show basic dynamics of depression or facilitation and recovery to the resting weight, albeit with a variability of the model parameters.

## 4. Discussion

The presented circuits provide a set of phenomenological models implemented in hardware, which can be used to construct neural circuits and networks. The implemented models include non-linear spiking and bursting dynamics of the cortical neurons, STDP and dopamine modulated STDP synaptic dynamics, and facilitating and depressing short-term dynamics of inhibitory and excitatory synapses. The experimental results obtained from the prototype ICs verify the operation of these models in silicon. When designing these circuits, we aimed to achieve the ability to model neuronal networks with a level of biological plausibility greater than many existing VLSI devices proposed to date, while

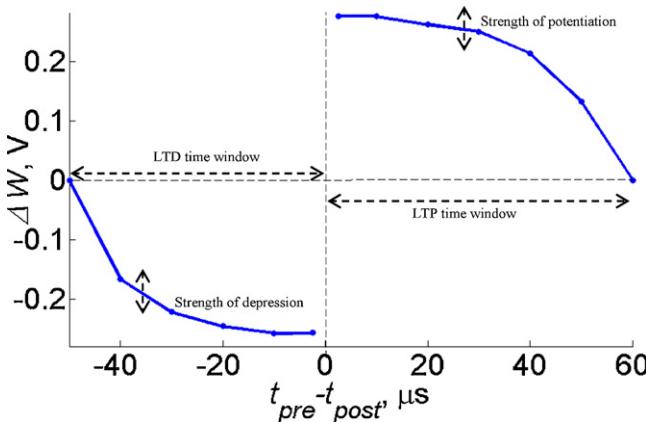


**Fig. 13.** Mismatch analysis of the STDP synapse circuit. Variation of spike timing history trace for (a) the pre-synaptic firing event; (b) the post-synaptic firing event. (c) Synaptic weight trace showing the mean (blue plot) and the standard deviation (red bars) of a response to a stimulus of 20 pairs of pre- and post-synaptic spikes as shown in Fig. 12 (for the initial 10 pairs pre-synaptic spike follows the post-synaptic spike, and then for the remaining 10 pairs the post-synaptic spike follows the pre-synaptic spike). The simulation results are obtained for 1000 Monte Carlo iterations, using Hspice simulator, with the standard mismatch models provided for the fabrication technology. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

