SAND2014-18438C

Development, Characterization, and Modeling of a TaOx ReRAM for a Neuromorphic Accelerator

Electrochemical Society Fall Meeting

Cancun

October 9, 2014

Matthew J. Marinella, Patrick R. Mickel, Andrew J. Lohn, David R. Hughart, Robert Bondi, Denis Mamaluy, Harry Hjalmarson, James E. Stevens, Seth Decker, Roger Apodaca, Brian Evans, J. Bradley Aimone, Fredrick Rothganger, Conrad James, and Erik P. DeBenedictis

*matthew.marinella@sandia.gov

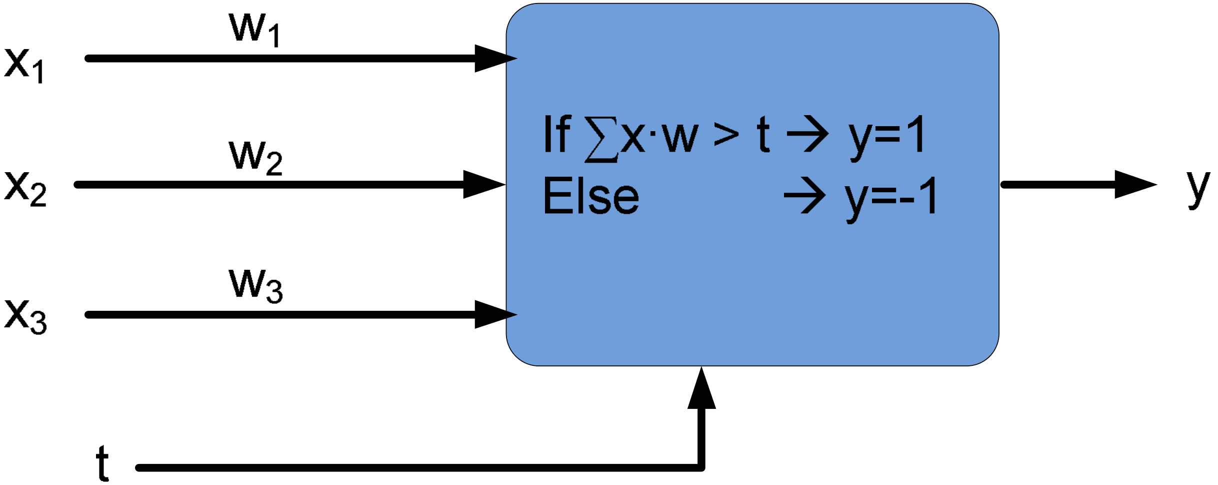
Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Outline

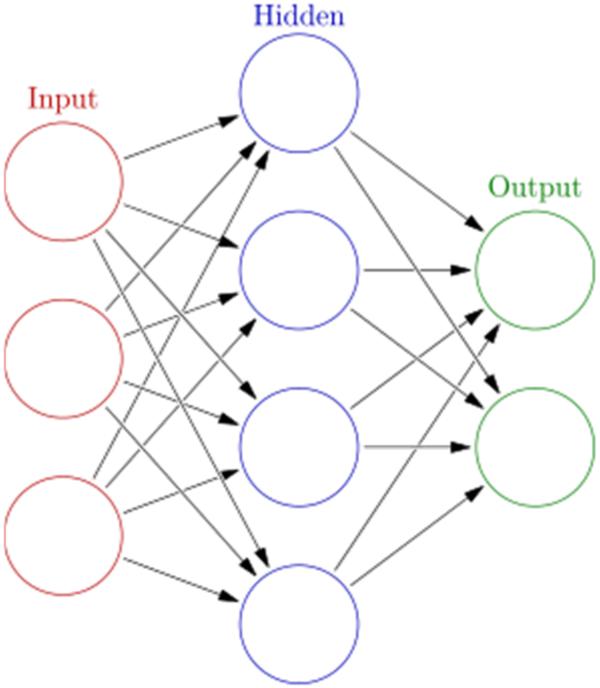
- Neuromorphic Computing with ReRAM
- Development of a CMOS/ReRAM Process
- Characterization of Device Behavior
- Modeling
  - Analytical
  - DFT
- Conclusion and Future Work

Perceptron Based Neural Nets

- Perceptron is the simplest model of a neuron
- Feed forward network
- Single node can learn linearly separable logic functions
- Applications:
  - Field programmable gate arrays
  - Pattern recognition

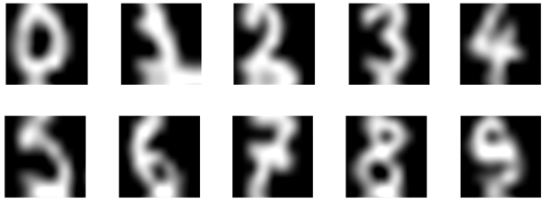


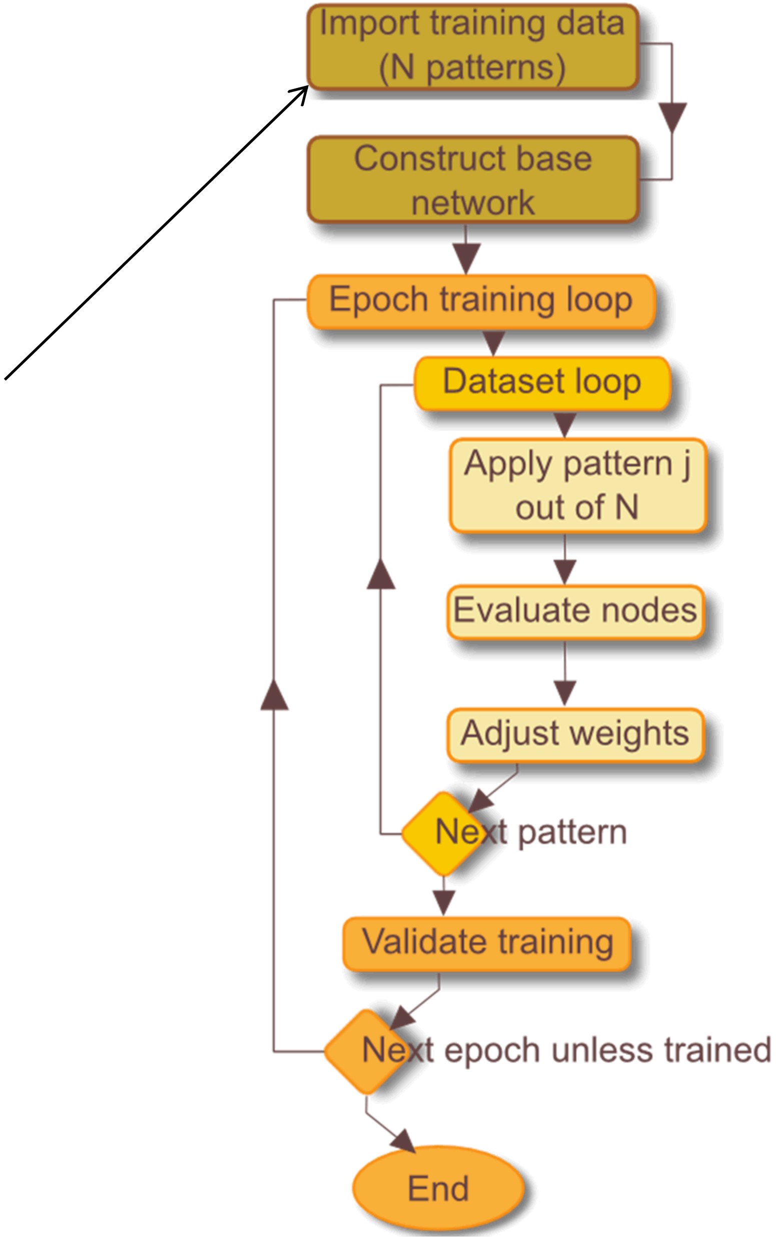
If $\sum x \cdot w > t$ → $y=1$
Else → $y=-1$



