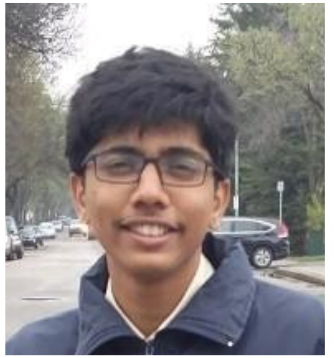
Circuit Cost Reduction for Online STDP Using NIPIN Selector as Timekeeping Device in RRAM Synapse

Ashwin Sanjay Lele, Anand Naik, Lakshya Bandhu, Bhaskar Das, Udayan Ganguly

Department of Electrical Engineering

Indian Institute of Technology Bombay



Speaker

Ashwin Sanjay Lele

2020 IEEE International Symposium of Circuits and Systems

Virtual, October 10-21, 2020

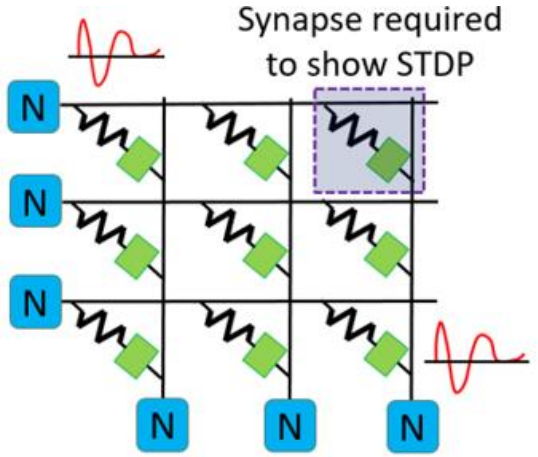
Contents

Introduction

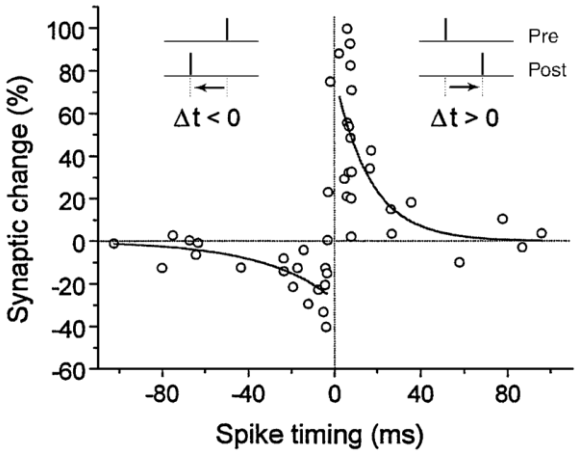
- Online learning in SNNs in crossbar synaptic array of NVM devices
- Pre-neurons drive current into the post neuron
- STDP characteristics map the spike-time difference to synaptic weight change
- Weight change in RRAMs requires programming voltage of appropriate magnitude
- STDP emulation requires mapping from spike time difference to programming voltage
- This is achieved using pulse shaping circuits

Pulse Shaping Circuits (PSCs) map the spike time difference to the programming voltage

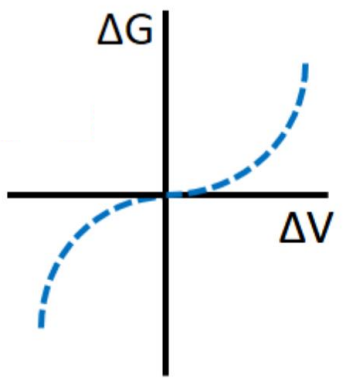
RRAM + Selector Crossbar Array [1]



Biological STDP [2]