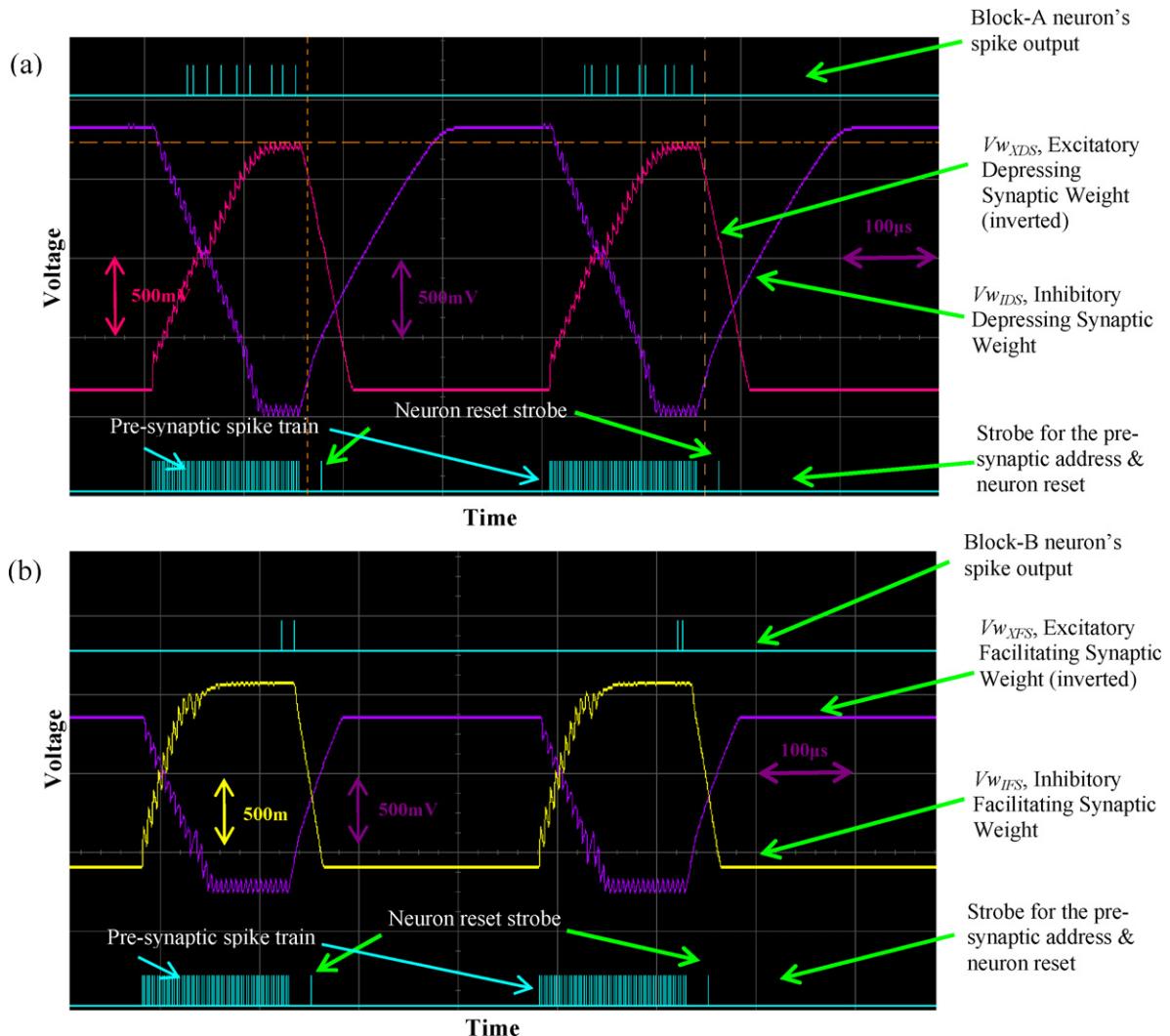
at the same time providing the level of integration and scalability that would allow the construction of relatively large systems. The devices are designed so that a system containing many ICs can be assembled; an off-chip spike routing hardware (Fasnacht and Indiveri, 2011) is envisaged for greatest flexibility of inter-neural connectivity.

The proposed hardware neural models operate in accelerated-time, which is beneficial from the point of view of computational efficiency. If biological-time operation is required, different specific circuit solutions might be needed, e.g. see Wijekoon and Dudek (2009) where we outline the design of a biological-time neuron cell.

The proposed VLSI circuits have been designed and fabricated in a standard 0.35 μm CMOS technology. The CNL IC comprises 120 individual neurons and 7650 distinct hardware synapses of various types. Several CNL ICs can be combined together to build small-scale networks. The feasibility of scaling this approach to a larger-scale system can be estimated. The CNL IC occupies 24 mm<sup>2</sup> of silicon. The IC area can be easily increased (at increased cost), providing more neurons and synapses on a single IC. Further device density improvements can be obtained by using a more modern silicon technology. We estimate, that when migrated to a 90 nm CMOS technology node, and with a chip area of 120 mm<sup>2</sup> (a size that can be easily fabricated with a reasonable yield) a future device similar to the CNL IC could contain about 5000 neurons and 300,000 synapses (when deriving these estimated figures we have taken into account the fact that the migration of 0.35 μm technology circuits into 90 nm technology would require redesigning the analogue neural circuits, considering higher leakage currents and mismatch problems that are inherent in deep sub-micron circuit implementations, and thus assumed a conservative density scaling factor). A hundred of such ICs could be integrated on a single printed circuit board, together with digital spike routing circuits, and ten boards assembled together in a system the size of a desktop-computer, and with



**Fig. 14.** The shape of the STDP curve obtained from the fabricated synapse circuit for a typical set of circuit tuning parameters. The STDP curve is adjustable using the four tuning parameters;  $V_{lkp}$  and  $V_{lkd}$  can be used to set the active time window of LTP and LTD;  $V_p$  and  $V_d$  can be used to set the strength of potentiation and depression, respectively.