Wikipedia.org

Perceptron Network (Simplified)





Import training data (N patterns)

Construct base network

Epoch training loop

Dataset loop

Apply pattern j out of N

Evaluate nodes

Adjust weights

Next pattern

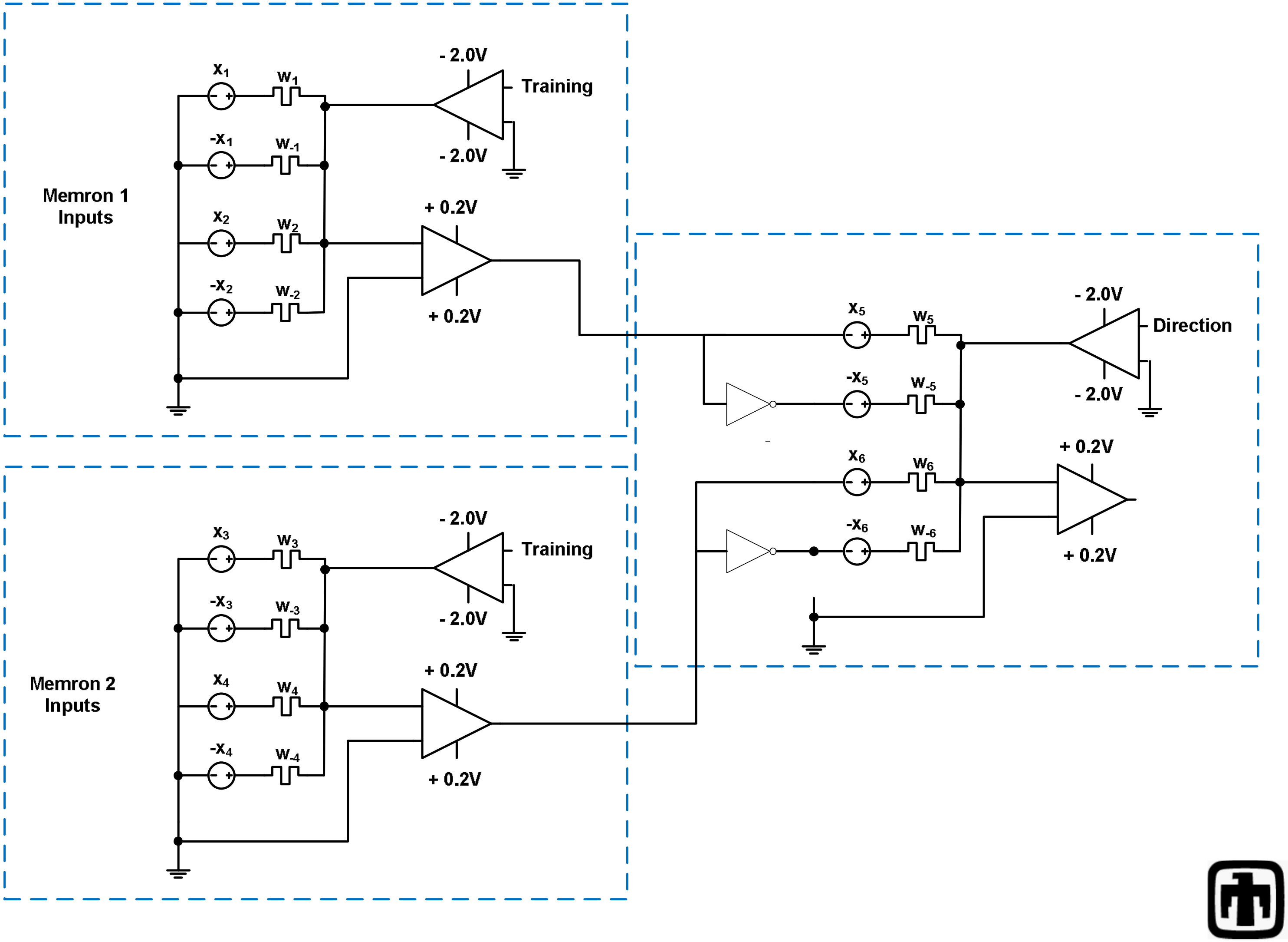
Validate training

Next epoch unless trained

End

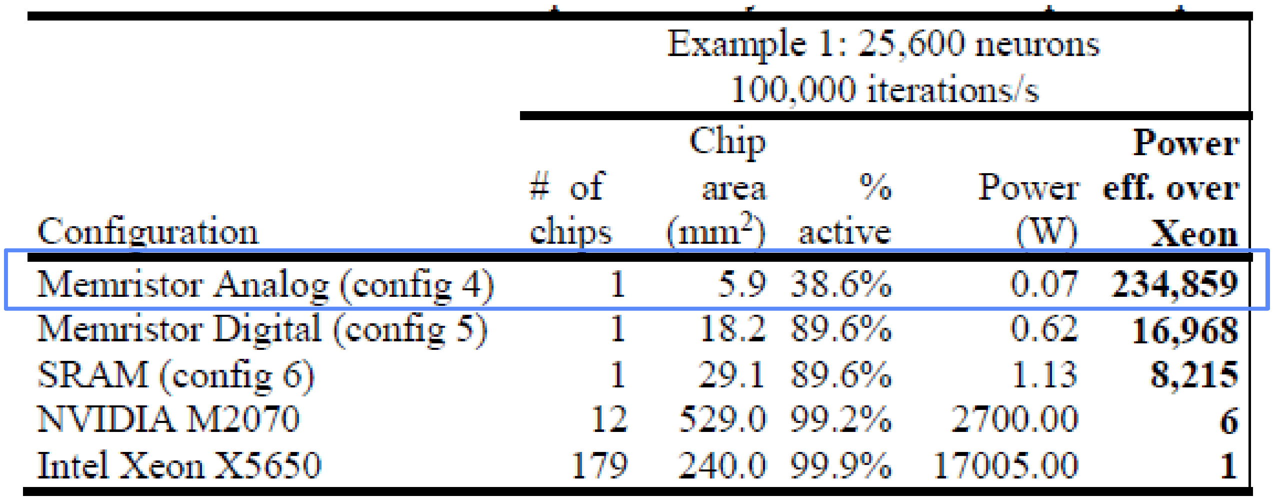
Memron

- **Requires 3 Memrons to learn XOR:**



Why Do We Need an HW Accelerator?

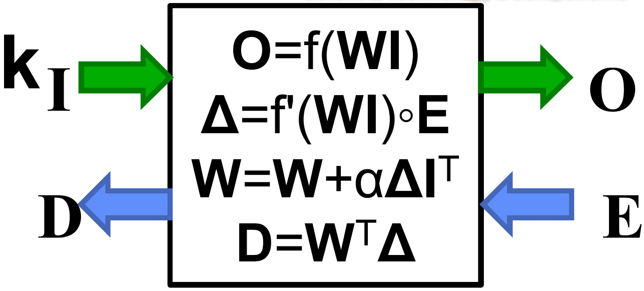
- Use simulation results for similar algorithm as example
- Significant power savings using a memristor-based HW accelerator :
- 16x reduction in power over SRAM ASIC
- 6x reduction in chip area over SRAM ASIC
  - Equivalent to 6x improvement in performance/area



	Example 1: 25,600 neurons 100,000 iterations/s				
Configuration	# of chips	Chip area (mm^2)	% active	Power (W)	**Power eff. over Xeon**
Memristor Analog (config 4)	1	5.9	38.6%	0.07	**234,859**