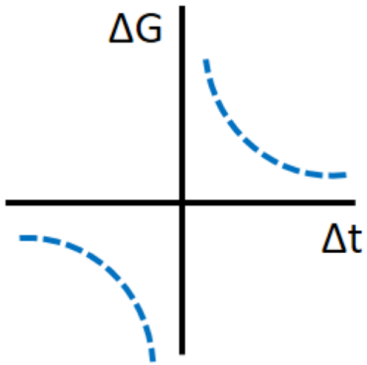


Device Programming Characteristics



ΔV = f(Δt)
PSC

STDP

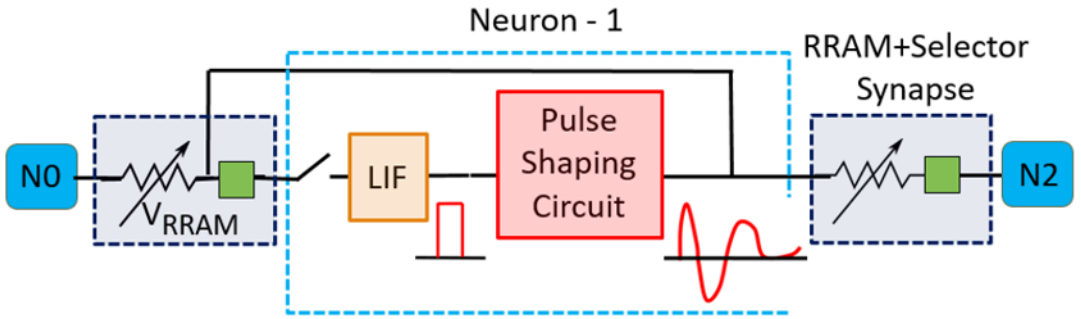


Pulse Shaping Circuits

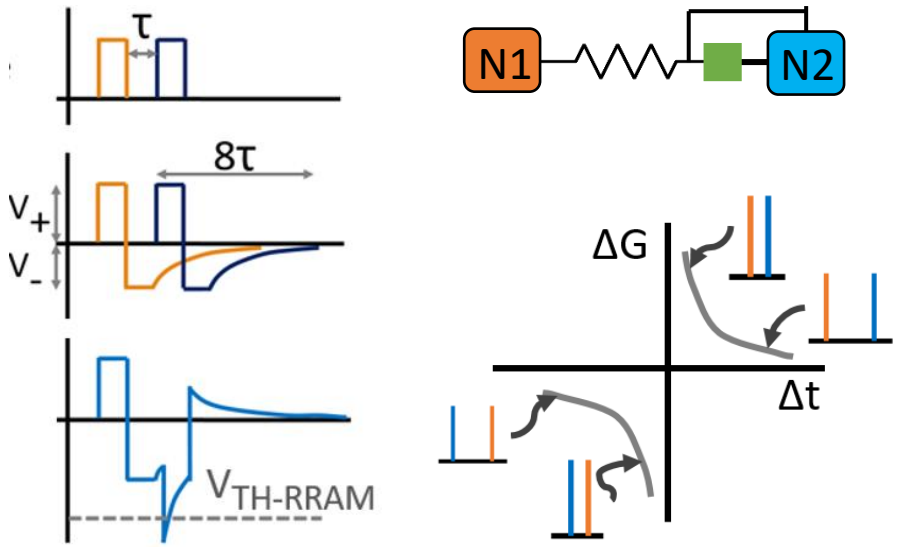
- PSCs are driven by the LIF block in the neuron
- The difference between temporally separated shaped voltage across the terminals causes spike-time dependent programming
- PSCs consume high fraction of circuit resources like energy and area
- Complexity of this circuits can negate the advantage gained by carefully designed LIF circuits [3]
- NIPIN selector device can eliminate the PSCs by doing the time-keeping (TK) using its hole storage mechanism [4]
- What is the quantitative resource consumption of PSCs and what is the advantage gained by TKD?

What is the quantitative advantage time-keeping device can provide?