**Fig. 15.** Experimental results from the fabricated short-term plastic synapse circuits, demonstrating the short-term facilitation and depression of the synaptic weights of the synapses, for a burst of pre-synaptic spikes followed by a silent period: (a) excitatory depressing and inhibitory depressing synapses, (b) excitatory facilitating and inhibitory facilitating synapses. Traces show (from the bottom to top): pre-synaptic spike train, weights of the two short-term plastic synapses and the post-synaptic response. For the excitatory synapses, inverted waveforms of the synaptic weight are shown. The results include two iterations of the same stimuli to demonstrate the degree of repeatability of the output responses. The strength of depression or facilitation per pre-synaptic spike, resting weight, the recovery time to reach the resting weight, and the post-synaptic current of the synapses can be controlled independently using the tuning parameters.

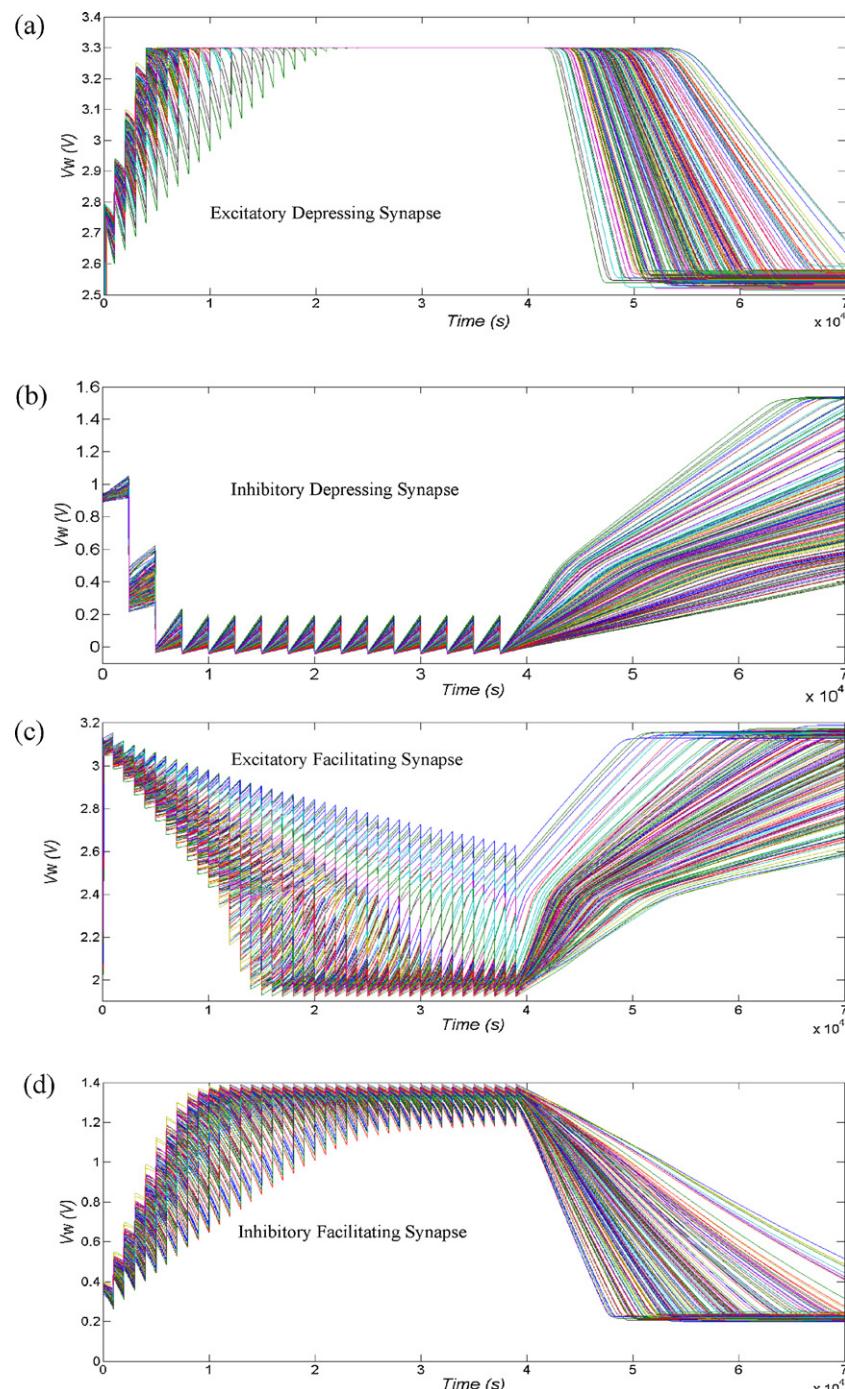
the power dissipation in the kilowatts range. Such system would enable the construction of networks comprising up to 5 million neurons and 300 million synapses.