Memristor Digital (config 5)	1	18.2	89.6%	0.62	**16,968**
SRAM (config 6)	1	29.1	89.6%	1.13	**8,215**
NVIDIA M2070	12	529.0	99.2%	2700.00	**6**
Intel Xeon X5650	179	240.0	99.9%	17005.00	**1**

T. Taha, R. Hasan, C. Yakopic, M. McLean, in Proc. IEEE Intl. Joint Conf. on Neural Networks, 2013.

Perceptron Network



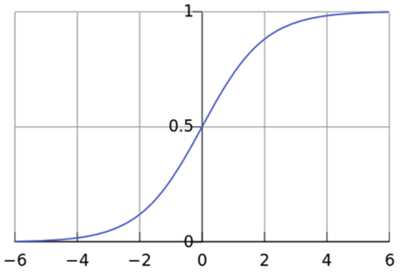
$O=f(\mathbf{WI})$

$\Delta=f'(\mathbf{WI})\circ\mathbf{E}$

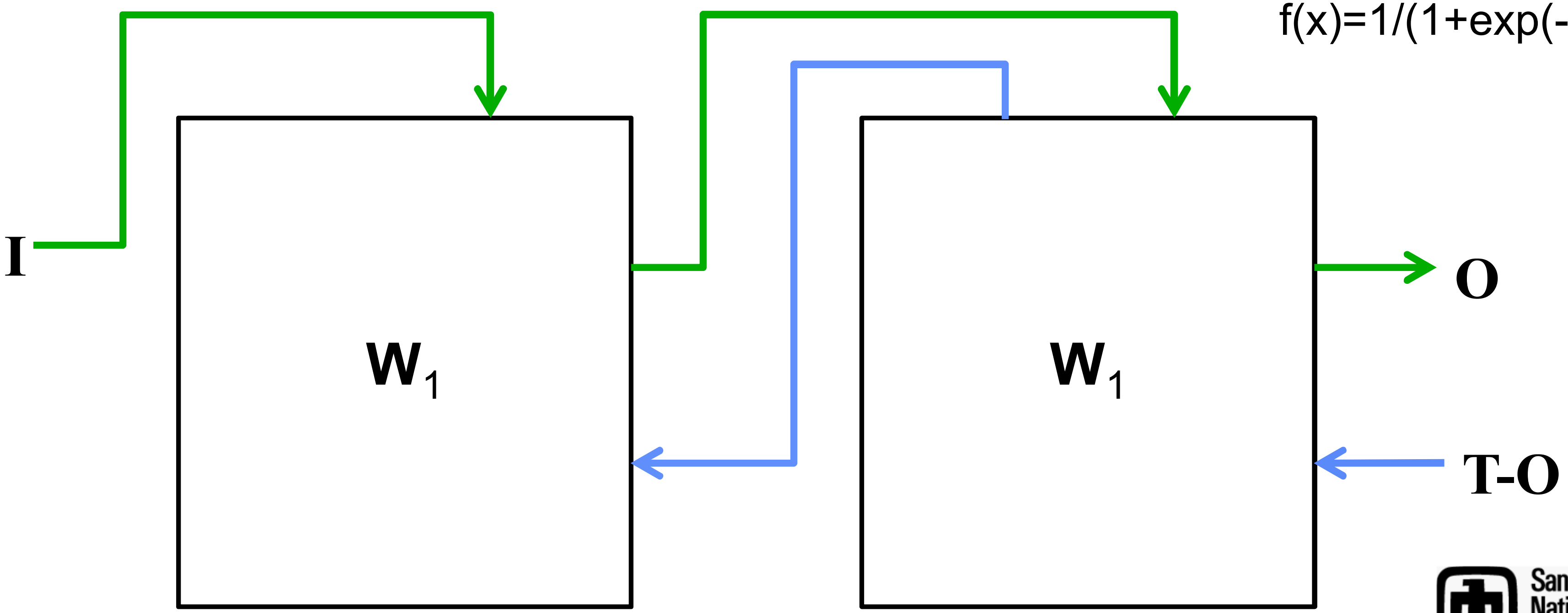
$\mathbf{W}=\mathbf{W}+\alpha\Delta\mathbf{I}^T$

$\mathbf{D}=\mathbf{W}^T\Delta$

- Compute output **O** based on input **I**, weight matrix **W**, and squashing function f().
- Delta-rule: Update **W** based on the outer product of **I** and the partial derivative **Δ**, scaled by learning rate α.
- Back-propagation: Allocate error on outputs (**E**) to error on inputs (**D**) by passing partial derivatives back through **W**.
- Training: Feed a bunch of (**I**,**T**) pairs, until convergence.