Synapse-Neuron Connection



Pulse Shaping Example

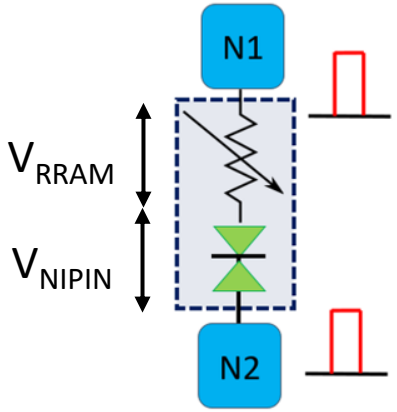


Contents

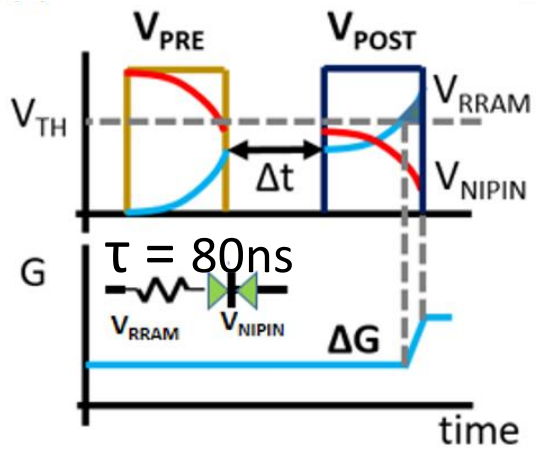
Time-Keeping Device - NIPIN Selector

- NIPIN time-keeping device produces STDP with square pulses
- The voltage across the pulse gets divided between NIPIN and RRAM
- Hole storage causes V_{NIPIN} to decrease with time
- Single pulse does not cause V_{RRAM} to exceed the programming threshold
- Closely space pulses cause RRAM programming
- This has been demonstrated experimentally
- A simulation model for circuit analysis is required

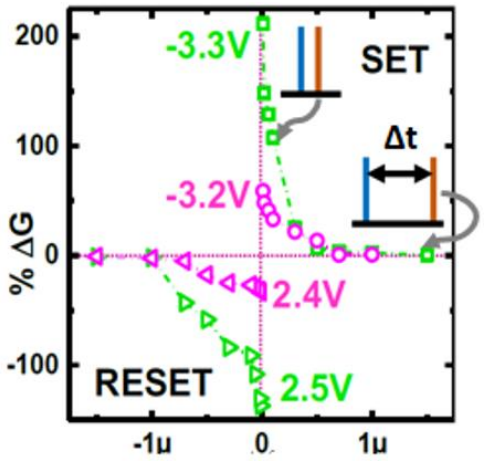
NIPIN device has hole storage physics that keeps time between pre and post spiking



Timing Diagram [5]



Experimental Characterization [4]



Device Modeling

Diffusion Current

$$I_D(t) = a_1 \times (exp^{-c_1\phi(t)} - exp^{-c_1\phi(0)})$$

Impact Ionization Current

$$I_{II}(t) = \frac{I_D(t)}{c_2} + c_3 \times (exp^{c_4 V_a} - 1)$$

Hole Current

$$I_{hl}(t) = c_5 \times (exp^{-c_6\phi(t)}) \times (\phi(0) - \phi(t))$$

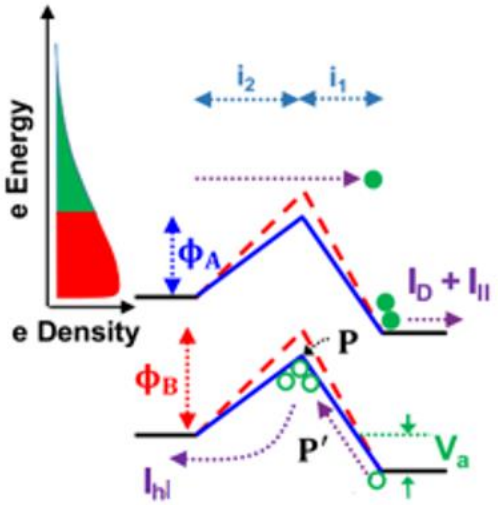
Charge Storage

$$\Delta Q_{stored}(t) = (I_{II}(t) - I_{hl}(t)) \times \Delta t$$

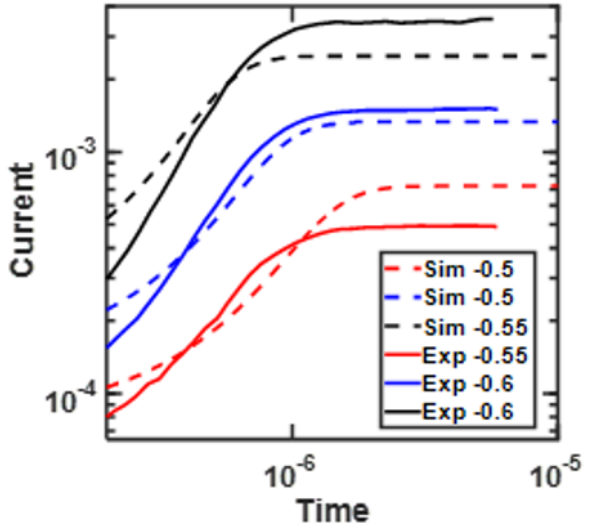
Voltage Build-up

$$\phi(t) = \phi(0) - \frac{Q_{stored}}{c_7 C}$$

Band Diagram Schematic [3]