Specialised packet routing schemes (Plana et al., 2007) and wafer-scale integration networks (Schemmel et al., 2008, 2010) have been proposed to support the high bandwidth requirements of inter-neuron communication in neuromorphic architectures. The three-dimensional stacking of silicon wafers (Topol et al., 2010) provides a promising solution to increasing the level of integration and high inter-die communication bandwidth. It should be noted that it is the communication infrastructure, not the implementation of the neuron/synapse model, that will most likely dominate the power and performance requirements in a large-scale system, and limit the complexity of the network models that can be implemented by the VLSI devices.

From the IC layouts (Fig. 10) it is evident that the synapse circuitry dominates the device area. Several hardware realisations of a “memristor” based synapse have been recently postulated (Strukov et al., 2008; Linares-Barranco and Serrano-Gotarredona, 2009), promising an extremely compact IC implementation. It has to be recognized, however, that these devices will be best suited to

a simple ‘weighted summation’ model of synaptic processing, providing little flexibility regarding the plasticity model, and inability to model complex synaptic dynamics. A number of solutions (e.g. Lin et al., 2006; Vogelstein et al., 2007) with apparently large number of synapses have been described in the literature, where the synapses are time-multiplexed or ‘virtualised’ (i.e. one hardware synapse is used to emulate multiple synapses). To some extent, these mechanisms can be also applied to the presented CNL IC, increasing the number of available synapses per neuron. It has to be emphasised, however, that the implementation of more elaborate learning rules, such as DA-STDP, and short-term synaptic dynamics, requires dedicated hardware resources for each synapse (as multiple state values have to be stored, and evolve, in each synapse). The circuits presented in this paper provide compact implementations of these models, indicating the limits of what can be achieved in current microelectronic technologies.

The circuits developed in this work implement models at a particular level of abstraction, and while we tried to capture the essential properties of neural dynamics that might be relevant to the computational capabilities of the neocortex, the choice of the models is somewhat arbitrary. While we attempted to



**Fig. 16.** Mismatch analysis of the short-term dynamic synapse circuit. Plots show the variation in the synaptic weight traces: (a) excitatory depressing (inverted weight shown), (b) inhibitory depressing, (c) excitatory facilitating (inverted weight shown), (d) inhibitory facilitating. Results were obtained from the Monte Carlo simulations in Hspice, using the standard mismatch models provided for the fabrication technology.

provide a level of generality and flexibility in our circuits (through adjustable parameters, a multitude of synapse types with programmable properties, generic off-chip connectivity, etc.), they remain, as all hardware-based solutions (especially based on analogue computation), committed to run a particular model. They can provide acceleration of the simulation times and significant power efficiency advantages, when compared with the execution of similar models in software on conventional computer hardware, but at the same time they offer a limited scope of simulation experiments, restricted to the parameter ranges and models that were foreseen at the IC design stage. Herein lies possibly the greatest disadvantage of

hardware-based models – they do necessitate a large development time (months to years) in case a different computational model should be required. For instance, it would be relatively easy to modify a software-based simulation to extend the STDP model to take into account spike triplets, or use an exponential non-linearity in the neuron cell instead of quadratic. In the case of a VLSI hardware model, the development of a new integrated circuit is needed. Even if a large library of “standard cells” could be developed in hardware, covering the most likely variants of the models that might be required, there will always be cases that a new model, or a new set of parameters not covered by the existing hardware, will be desired

by the experimenter. The addition of such a model to existing hardware, even if possible, requires many man-months of design effort. This problem also exists in the case of specialised digital hardware architectures, because even though a more established design flow for turning algorithms into hardware exists for such designs, it is likely that the entire system is built on certain assumptions (for example, data types, or bandwidth requirements) that become invalid as the implemented neural models are modified.

