

# A Current-mode Spiking Neural Classifier with Lumped Dendritic Nonlinearity

Amitava Banerjee, Sougata Kar\*, Subhrajit Roy, Aritra Bhaduri and Arindam Basu†

School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore  
 Email: \*sougata@ntu.edu.sg; †arindam.basu@ntu.edu.sg

**Abstract**—We present the current mode implementation of a spiking neural classifier with lumped square law dendritic nonlinearity. It has been shown earlier that such a system with binary synapses can be trained with structural plasticity algorithms to achieve comparable classification accuracy with less synaptic resources than conventional algorithms. Hence, in our address event based implementation, we save  $2 - 12X$  memory resources in storing connectivity information. The chip fabricated in  $0.35\mu\text{m}$  CMOS has 8 dendrites per cell and uses two opposing cells per class to cancel common mode inputs. Preliminary results show the chip is functional and dissipates  $30\text{nW}$  of static power per neuronal cell and  $422\text{pJ/spike}$ .

## I. INTRODUCTION AND MOTIVATION

Spiking Neural Networks (SNN), considered to be the third generation of neural networks, was proposed due to the advent of neurobiological evidences [1] which suggested that biological neural systems use timing of action potentials or ‘spikes’ to convey information. These types of network are considered to be more bio-realistic and computationally more powerful than its predecessors. The computational units of SNN communicate through spikes which are equivalent to noise robust digital pulses, so SNN are well suited for low-power, low-voltage very large scale integrated circuit (VLSI) implementations. Thus, while theoretical studies on SNN have progressed on one hand, in parallel, neuromorphic engineers have implemented low-power VLSI circuits that emulate sensory systems [2], [3] and higher cognitive functions like learning and memory.

With the arrival of Brain Machine Interfaces, wearable devices, internet of things, etc. there is also a pressing need for area and energy efficient on-chip implementation of supervised classifiers. To fulfill this demand we are looking for hardware implementation of area and power optimized spiking neural classifiers. Recently in [4], [5], structures have been proposed inspired by the non-linear properties of dendrites in neurons which require much less synaptic resources than other neuromorphic designs. Moreover, the learning of these structures involves network rewiring of *binary* synapses which is comparable to the structural plasticity observed in biological neural systems. Specifically the learning algorithm tries to find the best ‘combinations’ of input on each dendrite to improve the performance. This choice of connectivity can be easily implemented in hardware by exploiting address event representation (AER) protocols, commonly used in current Neuromorphic systems, where the connection matrix is stored in memory.

In this work, we present a current mode implementation of the above neural classifier and show preliminary results

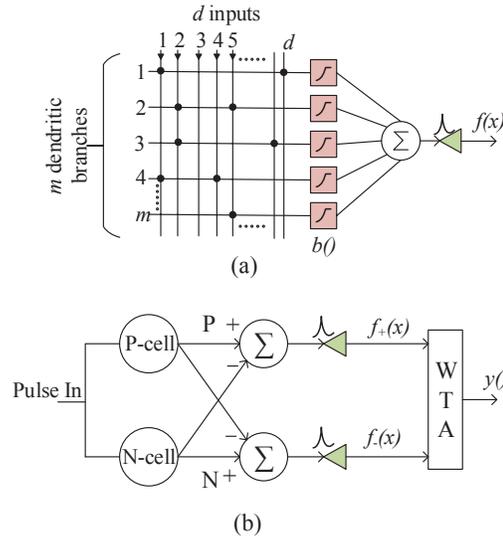


Fig. 1: Architecture of two-class pattern classifier that uses WTA to compare the outputs of two non-linear cells trained to respond to patterns of classes (+) and (-) respectively.

from a chip fabricated in  $0.35\mu\text{m}$  CMOS to prove functional correctness.

## II. BACKGROUND AND THEORY

In [4], [6], the authors have suggested classifiers employing neurons with non-linear dendrites (NNLD) and binary synapses. Due to the presence of binary synapses, the learning in these type of architectures happen not by weight update but by morphological changes of the connections between inputs and dendrites. Thus, these architectures are amenable for neuromorphic implementation employing AER protocols. We invite the reader to look into [4], [6] for a detailed description of the architectures and learning rules. In this paper, we present a circuit to implement the method proposed in [6] which has comparable performance as other spike based classifiers such as [7] but use  $12X$  less synaptic resources. Hence, our implementation requires correspondingly less memory.

