

A Synchronized Axon Hillock Neuron for Memristive Neuromorphic Systems

Ryan Weiss, Gangotree Chakma, and Garrett S. Rose

IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, Massachusetts, August 2017.

©2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Citation Information (BibTex):

```
@INPROCEEDINGS{WeissMWSCAS:2017,  
  author="Ryan Weiss and Gangotree Chakma and Garrett S.  
    Rose",  
  title="A Synchronized Axon Hillock Neuron for Memristive  
    Neuromorphic Systems",  
  booktitle="Proceedings of {IEEE} International Midwest  
    Symposium on Circuits and Systems{MWSCAS}"  
  month="August",  
  year="2017",  
  address="Boston, Massachusetts"  
}
```

A Synchronized Axon Hillock Neuron for Memristive Neuromorphic Systems

Ryan Weiss, Gangotree Chakma and Garrett S. Rose
Department of Electrical Engineering and Computer Science
University of Tennessee, Knoxville
Knoxville, Tennessee 37996 USA
Email: {rweiss2, gchakma, garose}@utk.edu

Abstract—In this paper we present circuit techniques to optimize analog neurons specifically for operation in memristive neuromorphic systems. Since the peripheral circuits and control signals of the system are digital in nature, we take a mixed-signal circuit design approach to leverage analog computation in multiplying and accumulating digital input spikes and generate binary spikes as outputs to be consistent with surrounding synchronous digital logic circuits. A novel approach for synchronization is leveraged based on domino logic. The principal advantage of utilizing analog neurons within an overall digital system design is to ensure efficiency in size and power consumption. Energy per spike was determined to be 20 fJ, based on Cadence Spectre simulations of the proposed domino-based neural circuit.

I. INTRODUCTION

The human brain represents a very complex architecture for computing and processing data. A deep understanding for how biological neural systems provide both accuracy and energy efficiency remains a goal for many researchers. During this continuing process, researchers have developed neural network models which emulate brain functionality, including deep learning techniques used to train such systems. However, offline training approaches are still inefficient compared to the low energy consumption exhibited by the human brain. Further, for many brain-inspired approaches, researchers often leverage von Neumann machines to implement neural networks. The training and execution of many machine learning systems require a large amount of energy and area resources which is not efficient for embedded systems applications. Thus, many researchers try to overcome such limitations by adapting unconventional computing architectures, such as neuromorphic computing [1]. Memristive neuromorphic computing leverages metal-oxide memristors [2], [3] to minimize the area and energy consumption.

In this research, the proposed memristive neuromorphic architecture is constructed from both analog and digital components. Further, communication between neurons is synchronous, enabling a simplified approach for controlling the timing of spiky information. Artificial synapses are implemented using metal-oxide memristors to store weight values and transmit analog weighted results to post-synaptic neurons. The neuron uses the analog output of the synapse to produce a firing event (or spike) that is synchronized with the system. Further, the system considered leverages the relative timing of synchronous digital spikes to perform spike time dependent

plasticity (STDP) based online learning. The synchronous nature of the system helps simplify the necessary control logic for STDP, thus reducing area utilization. Here we describe an analog axon hillock neuron design approach leveraging domino logic to synchronize output spikes. We further show that the proposed provides desired energy and area efficiency for emerging CMOS-memristive neuromorphic implementations.

II. MEMRISTIVE SYNAPSE

The synapse of a neural network stores a weight value that relates the strength of a firing spike of a pre-synaptic neuron to the following post-synaptic neuron. The weights characterize the required activation of the preceding neuron to activate the following neuron. This relationship is created by the output current of the synapse.

The synapse design considered here uses a twin memristor configuration (shown in Fig. 1) to store the weight value. Memristors are two terminal nanoscale and non-volatile devices first theorized by Chua [2] in 1971. The non-volatile nature of these devices provide opportunities for high efficiency in area and energy consumption. A memristor provides multiple resistance levels between the low resistance state (LRS) and high resistance state (HRS). The LRS and HRS of any memristor are dependent on the switching material, process conditions, noise and environmental conditions. Several materials used to build memristors for their switching behavior, including TaO_x [3], TiO₂ [4], and HfO_x [5]. All of these memristors are differentiated according to LRS values, LRS to HRS ratios, threshold voltage, and switching time. For this design a suitable range of LRS and HRS has been considered based on the literature.

