Ionic Analog Synapses for Deep Learning Accelerators

Kevin Limanta, Ahmad Zubair, Tomás Palacios

Sponsorship: Advanced Micro Devices Undergraduate Research and Innovation Scholar

The recent progress in novel hardware/software co-optimizations has led to tremendous improvement of the efficiency of neural networks. Nevertheless, the energy efficiency is still orders of magnitude lower than biological counterpart – the brain. Digital CMOS architecture has inherent limitations for deep learning applications due to their volatile memory, spatially separated memory and computation, and the lack of connectivity between nodes. Crossbar arrays of non-volatile memory devices, able to perform simple operations (e.g. bit multiplication), can potentially achieve a 30000× improvement in energy efficiency. State-of-the-art analog "synaptic" devices based on resistive memories suffer from stochastic, asymmetric, and non-linear weight updates, detrimental to training accuracy. Electrochemical ionic devices have been shown to be fast, energy efficient, and exhibit symmetric, linear weight updates. However, electrolytes used for the electrochemical l reaction are often CMOS incompatible and suffers from scalability.

We propose a simpler transistor-based analog synapse, consisting of a proton-doped SiO_2 gate oxide which electrostatically modifies the threshold voltage of the semiconductor channel, tuning the channel conductance (Figure 1). Non-volatility is maintained by trapping of protons in the oxide. Due to electrostatics, we expect to observe a symmetric and linear shift in threshold voltage, leading to linear weight updates. We study the proton diffusion and electrostatic effects through device simulation via Silvaco Atlas and analytical modeling. Simulations show a threshold voltage shift of the MOS gate stack due to the presence of ions in the gate oxide (Figure 2). We fabricate n-Si/ALD SiO_2/Al MOS capacitor and to demonstrate the feasibility of our ionic device. We observe that the MOS gate stack exhibits hysteretic behavior below 2V, indicating non-volatility and low-voltage operation. The results of this work will shed light on the feasibility of simple CMOS-compatible ionic devices for the next generation of neural network hardware accelerators.

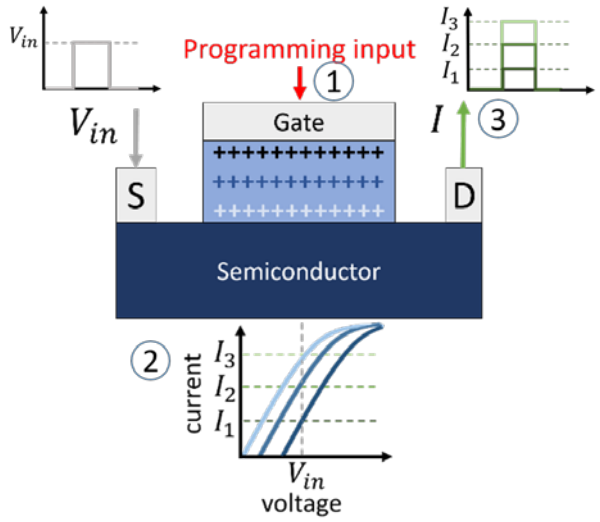

Figure 1: All inputs are based on simple square pulses. **1.** Programming input pulse accelerates protons toward semiconductor, changing device to a new state. **2.** Threshold voltage V_T shifts. **3.** For a given input voltage V_{in}, the conductance of the device (measured by output current) changes based on state of device.

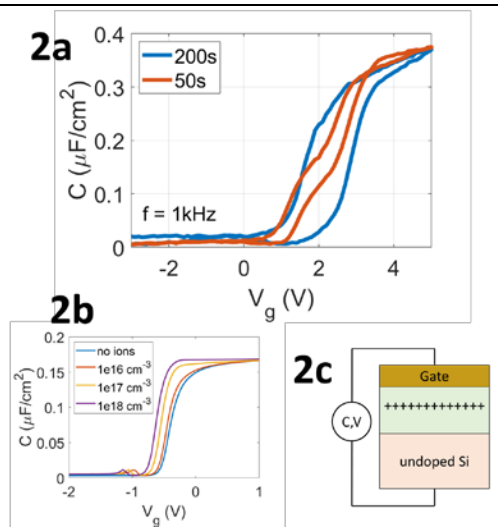

Figure 2: **a**. Capacitance-voltage measured on fabricated MOS gate stack. **b.** Device simulation of MOS gate stack doped with protons shows a threshold voltage shift. Effect is stronger with more ions implanted into device. **c.** Diagram of MOS structure.

Kevin Limanta
klimanta@mit.edu

Undergraduate supervised by Tomás Palacios
Starting PhD at MIT EECS in September 2020

Research Interests:
Electronics devices, nanotechnology

Profile: https://mtlresume.mit.edu/profile/kevin-limanta
PDF: https://mtlresume.mit.edu/system/files/resumes/limanta_resume.pdf