Briefly, the classifier architecture comprises of two NNLD, each of which has  $m$  identical dendritic branches with  $k$  excitatory synapses per branch. If a  $d$ -dimensional input vector  $\mathbf{x}$  ( $d \gg k$ ) is input to this system, then each synapse is excited by any one of the  $d$  dimensions and the output response of  $j^{\text{th}}$  dendritic branch is given by  $z_j = b(\sum_{i=1}^k w_{ij}x_{ij})$ . Here  $b()$  is a model of the dendritic nonlinearity given by  $b(x)=x^2/x_{thr}$ ,  $w_{ij}$  is the synaptic weight of the  $i^{\text{th}}$  synapse on

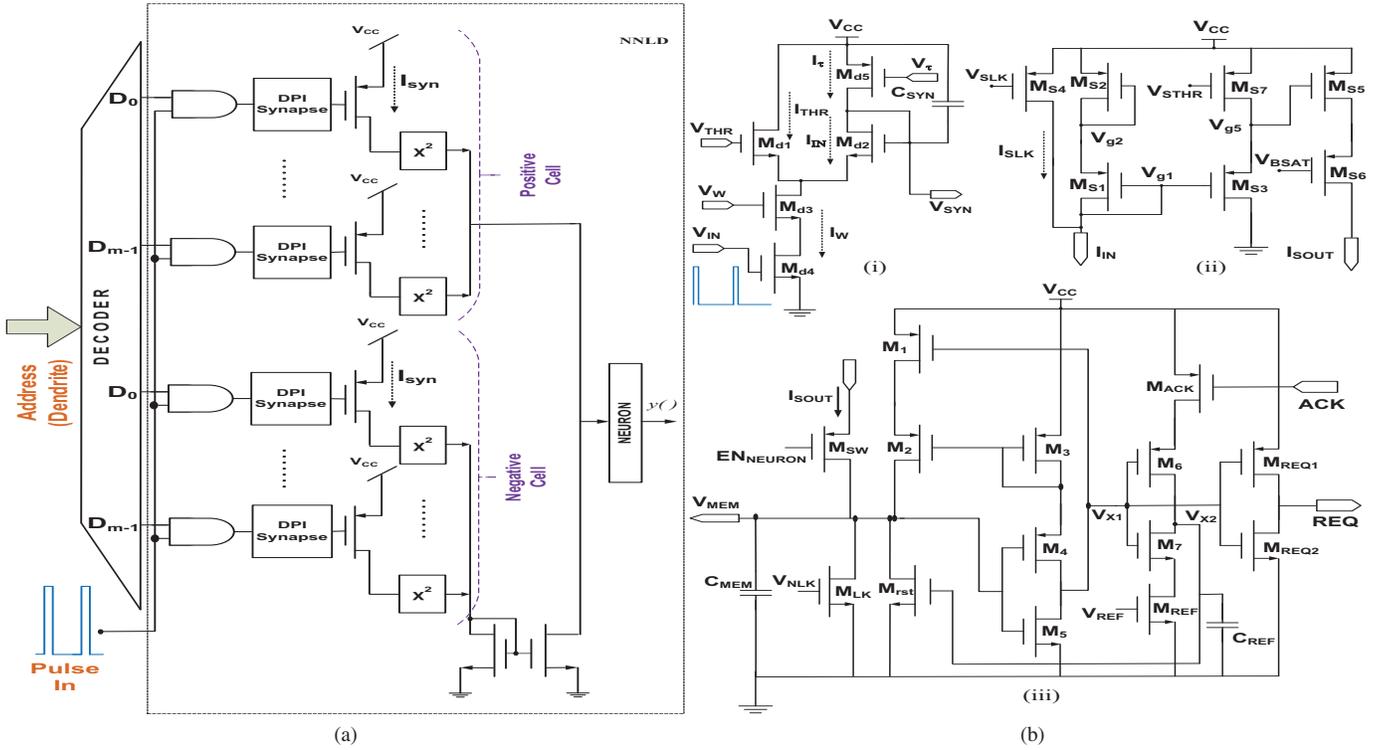


Fig. 2: (a) Architecture of the Neuromorphic IC and (b) Circuit diagram of sub-blocks (i) synapse, (ii) squaring and (iii) neuron.