A. Synapse Design and Timing

The synapse uses a digital logic block to drive each memristor in the pair to a high voltage. The two phases of operation examined here are the accumulation phase and learning phase. Other phases important in practice but beyond the scope of this work are the forming phase and programming phase, which enable memristor behavior and set the pair to an initial weight.

During the accumulation phase, the synapse produces an output current, I_{in} , that is fed into the neuron. The difference in the resistance values of the memristors in each pair creates

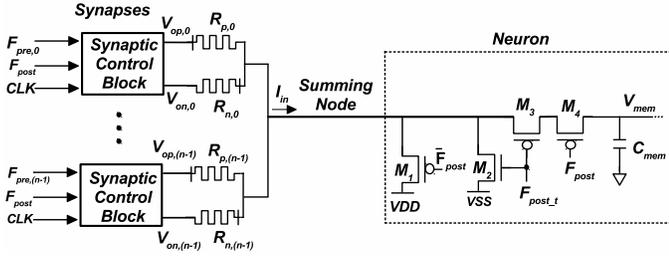


Fig. 1: Synaptic input illustrating the “twin memristor” approach to providing both positive and negative weights. Also shown are the FETs used in the learning phase and integrator representing the input of a neuron.

a positive or negative weight. Positive weights are created when the output current is adding charge into the neuron. This happens when R_p in Fig. 1 is less than R_n . Negative weights are created when the output current is pulling charge from the neuron. This is accomplished when R_n is less than R_p .

During the learning phase, the synapse updates the weights by changing the resistance of the memristors.

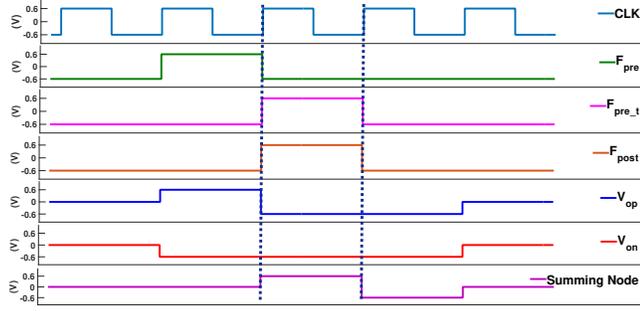


Fig. 2: Timing diagram of LTP

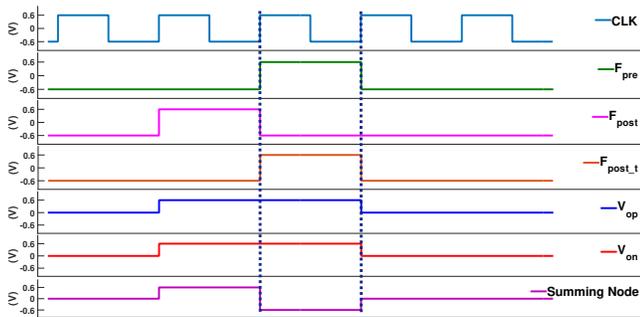


Fig. 3: Timing diagram of LTD

The waveforms in Fig. 2 and 3 show the signals used by the synapse. To update the weight of the synapse, the voltage across the memristor pair needs to be larger than the threshold voltage for the memristor [6]. To accomplish this the neuron drives the summing node while the synapse drives a voltage of opposite polarity from the left. Long term potentiation is initiated when the synapse causes the neuron to fire. This

occurs when the synapse fires (F_{pre}) the clock cycle before the neuron fires (F_{post}), illustrated in Fig. 2. Long term depression (Fig. 3) is performed on the synapse when it fires the clock cycle after the output neuron fires. During the neuron’s output fire, the neuron drives the summing node high for one clock cycle and then low for the next.