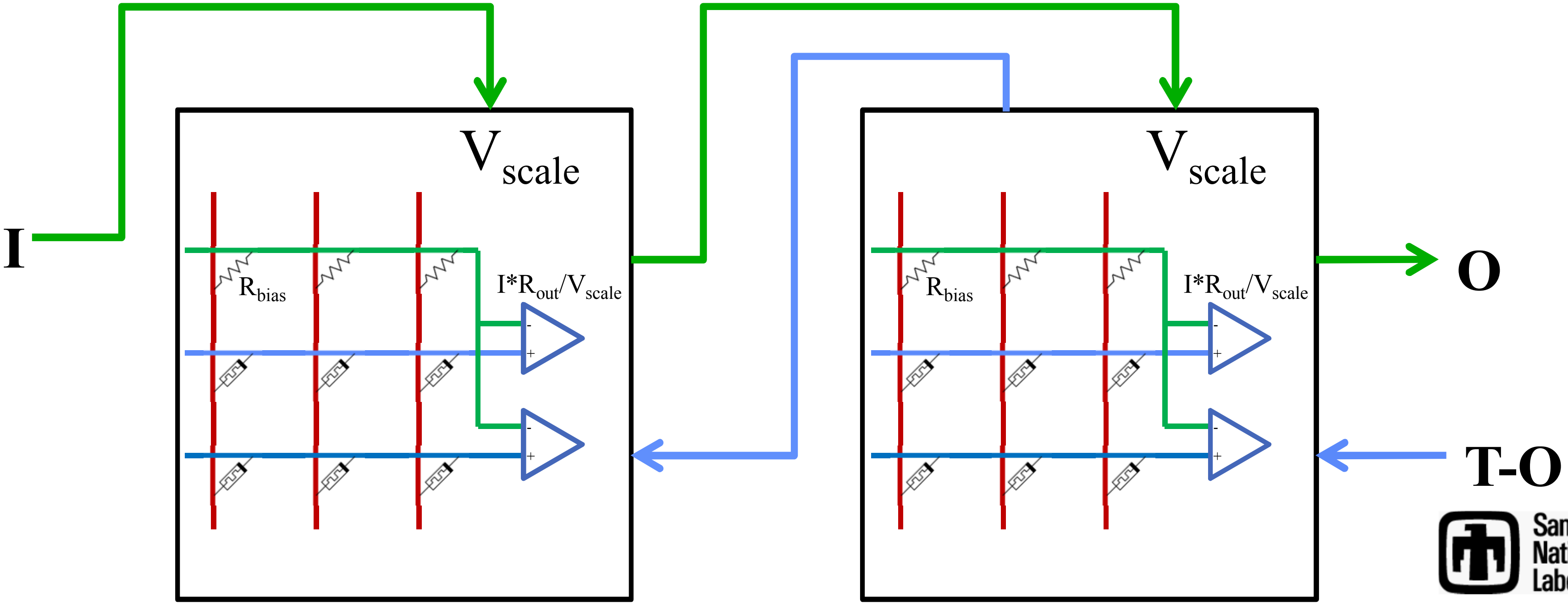


f(x)=1/(1+exp(-x))



ReRAM-Neural Hardware Simulation

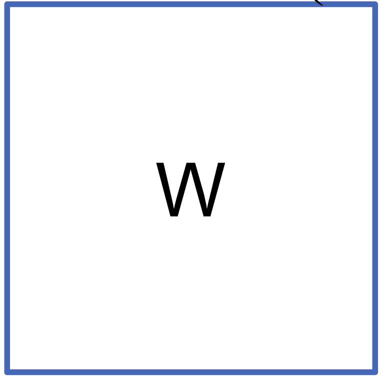
- Purely in software
- Uses data table to emulate each memristor
- Replaces MLP block with a crossbar and some hypothetical surrounding circuitry.
- Values passed between blocks in floating point.
- Converted internally to voltages via V_{scale} (at present simply 1).
- Weights mapped to numbers via R_{bias}=4,000 and R_{out}=40,000.
  - $R=1/(W/R_{out}+1/R_{bias})$



Performance of various network structures

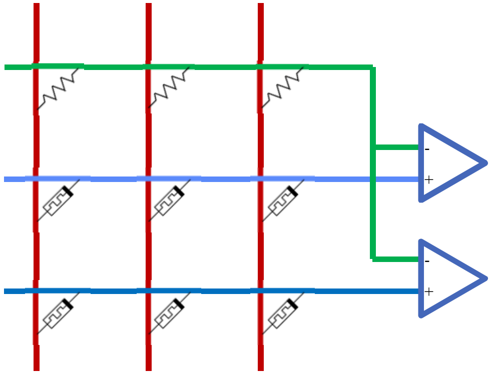
- **Train conventional perceptron.**
- **Convert weight matrix directly to memristor values.**
- **Test on MNIST handwriting dataset, with and without model noise.**

Conventional (CPU)



