

Discussion Leader: Mark Yu

Scribe Notes: Sam Thomas

Paper: Review of Memristor devices in neuromorphic computing: materials science and device challenges

1. Introduction

- a. Purpose: lower power consumption, increase efficiency, increase speed and density
- b. Doped and undoped oscillator regions with titanium oxide to conduct (doped side) and insulate (undoped side)
- c. Voltage grows the doped region, which also grows the resistance
- d. (Iris) How does the device keep memory? What is happening when voltage is applied? => the state is determined by larger voltage (on write), reading done by providing a small voltage and reading the different resistances depending on what was stored
- e. Physical properties
 - i. Supports both abrupt (binary) and gradual (analog) switching as a result of resistive switching => gradual state for neuromorphic computing application
 - ii. Multiple analog states allow for in-memory computing states
 - iii. High endurance and long retention => this seems to be contrary to what we have previously discussed as it relates to DRAM
 - iv. (Iris) Are there physical limitations of memristors to keep them from scaling to 10nm (like CMOS or silicon technologies)? => smaller scales would cause the reliability of the memristors to be much worse

2. Material system engineering

- a. Oxides can be an issue, certain oxides are incompatible with CMOS devices
- b. Heating effects are a concern (Iris) which might kill the density arguments => more density is better, but if the heat density is greater then you don't win regardless; (Jiwon) What does I-V nonlinearity refer to? => voltage can change resistance, comes up later :-)
- c. Phase change memristors: SET switching speed is slow for PCM, very high programming current for RESET, reliability is difficult

3. Challenges of memristor dielectric and electrode

- a. Stochasticity: variations from device to device in working voltages (bad), can be used for hardware security applications (good) => (Iris) You may write to produce certain resistance, but there's a change in that resistance over time? It has some nice properties when applied to a random number generator, but variability within a range => (Semanti) Does randomness hinder large-scale production and standardization? Is randomness bad?
- b. CMOS compatibility: Pt and Pd are not compatible, so TiN and TaN are alternatives
- c. Integration: parasitic wire resistance (device is compact, so extra effective capacitance), in addition to heat density issue

- i. Solution: 3D stacking to dissipate density
- 4. Performance expectations
- 5. Integration challenges and strategies
- 6. Application to neuromorphic computing
 - a. Networks trained by *in situ* (adjusted parallelly in the memristive crossbar) (Iris: weights calculated in a GPU and then compared to the memristor values) or *ex situ* (synaptic weights calculated by software) methods
 - i. CPU, GPU, MPU (Iris) All processing is done in-memory in MPU architecture, there is no external memory needed?
 - b. Units of memristors exist in the crossbars: where “synapse weights” live and can be computed in-memory
 - i. Weights adjusted by changing pulse and frequency of the voltage into the crossbars, which will cause different resistances firing from the crossbars => (Iris) due to randomness, the memristors take advantage of which resistor will fire first; the strength of the pulse can overpower the randomness from the stochasticity (Jiwon, Karpur, Semanti) How would you exploit randomness for neural networks? (Iris) Noise/randomness might make network more robust during training, but after that it's not good - in fact bad (Karpur) the more randomness can be realized in a state would lead to an energy barrier itself (Iris) if there's enough confirmation/signals coming in, then the noise will be overruled potentially? (Semanti) Not even all parts of training can be random, only drop out, etc... but not convolution or computation (Qishen) True randomness would help with the overall security of the system, but the paper instead focuses on neuromorphic computing where application of randomness is much more tenuous (Semanti) If encrypted files are sent across devices and other devices have equally variant systems, wouldn't that mean that files couldn't be decrypted?
 - ii. Analog histories could theoretically be used to make predictions, weight of resistance is important