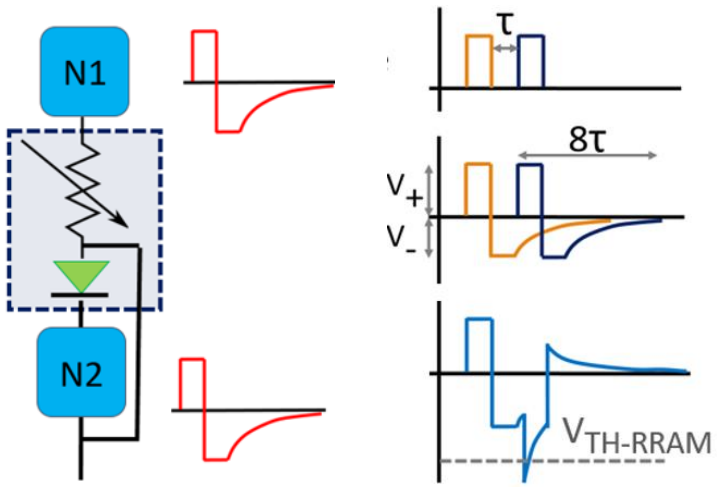
Transient Response [5]



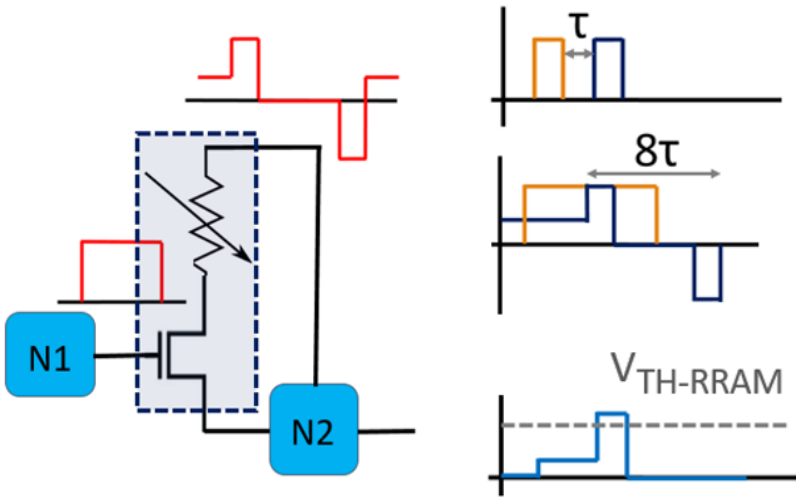
Simple Verilog-A model can describe the transient behaviour

Configurations

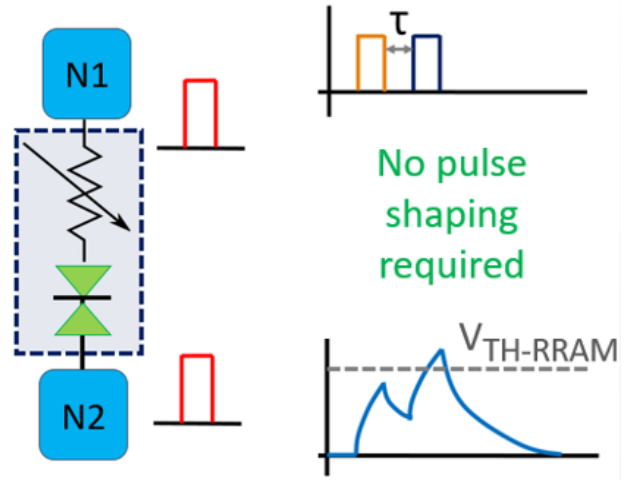
1D1R Synapse [3]