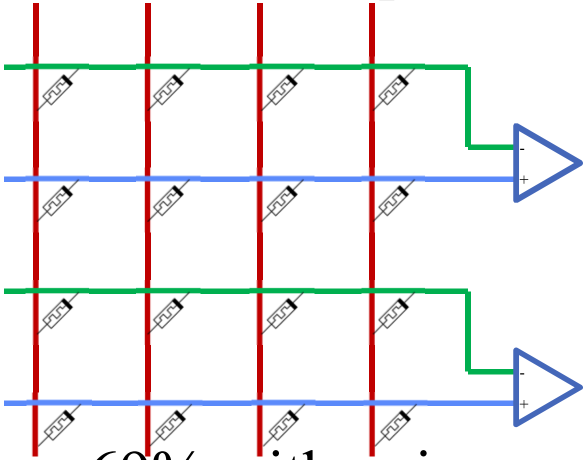
96%

Fixed bias resistors



88% with noise
96% noise-free

Bias memristor per element

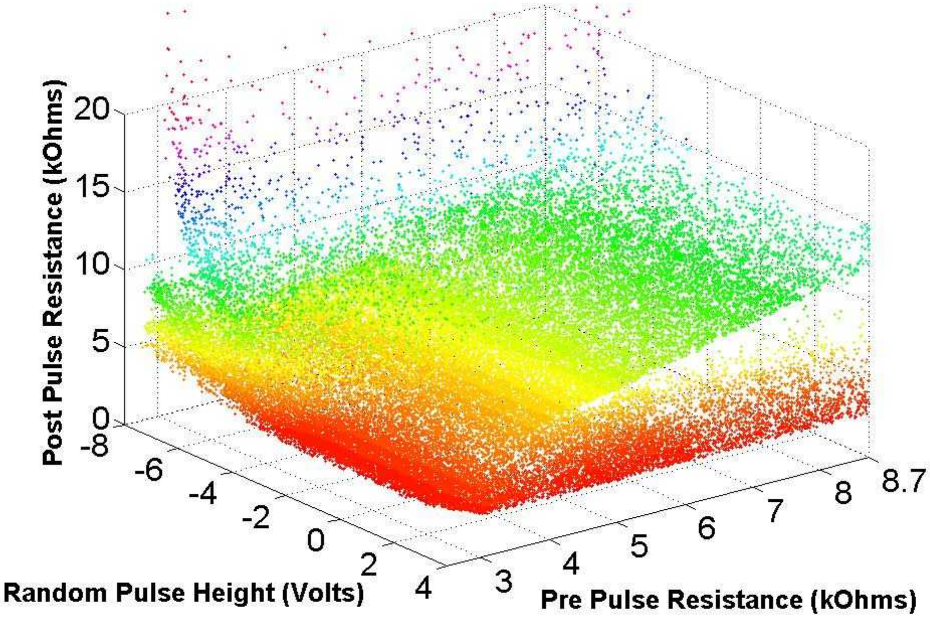


69% with noise
96% noise-free

Why? – Each memristor introduces more noise, so fixed-bias network is more accurate.

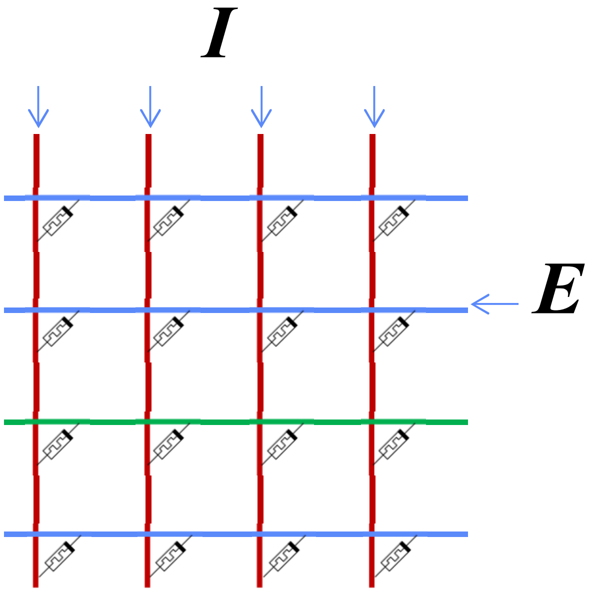
Empirical memristor model

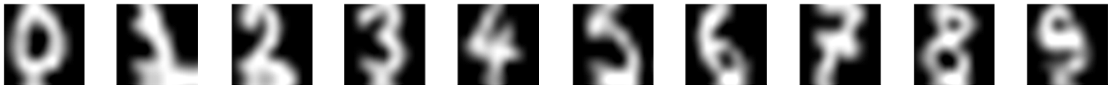
- Represented by a lookup table
- Binned 10M-sample data (100 Ω x 0.1 V bins)
- Table contains ΔΩ values
- Standard deviation of each cell used to simulate noise.
- Chart on left is slice through 4KΩ
- Flat between -1.6V and 1.6V



Empirical memristor model

- Backpropagation training
- $dW = \alpha E I^T$ where E is error vector and I is input
- Update each row separately so column values can be set appropriately for given error value.
- Test on handwritten digit recognition (MNIST dataset).
- Accuracy = 91.8% (classic methods reach 99.9%)





Representing Weight As Resistance

Relationship of weight to resistance: $W = R_{out}\left(\frac{1}{R} - \frac{1}{R_{bias}}\right)$

Sensitivity of weight to resistance (with Ra=Rb+R'):

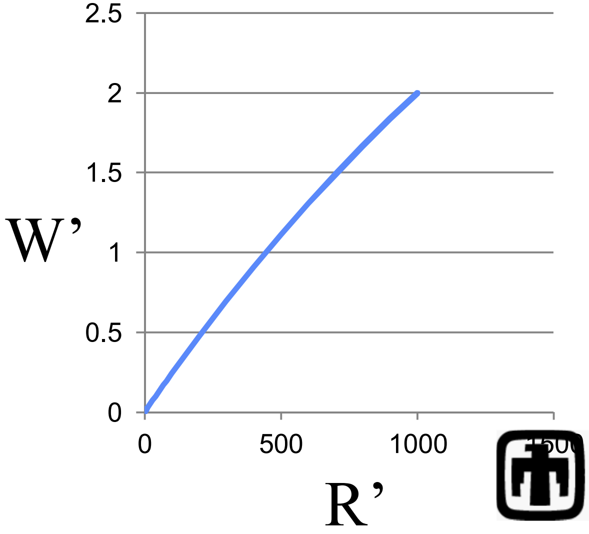
- $W' = W_a - W_b = R_{out}\left(\frac{1}{R_a} - \frac{1}{R_{bias}}\right) - R_{out}\left(\frac{1}{R_b} - \frac{1}{R_{bias}}\right)$

$$= R_{out}\left(\frac{1}{R_b + R'} - \frac{1}{R_b}\right)$$

$$= R_{out}\left(\frac{-R'}{R_b(R_b + R')}\right)$$

Example:

R_{bias}	= 4kΩ
R_{out}	= 40kΩ
R_b	= R_{bias}
R'	= 100Ω
W'	= 0.244



Sensitivity of Error to Resistance:

- Let WI be close to 0, so $f'(WI)=0.25$ (derivative of sigmoid function)
- Let $I=1$ (for simplicity)
- $\alpha=0.01$
- $W'=\alpha\Delta I=0.01*0.25*E*1$
- $E=W'/0.0025=0.244/0.0025=97.6$ (with $R'=100\Omega$ as on previous slide)

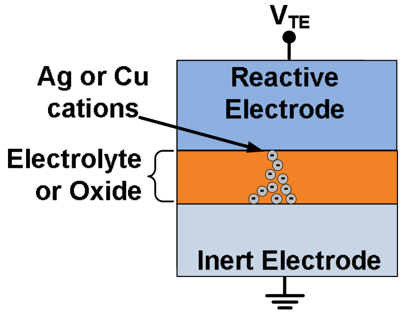
- That is, at typical values, the smallest error achievable is 2 orders of magnitude larger than typical output range for an MLP neural network!