III. NEURON

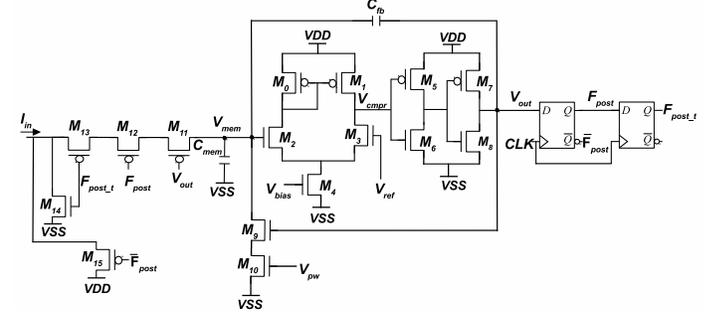


Fig. 4: Implementation of Axon-Hillock neuron with a comparator for variable threshold.

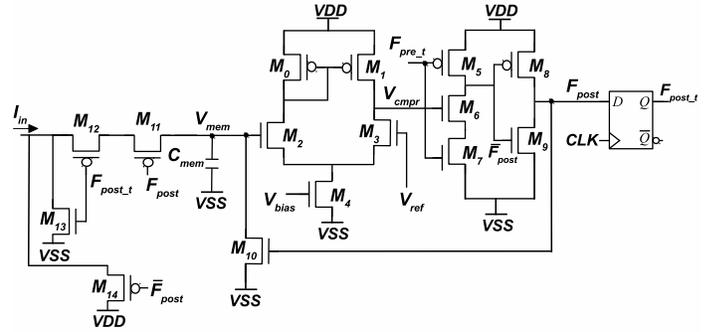


Fig. 5: Proposed Implementation of synchronous neuron with a comparator for variable threshold.

The axon hillock neuron first proposed by Carver Mead [1] takes an input current, integrates the input current on a capacitor and outputs a voltage spike upon crossing a threshold. The axon hillock circuit considered has a variable threshold that triggers a fire when the stored voltage on the capacitor C_{mem} reaches the voltage V_{ref} . The output spike width, refractory period and reset time is dependent on the sizing of the capacitors and transistors C_{mem} , C_{fb} , M_9 , and M_{10} . For this project the output spike must be synchronized with the system clock.

Neuron operation is determined by the pre-neuron synapses. If there are pre-neuron synapse fires, the neuron sums and integrates the weight dependent output currents of the synapses and holds the charge on the capacitor until there is another input into the neuron. When the charge on the capacitor goes above the threshold voltage, the neuron fires and resets the capacitor. The neuron has a refractory period of two clock cycles that it will not accumulate inputs from the pre-neuron

synapses. The output fire of the neuron drives the post-neuron synapses to fire into the next neuron layer in the network. The output of the neuron also controls the weight updates of its pre-neuron synapses.

The proposed neuron design leverages the synchronous system and control signals used to operate the system with the ideas presented for the axon hillock neuron. The design follows the same basic functionality of integrating the input current and triggering a firing event upon crossing the threshold. In the proposed design, the control signals synchronize the timing of neuron output spikes.

A. Neuron Input Circuit Design

The current input of the neuron flows from the memristors of the synapse through a series of PFETs into the input capacitor. The proposed design uses two PFETs that block the capacitor C_{mem} from the voltage drive of the summing node during learning but pass current into the capacitor during accumulation. Design specifications for the PFETs and capacitor are based on the available resistance levels of the memristors, as well as clock speed. Analogous to the standard axon hillock neuron, the proposed domino based design uses the PFETs at the input to define the refractory period. For the domino based neuron, the refractory period is defined by clock cycles needed in the learning phase and is not determined by sizing the capacitor C_{fb} or NFETs M_9 and M_{10} . Results presented for this implementation assumes a defined refractory period of two clock cycle.

For this design, the sizing of the PFETs and C_{mem} are intended to be small to allow for integrating many neurons per chip. A high speed clock is used to keep the size of C_{mem} small. The sizes of the PFETs are also small because as the r_{ds} resistance of the PFETs increases the input current, I_{in} , decreases. Keeping the size of the PFETs small also keeps the capacitance the neuron output needs to drive small.

B. Neuron Output Circuit Design

The neural network structure in this research uses a synchronous digital logic system to implement the weight update of the synapses. Because of this synchronous approach the neuron must output a synchronous signal to the synapse to update its weight upon a post-synapse neuron fire. To synchronize the asynchronous analog integration and comparison computations performed by the neuron, it outputs a clocked digital pulse. To align the output signal with the pre-neuron synapse a D-flip flop is used to capture the output pulse. The proposed design uses a domino logic inverter to control pulse timing in a way that reduces transistor count and energy.