the  $j^{th}$  branch and  $x_{ij}$  the corresponding input. The overall output  $f(\mathbf{x})$  of a single neuronal cell is given by:

$$f(\mathbf{x}) = f\left(\sum_{j=1}^m b\left(\sum_{i=1}^k w_{ij}x_{ij}\right)\right) = K \times \sum_{j=1}^m b\left(\sum_{i=1}^k w_{ij}x_{ij}\right) \quad (1)$$

where  $f()$  denotes the linear neuronal current-frequency conversion function. In [6], the output of the two cell architecture was calculated by noting the difference of the output of the two NNLD:

$$y = g[f_+(P) - f_-(N)] = g[K \times (P - N)] \quad (2)$$

where the function  $g$  depends on the task. For example, it is a heaviside function for classification.

In [6], each of  $f_+$  and  $f_-$  are spike trains obtained from a spiking neuron and the comparison is done in a WTA. However, for our implementation, we found this has the problem of high common mode current at the input of the spiking neurons leading to frequency saturation. Hence, we modified the architecture to a differential one as shown in Fig. 1 with the following equation:

$$y = g[f_+(P - N) - f_-(N - P)] = g[2K \times (P - N)] \quad (3)$$

### III. VLSI IMPLEMENTATION OF NEUROMORPHIC IC

#### A. Architecture

The VLSI architecture of the implemented neuromorphic IC is shown in Fig. 2, where AER is used to provide the

synaptic input. A Differential Pair Integrator (DPI) circuit has been used to implement the synaptic function of neuron where it is possible to achieve the linear filtering property of each synapse [8] by proper biasing. This linear filtering property is implemented here to replace all the synapses in one dendritic connection by a single shared synapse drastically reducing the effective layout area of the IC.

This IC is interfaced with an FPGA controller which generates input spikes and addresses as shown in Fig. 3. The learnt sparse connectivity matrix is stored inside the FPGA memory in a very compressed form by using two look up tables. Based on the input line address, the controller reads the connectivity information of the address line from the memory, and generates  $n$ -bit decoder address to route the spike to the proper dendrite. Spikes from the FPGA output reaches the synapse circuit input through a  $n$ -bit to  $2^n$ -bit address decoder followed by digital switches. There are total  $2^{(n-1)}$  P-cell and  $2^{(n-1)}$  N-cell dendritic branches connected to a NEURON block, where the shared synapse circuit in each dendrite is followed by a square-law nonlinear circuit. The difference of the total current of  $2^{(n-1)}$  P-cells and  $2^{(n-1)}$  N-cells appears to the input of the neuron block and is converted to equivalent output spikes. The output of the neuron block is digitized for the proper handshaking of the IC with the FPGA in the form of a REQ and ACK signal. We have only implemented one cell to compute one term in eq (3). The 2nd term is computed by reusing the same neuron with flipped inputs and the final decision is taken on the FPGA.

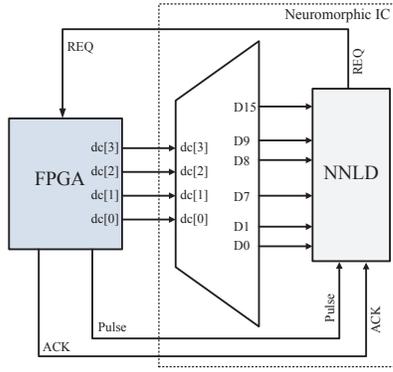


Fig. 3: FPGA-IC interface.