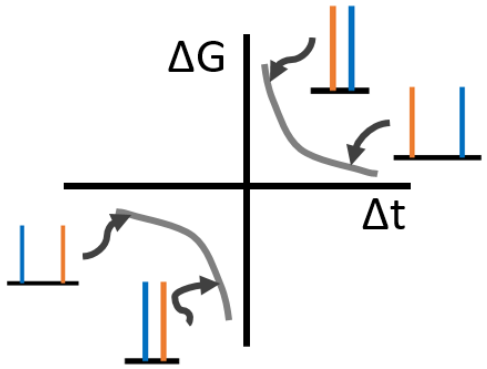
1T1R Synapse [6]



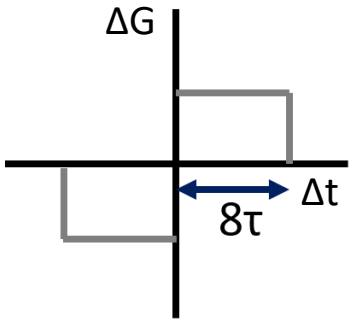
1S1R Synapse [4]



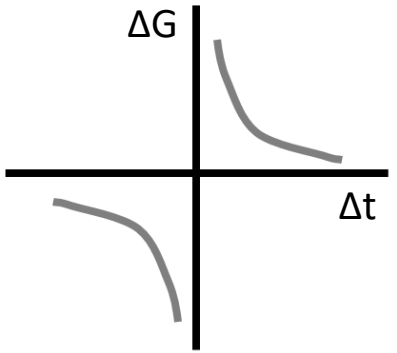
STDP



STDP

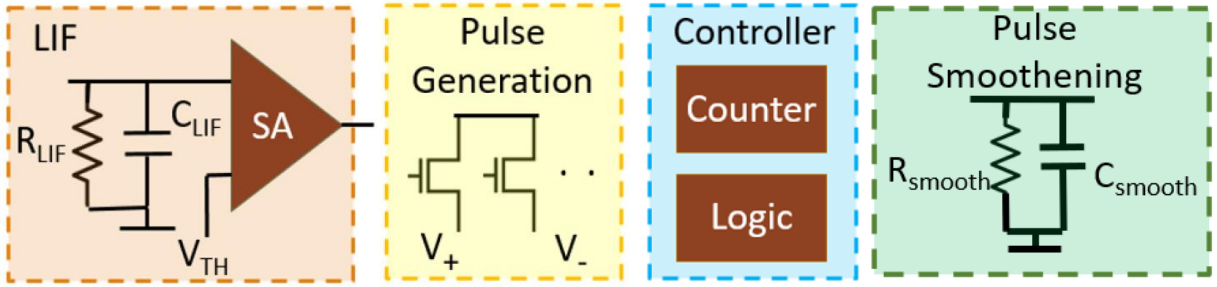


STDP



1D1R Configuration

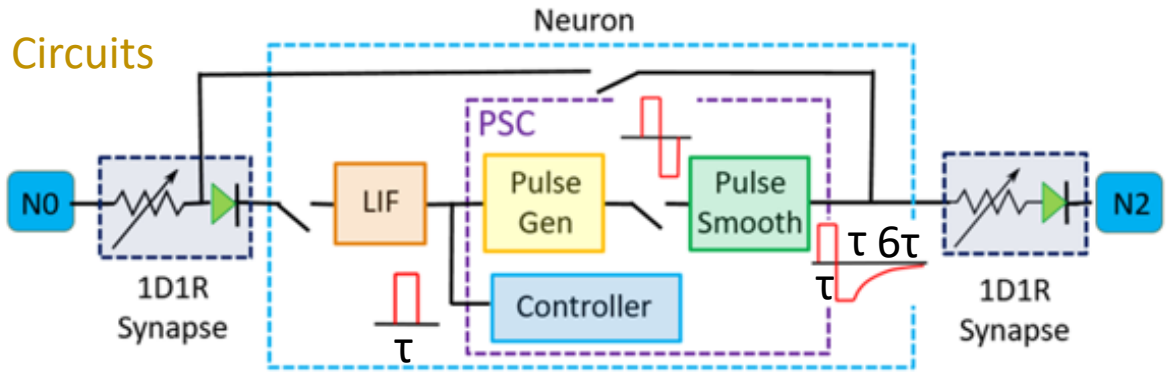
- Parallel RC circuit mimics LIF behaviour
- LIF spike activates the controller containing 3-bit counter
- Controller provides gate voltages to pulse generation transistors for pulse generation
- Pulse smoothening circuit passes the pulse through the RC filter
- The blocks are connected using CMOS switches