For comparison purposes, a flip flop synchronized axon hillock is considered. This neuron uses two inverters to drive the output to the positive rail when the voltage stored on C_{mem} crosses the threshold V_{ref} . The inverter drives the voltage V_{mem} higher than the threshold by positive feedback through C_{fb} . The output turns on NFET M_9 and stores the output on the first D flip flop. The D flip flop outputs the signal F_{post} . The voltage V_{pw} is set to equalize the output pulse width of

the neuron with the clock period. This allows the D flip flop to capture the output of the neuron only on the next clock cycle. If the pulse width of the neuron is not equal to the clock period, the D flip flop will either miss spikes or drive a high output for more than one clock cycle. The output fire signal F_{post} can also be delayed by a clock cycle. This occurs when the integrated capacitance crosses the threshold and the delay to V_{out} going high causes the D flip flop to capture it on the next clock cycle. The neuron needs a built in refractory period for resetting C_{mem} , which is overshadowed by the two clock cycle refractory period for learning.

C. Pulse Control Inverter

The proposed domino based design incorporates the control signals needed for the learning phase in the computing phase. The design feeds the delayed fire signal of the preceding neuron into the subsequent neuron. This signal drives a domino logic inverter that evaluates the neuron's output signal F_{post} at the appropriate time. Since the evaluation of the comparison is performed using domino logic, the output signal will not fall to negative rail until the end of the clock period when $F_{pre,t}$ goes low. This means the functionality of the output is restricted to driving the necessary gates for learning and computing. The output of the neuron is not internally connected to the input to set the refractory period. The reset of the capacitor is immediately driven, while the refractory period is still defined by F_{post} and $F_{post,t}$ signals on the input PFETs.

IV. SIMULATION RESULTS

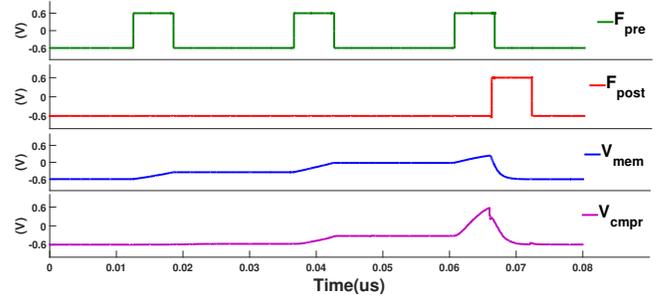


Fig. 6: Simulation result for the analog neuron with domino logic.

In the examples provided in Fig. 6 and Fig. 7 the neurons both take three spikes through the synapse to cause an output fire. Both designs function as intended, taking inputs from the synapses and integrating them on V_{mem} and generating an output fire pulse F_{post} . For the standard axon hillock, V_{out} is captured on a D flip-flop to produce F_{post} . As described above, V_{out} is designed to be high for the one clock period. The concern for the differential amplifier in the standard axon hillock is the rise time, while for the proposed domino based design it is the fall time. Both designs require that the voltage V_{cmpr} is quickly driven high so as not to miss a fire. The proposed domino design does not require the same high speed, and is limited by the reset time of V_{cmpr} .

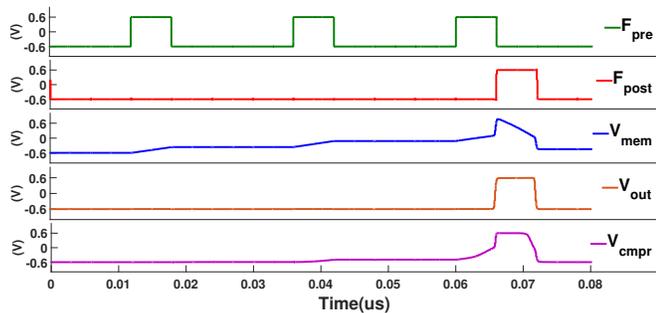


Fig. 7: Simulation result for the analog axon-hillock neuron.