Outline

- Neuromorphic Computing with ReRAM
- Development of a CMOS/ReRAM Process
- Characterization of Device Behavior
- Modeling the Source of Intradevice Variation
- Conclusion and Future Work

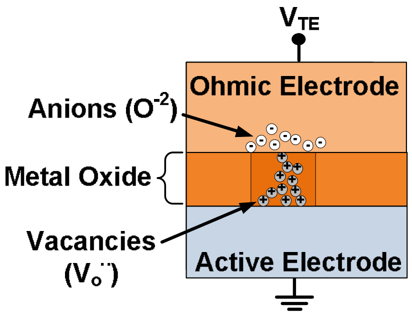
Resistive Random Access Memory

Electrochemical Metallization Bridge



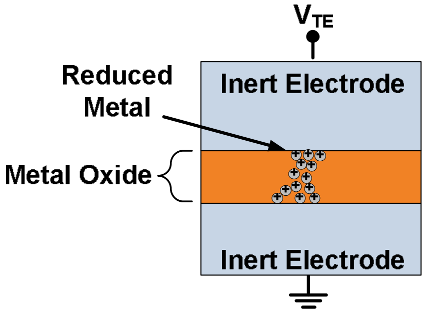
- Switching: Electrochemical formation and dissolution of Ag or Cu filament
- Cation motion (Ag or Cu)
- Chalcogenide or oxide insulating layer
- Switching depends on E-field direction
- R/W current independent of device area

Metal Oxide: Bipolar Filamentary



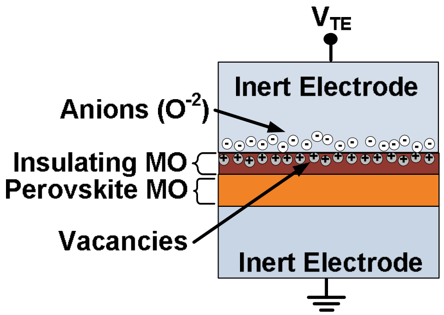
- Switching: Valence change and migration of oxygen vacancies
- Anion motion (O^{2-})
- HfO_x, TaO_x most common insulators
- Switching depends on E-field direction
- R/W current independent of device area

Metal Oxide: Unipolar Filamentary



- Switching: Thermochemical change in oxide valence state
- Anion motion (O^{2-})
- Symmetric structure
- NiO_x most common material
- Switching independent of E-field direction
- R/W current independent of device area

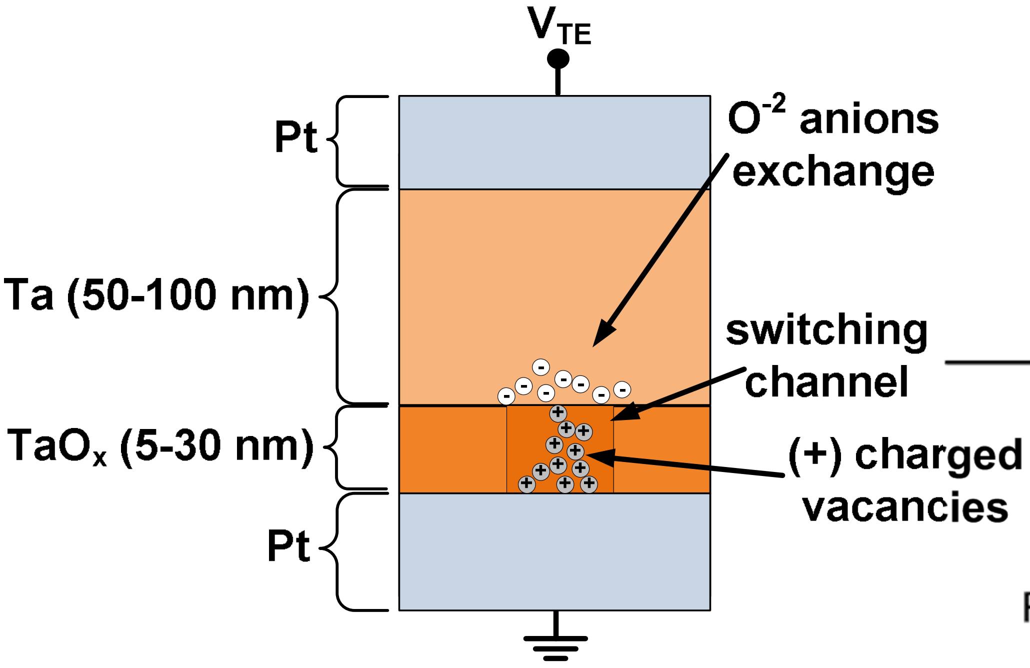
Metal Oxide: Bipolar Non-Filamentary

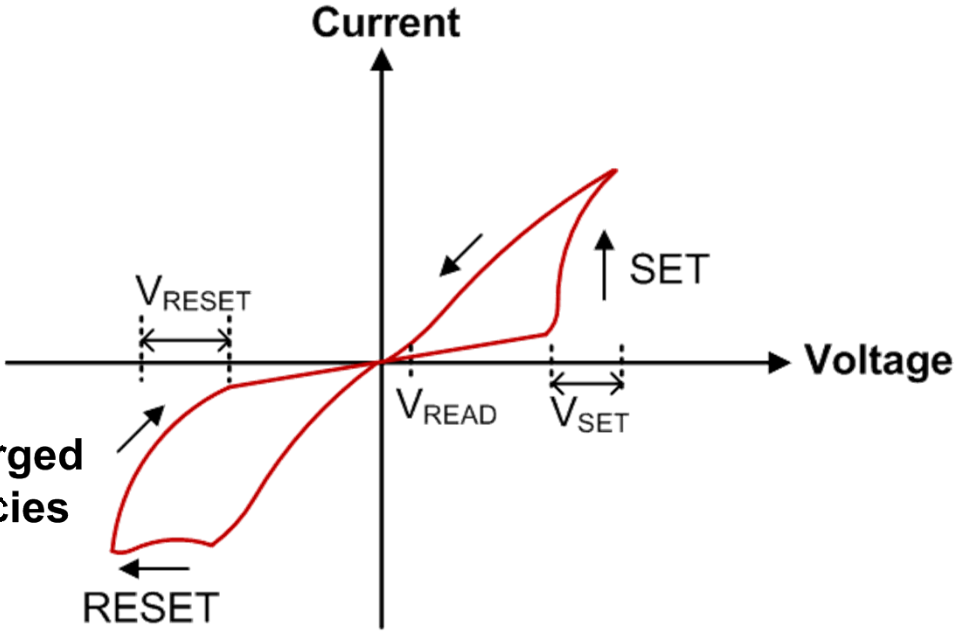


- Switching: Oxygen exchange causes Schottky barrier height at interface
- Anion motion (O^{2-})
- Perovskite and insulating metal oxide
- Switching depends on E-field direction
- R/W currents depend on device area