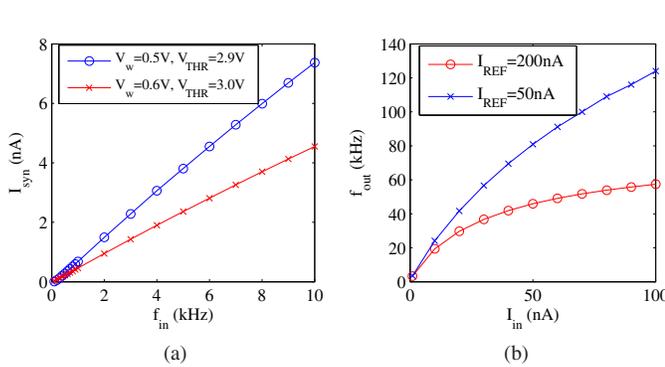


Fig. 4: Simulation Results of (a) Synapse Block showing linearity of Average Current with Input Frequency and (b) Neuron Block showing Output Frequency vs Input Current Relationship.

The neuromorphic IC has been fabricated in AMS 0.35 $\mu\text{m}$  CMOS technology and occupies 850 $\mu\text{m}^2$  area. In this section and in the following sections, we will briefly describe the circuit implementation of these blocks and show simulations and experiment results of the effect of different biasing conditions.

### B. Synapse Circuit

The circuit schematic shown in Fig. 2(b) is a DPI circuit that converts presynaptic voltage pulses into postsynaptic currents ( $I_{syn}(t)$ ). In this circuit, transistors operate in subthreshold regime. As mentioned in [8], the bias voltages  $V_w$  and  $V_\tau$  are used to set  $I_w \gg I_\tau$ , which simplifies this nonlinear circuit to a canonical first-order low pass filter. For an input spike arriving at  $t_i^-$  and ending at  $t_i^+$  to the DPI synapse, the rise time of  $I_{syn}$  is very small. The discharge profile can be modeled by the Equation 4.

$$I_{syn}(t) = I_0 e^{-\frac{t-t_i^+}{\tau_s}} \quad (4)$$

where  $\tau_s = \frac{C_{SYN} U_T}{\kappa I_\tau}$ ,  $\kappa$  is the subthreshold slope factor, and  $U_T$  is the thermal voltage. This synapse circuit when typically excited with continuous pulse train of average frequency of  $f_{in}$  and pulse duration  $\Delta t$ , its steady-state output current (average of the output transient current) will hold a linear relationship with the input frequency. The relationship has been derived in [8] as

$$I_{syn}^-(t) = K_{syn} \times (\Delta t) \times f_{in} \quad (5)$$

Here  $K_{syn}$  is a constant set by  $V_w$ ,  $V_{THR}$  and  $V_\tau$ . Simulation results for this block shown in Fig. 4(a) demonstrate the expected linearity.

### C. Square Block Circuit

We have designed the current mode squaring circuit given in Fig. 2(c) as described in [5]. Transistors  $M_{S2}$ ,  $M_{S1}$ ,  $M_{S3}$  and  $M_{S5}$  form a translinear loop. Hence, the current through  $M_{S5}$  is expressed as given in Equation 6. The transistor  $M_{S5}$  is biased to pass a maximum current of  $I_{sat}$  (set by  $V_{BSAT}$ ).

$$I_{SOUT} = \frac{(I_{in} - I_{SLK})^2}{I_{STHR}} \quad (6)$$

$I_{STHR}$  is the dc current through  $M_{S4}$  set by its Gate voltage ( $V_{STHR}$ ).