Another limitation of the analogue VLSI approach to neural modelling, clearly demonstrated by our results, is the problem of accuracy and repeatability of the model. Both random electronic noise (introducing temporal variability), and the mismatch of transistor parameters (introducing spatial variability) are unavoidable in practical hardware implementations. As can be seen from Table 1 and Figs. 13 and 16, through careful design it can be ensured that the device variability does not push the model outside its operating ranges, allowing the system to execute a valid neural model. Nevertheless, the spread of results caused by device variability is relatively high (and will only get worse as technology scales to smaller dimensions), inducing uncertainty into the parameters of the executed model. This uncertainty and noise are a great disadvantage of neuromorphic hardware approach to neural modelling. It could be said that the noise and variability of the analogue circuits make them more analogous to biology, as compared with error-free and numerically accurate software simulation. However, the absolute controllability of the parameters, the ability to execute a model with an arbitrary accuracy, and the ability to introduce noise and variability in a controllable manner if desired, are the significant advantages of a digital modelling solution. It should be noted that the accuracy of analogue circuit models can be made higher than in the case of circuits presented in this paper, and a very good level of agreement between the hardware model and numerical simulation have been demonstrated in the literature (Saighi et al., 2011) – but at a cost of significantly increased circuit size and complexity, obliterating any performance and power advantages of analogue VLSI approach as compared with digital simulation.

The restricted flexibility of the hardware resource, and the limitations on the connectivity imposed by the communication fabric, pose problems related to the optimal mapping of the desired network structure onto the available hardware. As the actual model implemented by the VLSI device cannot be expressed analytically, or even approximated numerically with satisfactory accuracy, due to a large number of second-order effects that are difficult to account for, the validation of the model becomes problematic. A systematic approach to tuning of parameters, and their correspondence to the biological variables, is also difficult to establish. Further mapping issues are related to providing increased robustness against device mismatch (Neftci et al., 2011).

The observability of the internal variables in a hardware neural model is limited to the circuit nodes that have been designed to be available for this purpose. We have provided several mechanisms for outputting internal signals off-chip, however it is impossible to provide a comprehensive solution to monitoring all internal signals due to additional circuit complexity that this introduces.

The difficulties outlined above, which stem from the fundamental limitations inherent in the neuromorphic VLSI approach to neural modelling, point in our view to a conclusion that a great degree of biological realism should not be expected from such devices. Where precise models are required for the study of the brain (e.g. with applications in medicine or neurobiology), numerical methods implemented on digital computers are clearly more suitable.

Nevertheless, the neuromorphic devices should prove useful as tools for carrying out investigations into neural computation, and properties of spiking neural networks. While these have to be investigated at various levels of abstraction and biological plausibility,

the neural and synaptic dynamics based on models of cortical cells might be critical to the intelligent information processing carried out by the biological brains. The circuits presented in this paper offer compact and configurable network implementations, suitable for the exploration of the computational properties of systems with dynamics relevant to cortical network models. In this context, they offer performance and power advantages over digital hardware/simulation solutions. As such, the circuit techniques we have developed should also find use in the construction of neuromorphic systems applied to sensory information processing or biologically inspired robot control, and provide foundations for the design of efficient brain-inspired computational hardware once the operation of the brain and the underlying computational paradigm is better understood.

The physical limitations of silicon hardware are not that different to the limitations imposed on a biological system. The noise and variability are present in both, and so is the need for physical layout constraint that favours local connectivity. The biological solutions to noise tolerance, robustness against faults, and plasticity mechanisms, have evolved being shaped up by a similar set of constraints that the hardware designer is faced with. Hence, there is hope that a better understanding of analogue hardware implementations of neural systems will help to form the hypotheses and computational models that emphasise these important physical implementation constraints, and ultimately help to understand the information processing principles present in the brain, and conversely, that biology will provide the inspiration for the design of future intelligent computing machines.

## Acknowledgements

The authors thank Thomas Wennekers and Arnd Roth for helpful discussions that led to the development of the specifications for the CNL chip. This work has been supported by the UK Engineering and Physical Sciences Research Council (EPSRC), grant number EP/C010841/1.