Valence Change ReRAM

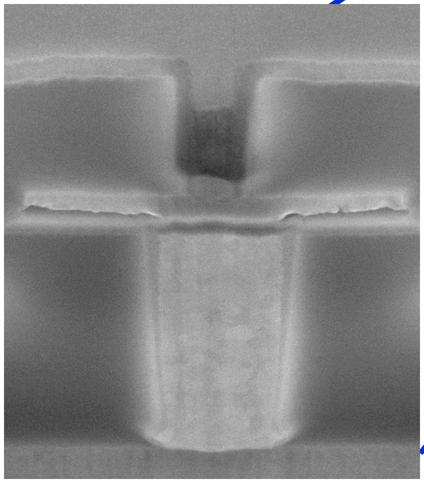
- "Hysteresis loop" is simple method to visualize operation
  - (real operation through positive and negative pulses)
- Resistance Change Effect (polarities depend on device):
  - Positive voltage/electric field: low R – O^{-2} anions leave oxide
  - Negative voltage/electric field: high R – O^{-2} anions return
- Common switching materials: TaO_x, HfO_x, TiO_2, ZnO

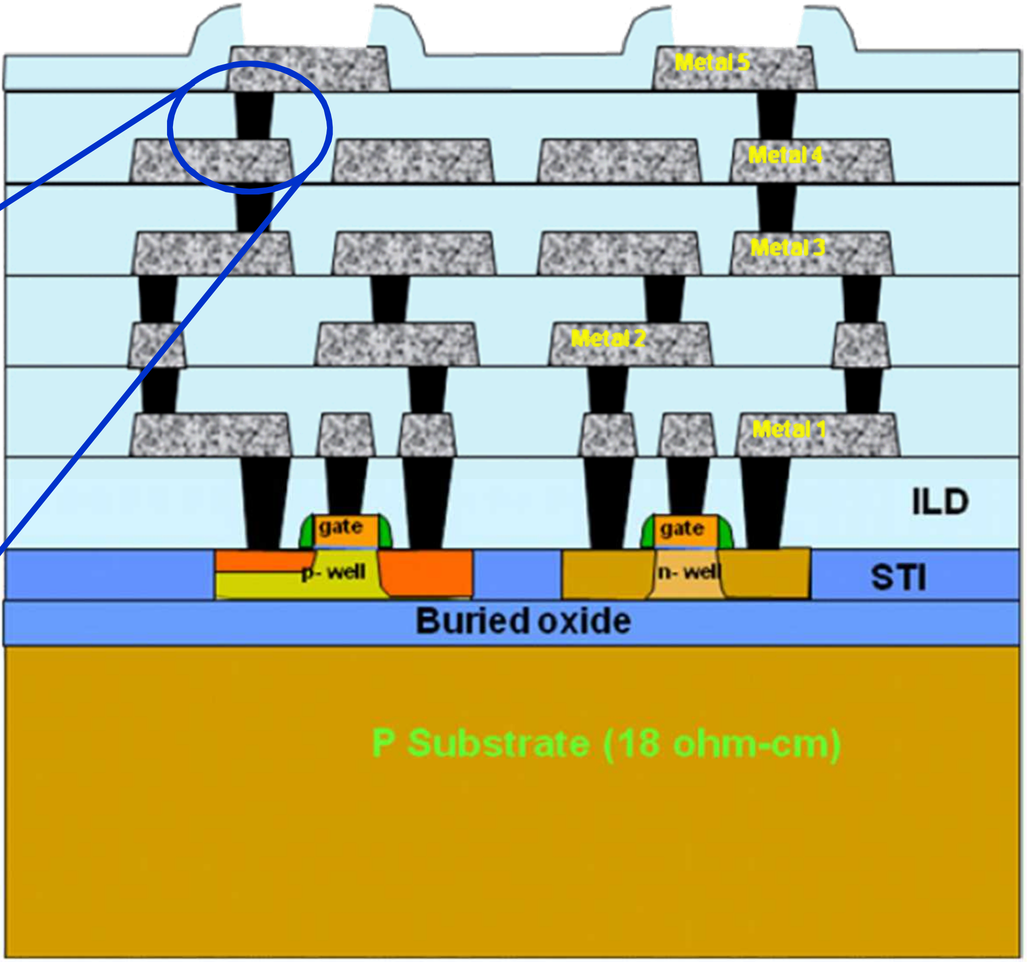




Memristors + CMOS

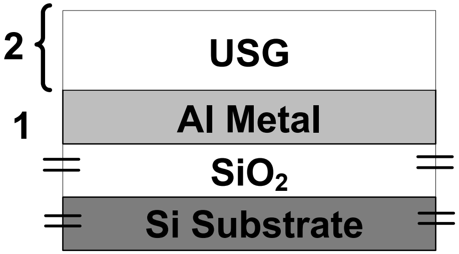
- Sandia CMOS7 Process
  - 3.3V, 350 nm, MOSFETs
  - SOI substrate
- Baseline for memristor integration



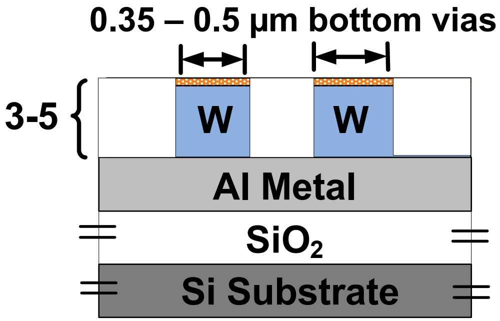


Process Flow

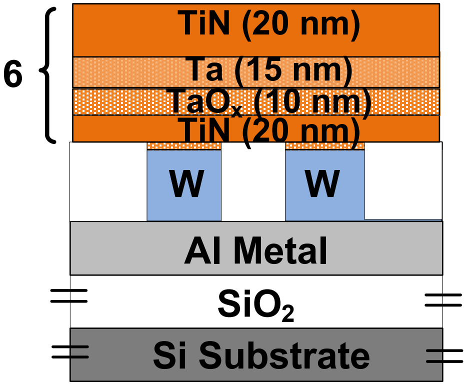
1. Deposit Bottom Metal (Al)
2. Deposit USG