R_{LIF} = 1 MΩ
C_{LIF} = 1 fF

R_{smooth} = 100 kΩ
C_{smooth} = 10 fF

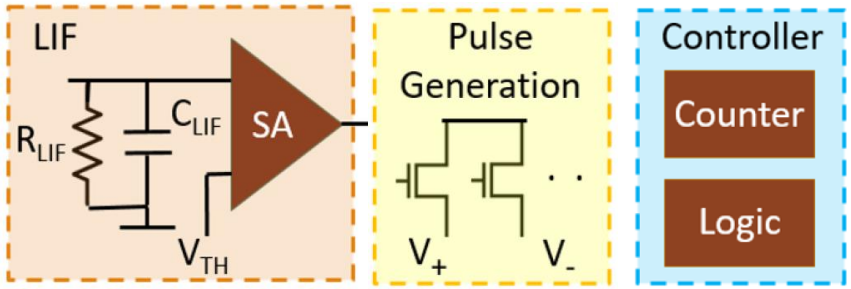
Circuits



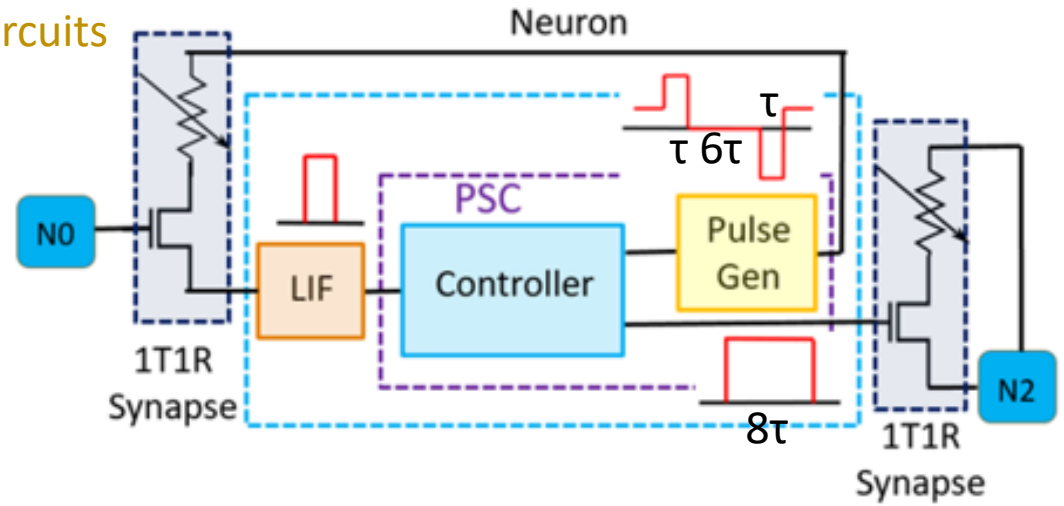
1D1R configuration requires multi-block shaping circuit

1T1R Configuration

- Similar LIF block drives the controller and pulse generation circuit
- Controllers generates 2 types of pulses driving the RRAM and selector transistor
- Pulse smoothening RC block is eliminated
- The circuit gets simplified at the cost of square STDP window