### D. Neuron Circuit

The circuit diagram of the implemented neuron has been shown in Figure 2(d). Here is one integrating capacitor  $C_{MEM}$ , inverter ( $M_4$  and  $M_5$ ) with positive feedback circuit ( $M_1$ ,  $M_2$  and  $M_3$ ) to  $V_{MEM}$ ,  $M_{REF}$  and  $M_{rst}$  to control the refractory period of the neuron, leakage transistor ( $M_{LK}$ ) for controlling the minimum  $I_{OUT}$  to start charging of  $C_{MEM}$ ,  $M_{REQ1}$  and  $M_{REQ2}$  to generate REQ signal to the FPGA, and  $M_{ACK}$  to be turned on for an ACK from FPGA. The output current of the square block ( $I_{SOUT}$ ) is integrated by  $C_{MEM}$  to generate the profile of the Membrane voltage ( $V_{MEM}$ ). As  $V_{MEM}$  increases and approaches the switching voltage of the inverter ( $M_4$ - $M_5$ ), the feedback current starts flowing through  $M_1$ - $M_2$  causing a sudden rapid increase of the  $V_{MEM}$  profile.  $C_{MEM}$  is quickly discharged back to ground through the reset transistor  $M_{rst}$ , the on-resistance of which is controlled by  $V_{X2}$ . When  $C_{MEM}$  is fully discharged,  $V_{X1}$  is driven back to  $V_{CC}$  causing turning on of  $M_7$ . As long as  $V_{X2}$  is sufficiently high,  $M_{rst}$  is active and is clamped to ground. This is called the refractory period of the neuron, and hence there will be no REQ pulse generated by the Neuron during this period. The entire current at the output of the square block will pass through  $M_{rst}$ . The input-output characteristic (for negligible refractory period) can be represented by equation 7.

$$f_{OUT} = K_{neu} \times (I_{SOUT} - I_{NLK}) \quad (7)$$

Figure 4(b) displays the above relation of  $f_{OUT}$  and  $I_{SOUT}$  and shows the increased linearity when refractory period is reduced.

Hence, combining equations 5, 6, and 7 we get the nonlinear relationship of the IC as given in equation 8 -

$$f_{OUT} = K_{neu} \left[ \frac{(K_{syn} \times f_{in} - I_{SLK})^2}{I_{STHR}} - I_{NLK} \right] \quad (8)$$

The neuromorphic IC and its FPGA controller interfacing is shown in Figure 3.

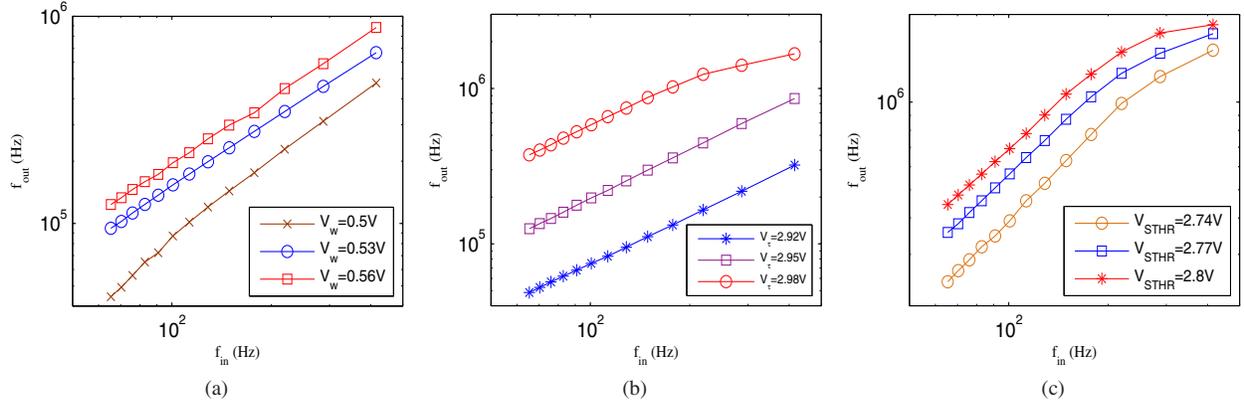


Fig. 5:  $f_{out} - f_{in}$  characteristics with parameter variation. (a)  $V_w$  variation, (b)  $V_\tau$  variation, (c)  $V_{STHR}$  variation

#### IV. MEASUREMENT RESULTS

##### A. Characterization

Microphotograph of the fabricated chip is shown in Fig. 6. The DPI synapse, square block and the neuron combined together forms a basic unit of the fabricated chip governed by equation (8). We have found all blocks to be functional and show some characterization results here. First, we show the effect of increasing synaptic weight by changing  $V_w$  from 0.5V to 0.56V. Figure 5a shows the measured plots where the slope increases as expected. Next, we changed  $V_\tau$  of the synapse from 2.92V to 2.98V. As this is a PMOS transistor, the fall time constant of the synapse increases with  $V_\tau$ . This increases the average output current of the synapse as well. Fig. 5b plots the desired transfer characteristics for different  $V_\tau$ . In Fig. 5c, the input and output frequencies are plotted for different threshold voltage of the square block. As the threshold voltage increases, current decreases (PMOS transistor). Therefore the output current of the square block also increases which increases the output frequency as well.