TABLE I: Performance Metrics for Synchronous AH Neurons

Metric	Axon Hillock (Flop)	Axon Hillock (Domino)
Transistor Count	41	28
Energy per Spike (fJ)	140	20
Idle Power (μ W)	0.417	2.55
Accumulation Power (μ W)	4.76	6.10
Frequency (MHz)	167	167

Table I shows simulation results for the neurons presented in the 65nm process. The simulation setup between neurons is intended to be equivalent. The idle power is an average of power used between accumulations and fires. The idle power is much lower for the domino logic because there is almost no current flow through M5, M6, and M7. In the standard axon hillock, as the differential amplifier output voltage increases, the current through M5 and M6 increases. This causes a linear increase in power with accumulated voltage for both idle and accumulation power. The accumulation power is close to the same since the same differential amplifier is used. However, because the differential amplifier needs to reliably control an inverter in the axon hillock, the bias current is greater. The accumulation power averages all time where the voltage V_{mem} is accumulating current from the synapse output. For the domino, this constitutes the entire clock cycle for every F_{pre} . For the flop based approach, the accumulation power includes the accumulation of V_{mem} from the synapse fire, and truncates the power consumed when V_{mem} is driven high by the positive feedback loop C_{fb} . This is not included in the accumulation power because it is considered part of the energy per spike. Energy per spike for the domino includes the energy used in the clock cycles that produce F_{post} and $F_{post,t}$, while for the flop approach it is measured from when V_{out} goes high.

V. DISCUSSION

The importance of synchronizing is the ability to reliably use the output to perform learning for memristive synapses. The synchronous nature of our system allows for advantages in terms of power and reliability. However, since the system requires a high speed clock to keep the feature size small,

issues can arise from the clock. The affects of clock skew could require control of delay of clock signals used for capturing the outputs. Also not rigorously defined is the fan out and fan in limits of the circuit. The fan in into a neuron should be able to handle a multitude of synapses since the input is summed currents. The fan out capability is determined by the strength of the domino logic inverter. If the load capacitance is excessively high, the drive strength of the inverter will need to be increased and require larger devices and more energy.

VI. CONCLUSION

In this paper, we have presented the design of a synchronous analog neuron for robust low power activity. The neuron was designed to work for memristive synapses that perform online learning. Given the synchronous nature of our system, we take advantage of the signals necessary for the learning and utilize them in reducing power and guaranteeing timing. From our simulation results we see that the proposed circuit reduces energy and power along with transistor count.

ACKNOWLEDGMENT

The authors would like to thank Dr. Mark Dean, Md. Musabbir Adnan, Sagarvarma Sayyaparaju, and Sherif Amer from the University of Tennessee, Knoxville for interesting and useful discussions on this topic.

This material is based in part upon research sponsored by Air Force Research Laboratory under agreement number FA8750-16-0065 and the National Science Foundation under Grant No. NCS-FO-1631472. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

REFERENCES

- [1] C. A. Mead, in *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989.
- [2] L. O. Chua, "Memristor-the missing circuit element," *IEEE Transactions on Circuit Theory*, vol. 18, no. 5, pp. 507–519, September 1971.
- [3] J. J. Yang, M. Zhang, J. P. Strachan, F. Miao, M. D. Pickett, R. D. Kelley, G. Medeiros-Ribeiro, and R. S. Williams, "High switching endurance in taox memristive devices," *Applied Physics Letters*, vol. 97, no. 23, p. 232102, 2010.
- [4] G. Medeiros-Ribeiro, F. Perner, R. Carter, H. Abdalla, M. D. Pickett, and R. S. Williams, "Lognormal switching times for titanium dioxide bipolar memristors: origin and resolution," *Nanotechnology*, vol. 22, no. 9, p. 095702, 2011.
- [5] H. Lee, Y. Chen, P. Chen, T. Wu, F. Chen, C. Wang, P. Tzeng, M.-J. Tsai, and C. Lien, "Low-power and nanosecond switching in robust hafnium oxide resistive memory with a thin ti cap," *IEEE Electron Device Letters*, vol. 31, no. 1, pp. 44–46, 2010.
- [6] S. Kvatinsky, M. Ramadan, E. G. Friedman, and A. Kolodny, "Vteam: A general model for voltage-controlled memristors," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, no. 8, pp. 786–790, 2015.