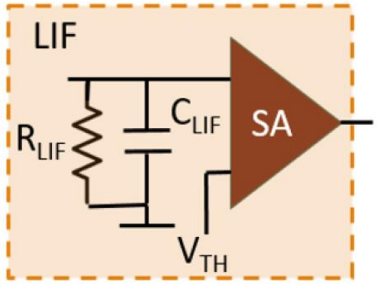
Circuits



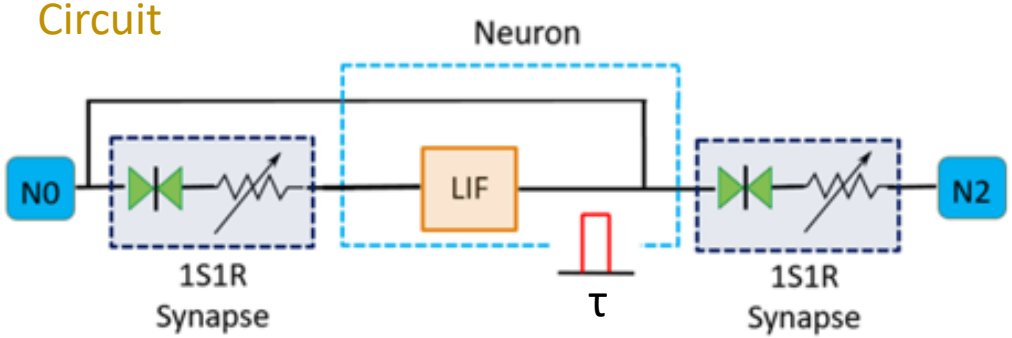
Circuit complexity gets slightly reduced compared to the 1D1R configuration

1S1R Configuration

- Contains only the LIF block
- Circuits are simulated using the shown NIPIN model, RRAM model [7] and 180 nm CMOS model [8]



Circuit



1S1R configuration eliminates the PSCs fully

Contents

Results

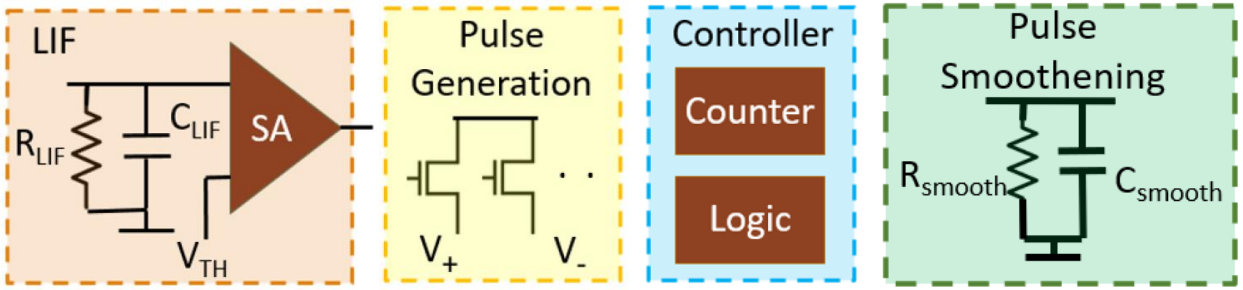
Energy per spike →

- Small R_{smooth} allows large current and high energy consumption
- All circuit blocks get eliminated in 1S1R scheme causing 133x energy per spike saving

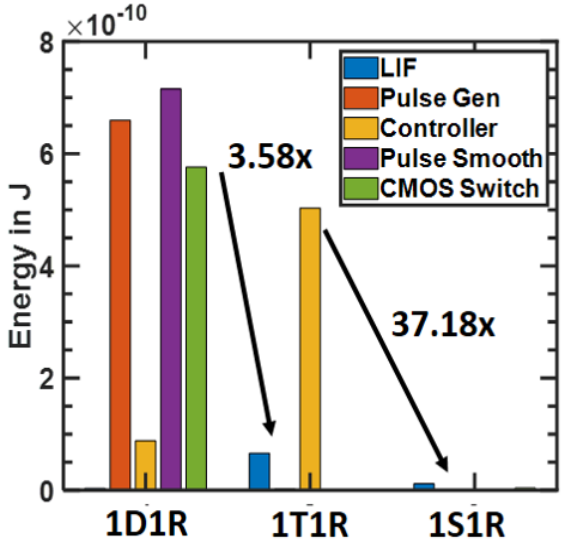
Area →

- Shaping RC requires high area with large capacitance requirement
- Controller area gets eliminated further from going to 1D1R from 1S1R with total 10x area saving