## References

- Abbott LF, Varela JA, Sen K, Nelson SB. Synaptic depression and cortical gain control. *Science* 1997;275(5297):221–4.
- Ananthanarayanan R, Esser SK, Simon HD, Modha DS. The cat is out of the bag: cortical simulations with 109 neurons and 1013 synapses. In: Supercomputing 09: proceedings of the ACM/IEEE SC2009 conference on high performance networking and computing; 2009. p. 14–20.
- Arthur JV, Boahen K. Recurrently connected silicon neurons with active dendrites for one-shot learning. *IEEE Int Joint Conf Neural Networks* 2004;3:1699–704.
- Ascoli GA, Alonso-Nanclares L, Anderson SA, Barrios-Garcia G, Benavides-Piccione R, Burkhalter A, et al. Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat Rev Neurosci* 2008;9(7):557–68.
- Bi G-Q, Poo M-M. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J Neurosci* 1998;18:10464–72.
- Boahen K. Point to point connectivity between neuromorphic chips using address-events. *IEEE Trans Circuits Syst* 2000;47(5):416–34.
- Bofill-i Petit A, Murray AF. Synchrony detection and amplification by silicon neurons with stdp synapses. *IEEE Trans Neural Netw* 2004;15:1296–304.
- Brumby SP, Galbraith AE, Ham M, Kenyon G, George JS. Visual cortex on a chip: large-scale, real-time functional models of mammalian visual cortex on a GPGPU. In: GPU technology conference; 2010. p. 20–3.
- Connors BW, Gutnick MJ. Intrinsic firing patterns of diverse neocortical neurons. *Trends Neurosci* 1990;13(3):99–104.
- Destexhe A, Crunelli V. Methods for computational neuroscience. *J Neurosci Methods* 2008;169(2):269–70.
- Farquhar E, Hasler P. A bio-physically inspired silicon neuron. Regular Papers. *IEEE Trans Circuits Syst I* 2005;52(3):477–88.
- Fasnacht DB, Indiveri G. A PCI based high-fanout AER mapper with 2 GiB RAM look-up table, 0.8 μs latency and 66 MHz output event-rate. In: Proceedings of conference on information sciences and systems; 2011. p. 1–6.
- Fellous JM, Suri RE. The Roles of Dopamine. In: Arbib MA, editor. *Hand book of brain theory and neural networks*. 2nd Cambridge, MA: MIT Press; 2003. p. 361–5.
- Fidjeland AK, Shanahan MP. Accelerated simulation of spiking neural networks using GPUs. *Proc IEEE World Congress Comput Intell* 2010:536–43.