##### B. Pattern Classification

For a preliminary case study, a two-class classification problem of 400 dimensional binary spike trains as discussed in [6] is shown here. The connectivity matrix for the task is learnt in software and stored in FPGA memory. A 400 ms sequence of pulse trains along with the input line addresses are generated as input where the two binary values are mapped to poisson spike trains with mean firing rates of  $f_{HIGH} = 250$  Hz and  $f_{LOW} = 1$  Hz respectively. The computed classification error for 100 random patterns split equally into the two classes is found to be 16%, where as the software shows an error of 8%. The difference is attributed to mismatch among dendritic branches and non square law behavior of the nonlinearity. Both of these will be addressed in future through calibration.

#### V. CONCLUSION AND DISCUSSION

In this paper, we present the VLSI circuit design in  $0.35\mu\text{m}$  CMOS of a neuromorphic spike based classifier with 8 non-linear dendrites per neuron and 2 opponent neurons per class.

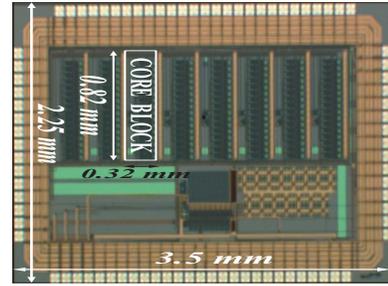


Fig. 6: Microphotograph of the neuromorphic IC.

We present characterization results to prove functionality of all sub-blocks and finally some preliminary results of classifying complex spike-based high dimensional binary patterns. The classification accuracy is 84% for 100 patterns into two classes—8% less than software due to transistor mismatch and non-square law behavior of dendrites. Future work will use calibration to improve accuracy.

#### REFERENCES

- [1] Fred Rieke, Davd Warland, Rob de Ruyter van Steveninck, and William Bialek, *Spikes: Exploring the Neural Code*, MIT Press, Cambridge, MA, USA, 1999.
- [2] V. Chan, S.C. Liu, and A.V. Schaik, "AER EAR: A matched silicon cochlea pair with address event representation interface," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 1, pp. 48–59, jan. 2007.
- [3] E. Culurciello, R. Etienne-Cummings, and K.A. Boahen, "A biomorphic digital image sensor," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 2, pp. 281–294, feb 2003.
- [4] S. Roy, A. Basu, and S. Hussain, "Hardware efficient, Neuromorphic Dendritically Enhanced Readout for Liquid State Machines," in *IEEE BioCAS*, Nov 2013, pp. 302–305.
- [5] S. Hussain, R. Gopalakrishnan, A. Basu, and S. C. Liu, "Morphological Learning: Increased Memory Capacity of Neuromorphic Systems with Binary Synapses Exploiting AER Based Reconfiguration," in *IEEE Intl. Joint Conference on Neural Networks*, 2013.
- [6] S. Hussain, A. Basu, and S. C. Liu, "ImprovedMargin Multi-Class Classification using Dendritic Neurons with Morphological Learning," in *IEEE ISCAS*, May 2014, pp. 2640–43.
- [7] P. O'Connor, D. Neil, S. C. Liu., T. Delbruck, and M. Pfeiffer, "Real-Time Classification and Sensor Fusion with a Spiking Deep Belief Network," *Frontiers in Neuroscience*, vol. 7, pp. 1–13, Oct. 2013.
- [8] C. Bartolozzi and G. Indiveri, "Synaptic Dynamics in Analog VLSI," *Neural Computation*, vol. 19, no. 10, pp. 2581–2603, 2007.