Time keeping selector saves significant circuit resources

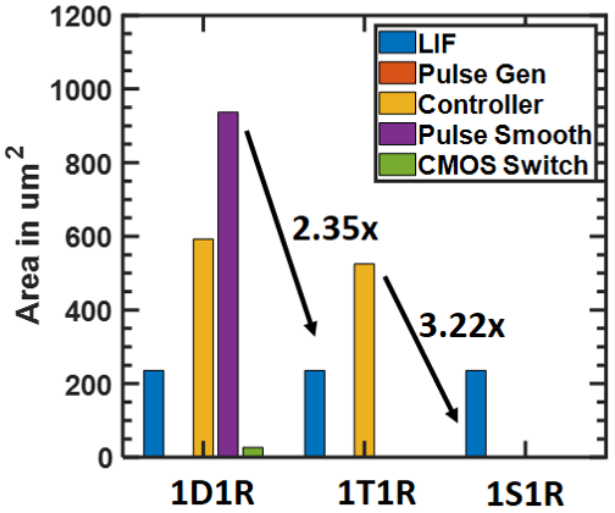


Energy Consumption



133x Reduction

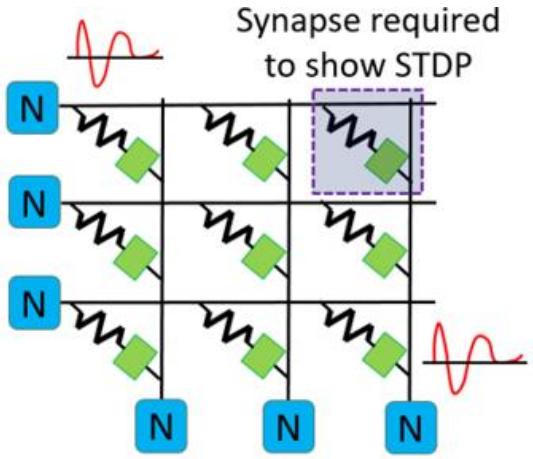
Area Consumption



10x Reduction

Discussion

- Multiple output resistances cause loading in high impedance 1D1R scheme requiring additional drivers
- Non-biological STDP may cause accuracy degradation in learning tasks
- With both positive and negative pulses 1T1R requires high V_{DD}
- Multiple voltage levels need to be generated on the chip for 1T1R scheme



Parameter	1D1R	1T1R	1S1R
Energy/spike	133×	37×	1×
Area	8×	3×	1×
Loading	High	Low	Low
STDP	Biological	Square	Biological
V_{DD}	2	3.6	2.6
Voltage Levels	3	4	2

1S1R configuration outperforms in multiple circuit requirements

Conclusion

- Implemented neuron circuits for different RRAM synapses reported in the literature
- Established the superiority of timekeeping device (NIPIN selector) based synapse in circuit cost reduction
- Showed 133x reduction in energy per spike and 8x reduction in area
- Provided a transient model for the device for future circuit level explorations in this device
- Highlighted the heavy overhead of pulse shaping circuits in neurons to encourage exploration in timekeeping devices as selectors

References

1. Prezioso, M., et al. "Spike-timing-dependent plasticity learning of coincidence detection with passively integrated memristive circuits." *Nature communications* 9.1 (2018): 1-8.
2. Bi, Guo-qiang, and Mu-ming Poo. "Synaptic modification by correlated activity: Hebb's postulate revisited." *Annual review of neuroscience* 24.1 (2001): 139-166.
3. Rajendran, Bipin, et al. "Specifications of nanoscale devices and circuits for neuromorphic computational systems." *IEEE Transactions on Electron Devices* 60.1 (2012): 246-253.
4. Das, Bhaskar, et al. "$Pr_xCa_{1-x}MnO_3$-Based Memory and Si Time-Keeping Selector for Area and Energy Efficient Synapse." *IEEE Electron Device Letters* 40.6 (2019): 850-853.
5. Das, Bhaskar, Jörg Schulze, and Udayan Ganguly. "Transient phenomena in sub-bandgap impact ionization in Si nipin diode." *IEEE Transactions on Electron Devices* 65.8 (2018): 3414-3420.
6. Pedretti, G., et al. "Memristive neural network for on-line learning and tracking with brain-inspired spike timing dependent plasticity." *Scientific reports* 7.1 (2017): 1-10.
7. Chen, Pai-Yu, and Shimeng Yu. "Compact modeling of RRAM devices and its applications in 1T1R and 1S1R array design." *IEEE Transactions on Electron Devices* 62.12 (2015): 4022-4028.
8. Zhao, Wei, and Yu Cao. "New generation of predictive technology model for sub-45 nm early design exploration." *IEEE Transactions on Electron Devices* 53.11 (2006): 2816-2823.

Thank You