- Giulioni M, Camilleri P, Dante V, Badoni D, Indiveri G, Braun J, et al. A VLSI network of spiking neurons with plastic fully configurable stop-learning synapses. *IEEE Int Conf Electron Circuits Syst* 2008;678–81.
- Haflinger P, Mahowald M, Watts L. A spike-based learning neuron in analog VLSI. In: Mozer MC, Jordan MI, Petsche T, editors. Advances in neural information processing systems, vol. 9; 1997. p. 692.
- Indiveri G. A low-power adaptive integrate-and-fire neuron circuit. *IEEE Int Symp Circuits Syst* 2003;820–3.
- Indiveri G, Linares-Barranco B, Hamilton TJ, van Schaik A, Etienne-Cummings R, Delbrück T, et al. Neuromorphic silicon neuron circuits. *J Front Neuromorph Eng* 2011;5:1–23.
- Indiveri G, Chicca E, Douglas RA. VLSI array of low-power spiking neurons and bistable synapses with spike-timing dependent plasticity. *IEEE Trans Neural Netw* 2006;17:211–21.
- Izhikevich EM. Simple model of spiking neurons. *IEEE Trans Neural Netw* 2003;14(6):1569–72.
- Izhikevich EM. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb Cortex* 2007;17:2443–52.
- Jin X, Lujan M, Plana LA, Davies S, Temple S, Furber S. Modeling spiking neural networks on SpiNNaker. *Comput Sci Eng* 2010;12(5):91–7.
- Koickal TJ, Hamilton A, Tan SL, Covington JA, Gardner JW, Pearce TC. Analog VLSI circuit implementation of an adaptive neuromorphic olfaction chip. *IEEE Trans Circuits Syst I: Regular Papers* 2007;54(1):60–73.
- LeMasson G, Renaud S, Debay D, Bal T. Feedback inhibition controls spike transfer in hybrid thalamic circuits. *Nature* 2002;417:854–8.
- Lichtsteiner P, Delbrück T. A  $64 \times 64$  AER logarithmic temporal derivative silicon retina. *IEEE PRIME 2005*, EPFL. Res Microelectron Electron 2005;2:202–5.
- Lin J, Merolla P, Arthur J, Boahen K. Programmable connections in neuromorphic grids. In: 49th IEEE midwest symposium on circuits and systems. IEEE Press; 2006. p. 80–4.
- Linares-Barranco B, Sanchez-Sinencio E, Rodriguez-Vazquez A, Huertas JL. A CMOS implementation of FitzHugh-Nagumo neuron model. *IEEE J Solid State Circuits* 1991;26(7):956–65.
- Linares-Barranco B, Serrano-Gotarredona T. Memristance can explain spike-time-dependent-plasticity in neural synapses. In: Nature proceedings; 2009.
- Liu SC, van Schaik A, Minch B, Delbrück T. Event-based 64-channel binaural silicon cochlea with Q-enhancement mechanisms. *IEEE Int Symposium Circuits Systems* 2010;2027–30.
- Mahowald MA, Douglas R. A silicon neuron. *Nature* 1991;354(6354):515–8.
- Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. Interneurons of the neocortical inhibitory system. *Nat Rev Neurosci* 2004;5:793–807.
- Markram H. The blue brain project. *Nat Rev Neurosci* 2006;7(2):153–60.
- Mead CA. Analog VLSI and neural systems. Addison-Wesley; 1989, ISBN 0201059924, 9780201059922.
- Mead CA. Neuromorphic electronic systems. *Proc IEEE* 1990;78:1629–36.
- Merolla P, Arthur J, Shi BE, Boahen K. Expandable networks for neuromorphic chips. *IEEE Trans Circuits Syst I* 2007;54(2):301–11.
- Merolla P, Arthur J, Akopyan F, Imam N, Manohar R, Modha D. A digital neurosynaptic core using embedded crossbar memory with 45 pJ per spike in 45 nm. In: Proc. IEEE custom integrated circuits conference; 2011.
- Millner S, Grubl A, Meier K, Schemmel J, Schwartz M-O. A VLSI implementation of the adaptive exponential integrate-and-fire neuron model. *Adv Neural Inf Process Syst* 2010;23:1642–50.
- Morrison A, Aertsen A, Diesmann M. Spike-timing dependent plasticity in balanced random networks. *Neural Comput* 2007;19:1437–67.
- Morrison A, Diesmann M, Gerstner W. Phenomenological models of synaptic plasticity based on spike timing. *Biol Cybern* 2008;459–78.
- Nakada K, Asai T, Hayashi H. A silicon resonate-and-fire neuron based on the Volterra system. In: Int. symp. on nonlinear theory and its applications; 2005. p. 2–85.
- Neftci E, Chicca E, Indiveri G, Douglas R. A systematic method for configuring VLSI networks of spiking neurons. *Neural Comput* 2011;23(10):2457–97.
- Nowak LG, Azouz R, Sanchez-Vives MV, Gray CM, McCormick DA. Electrophysiological classes of cat primary visual cortical neurons in vivo as revealed by quantitative analyses. *J Neurophysiol* 2003;89:1541–66.
- Patel GN, DeWeerth SP. Analogue VLSI Morris-Lecar neuron. *Electron Lett* 1997;33(12):997–8.
- Pelgrom MJM, Duinmaijer ACJ, Welbers APG. Matching properties of MOS transistors. *IEEE J Solid State Circuits* 1989;24:1433–40.
- Plana LA, Furber SB, Temple S, Khan M, Yebin S, Jian W, et al. A GALS infrastructure for a massively parallel multiprocessor. *IEEE Des Test Comput* 2007;24(5):454–63.
- Rangan V, Ghosh A, Aparin V, Cauwenberghs G. A subthreshold: a VLSI implementation of the Izhikevich simple neuron model. In: Conference proceedings of the international conference of IEEE engineering in medicine and biology society; 2010. p. 4164–7.
- Saighi S, Bornat Y, Tomas J, Le Masson G, Renaud S. A library of analog operators based on the Hodgkin-Huxley formalism for the design of tunable, real-time, silicon neurons. *IEEE Trans Biomed Circuits Syst* 2011;5(1):3–19.
- Schemmel J, Fieres J, Meier K. Wafer-scale integration of analog neural networks. *IEEE Int Joint Conf Neural Netw* 2008;431–8.
- Schemmel J, Bruderle D, Grubl A, Hock M, Meier K, Millner S. Modeling synaptic: a wafer-scale neuromorphic hardware system for large-scale neural modeling. *Proc IEEE Int Symposium Circuits Syst* 2010;1947–50.
- Schoenauer T, Atasoy S, Mehrtash N, Klar H. NeuroPipe-Chip: a digital neuro-processor for spiking neural networks. *IEEE Trans Neural Netw* 2002;13(1):205–13.
- Schultz SR, Jabri MA. Analogue VLSI ‘integrate and-fire’ neuron with frequency adaptation. *Electron Lett* 1995;31(16):1357–8.
- Serrano-Gotarredona R, Oster M, Lichtsteiner P, Linares-Barranco A, Paz-Vicente R, Gomez-Rodriguez F, et al. CAVIAR: a 45k-neuron, 5M-synapse, 12G-connects/s AER hardware sensory-processing-learning-actuating system for high speed visual object recognition and tracking. *IEEE Trans Neural Netw* 2009;20(9):1417–38.
- Sharp T, Patterson C, Furber S. Distributed configuration of massively-parallel simulation on SpiNNaker neuromorphic hardware. *Int Joint Conf Neural Netw* 2011;1099–105.
- Simoni MF, DeWeerth SP. Adaptation in a VLSI model of a neuron. *IEEE Trans Circuits Syst-II: Analog Digital Signal Process* 1999;46(7):967–70.
- Strukov DB, Snider GS, Stewart DR, Williams SR. The missing memristor found. *Nature* 2008;453(7191):80–3.
- Tanaka H, Morie T, Aihara K. A CMOS circuit for STDP with a symmetric time window. *Int Congress Ser* 2007;1301:152–5.
- Toledo-Rodriguez M, Gupta A, Wang Y, Wu C, Markram H, Arbib MA, editor. Neocortex basic neuron types. The handbook of brain theory and neural networks. 2nd ed. Cambridge, MA: MIT Press; 2003. p. 719–25.
- Topol AW, La Tulipe DC, Shi L, Frank DJ, Bernstein K, Steen SE, et al. Three-dimensional integrated circuits. *IBM J Res Dev* 2010;50(4.5):491–506.
- Tsodyks M. Spike-timing-dependent synaptic plasticity – the long road towards understanding neuronal mechanisms of learning and memory. *Trends Neurosci* 2002;25(12):599–600.
- van Schaik A, Jin C, Hamilton TJ. A log-domain implementation of the Izhikevich neuron model. *IEEE Int Symposium Circuits Syst* 2010;4253–6.
- Vogelstein RJ, Mallik U, Vogelstein JT, Cauwenberghs G. Dynamically reconfigurable silicon array of spiking neurons with conductance-based synapses. *IEEE Trans Neural Netw* 2007;18(1):253–65.
- Wang YX, Liu S-C. A two-dimensional configurable active silicon dendritic neuron array. *IEEE Trans Circuits Syst I* 2011;58(9):1549–8328.
- Wijekoon JHB, Dudek P. A simple analogue VLSI circuit of a cortical neuron. *IEEE Int Conf Electron Circuits Syst* 2006;1344–7.
- Wijekoon JHB, Dudek P. Integrated circuit implementation of a cortical neuron. *IEEE Int Symposium Circuits Syst* 2008a;1784–8.
- Wijekoon JHB, Dudek P. Compact silicon neuron circuit with spiking and bursting behaviour. *Neural Netw* 2008b;21(2–3):524–34.
- Wijekoon JHB, Dudek P. A CMOS circuit implementation of a spiking neuron with bursting and adaptation on a biological timescale. *IEEE Biomedical Circuits and Systems Conference, BioCAS*, 2009, pp. 193–196.
- Wijekoon JHB. Mixed signal VLSI circuit implementation of the cortical microcircuit models. PhD thesis. School of Electrical and Electronic Engineering, The University of Manchester, UK, 2011.
- Wijekoon JHB, Dudek P. Analogue CMOS circuit implementation of a dopamine modulated synapse. *IEEE Int Symposium Circuits Syst* 2011;877–80.
- Zaghlool KA, Boahen K. Optic nerve signals in a neuromorphic chip II: testing and results. *IEEE Trans Biomed Eng* 2004;51(4):667–75.
- Zaghlool KA, Boahen K. A silicon retina that reproduces signals in the optic nerve. *J Neural Eng* 2006;3:257.