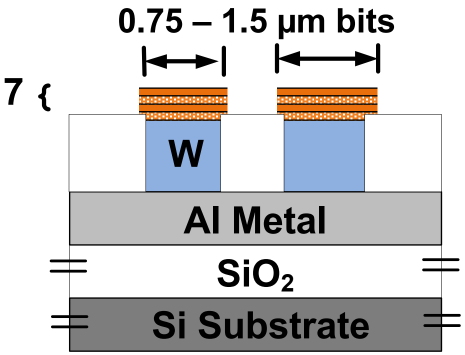
3. Etch via holes in USG
4. Deposit W and TiN layers
5. CMP



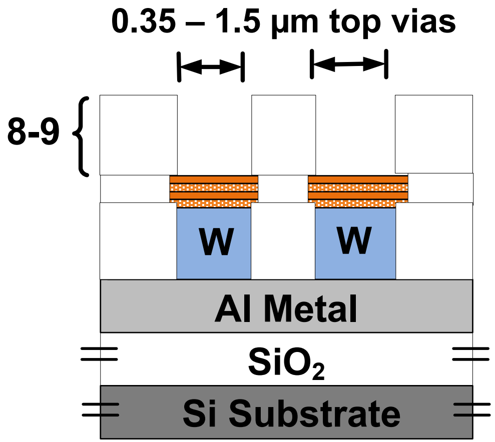
6. Deposit bit stack
(layers enlarged for clarity)



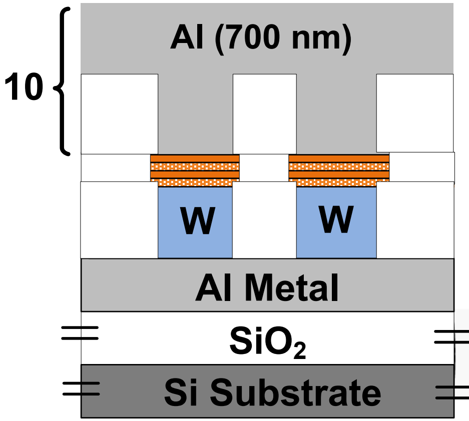
7. Etch bits



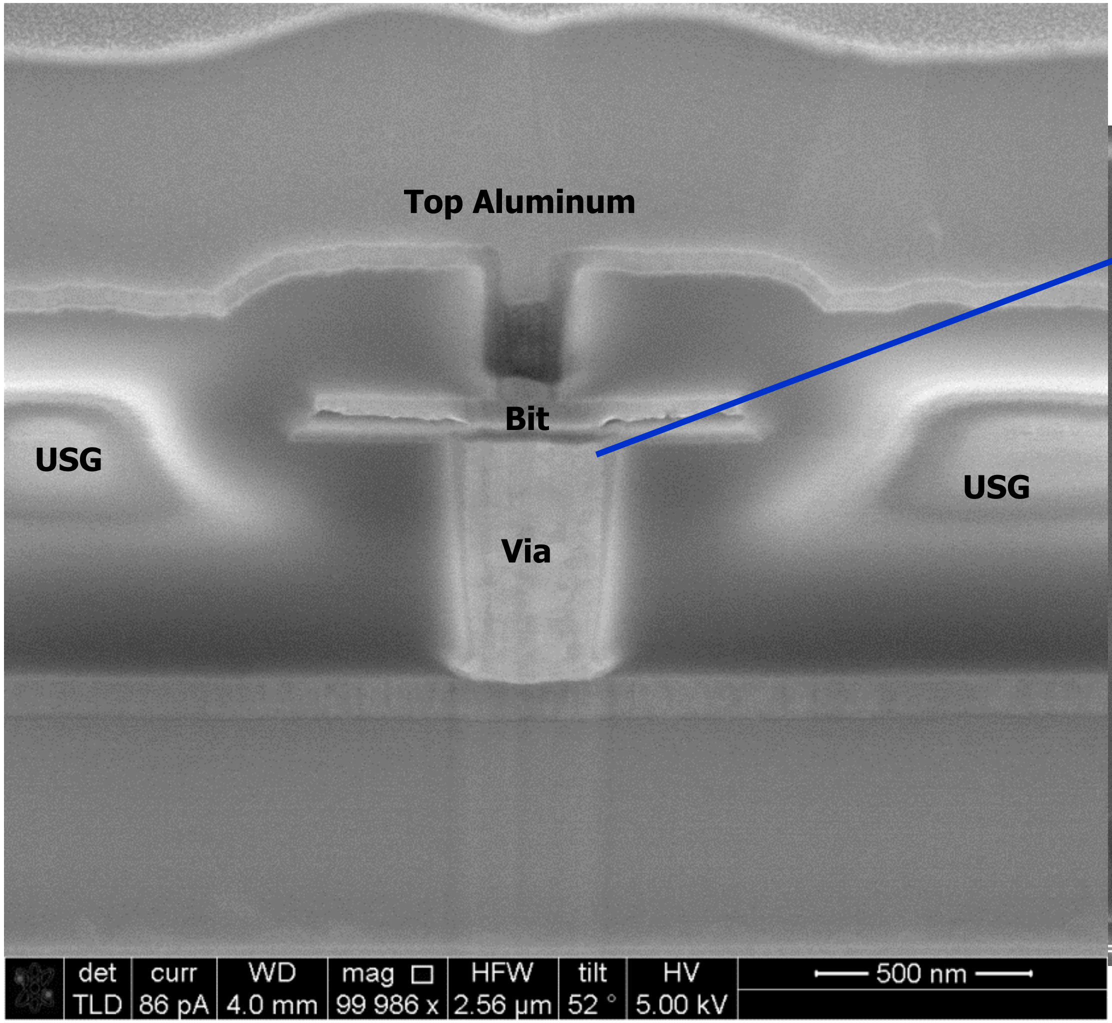
8. Deposit top USG
9. Etch top via holes in USG



10. Deposit top Al

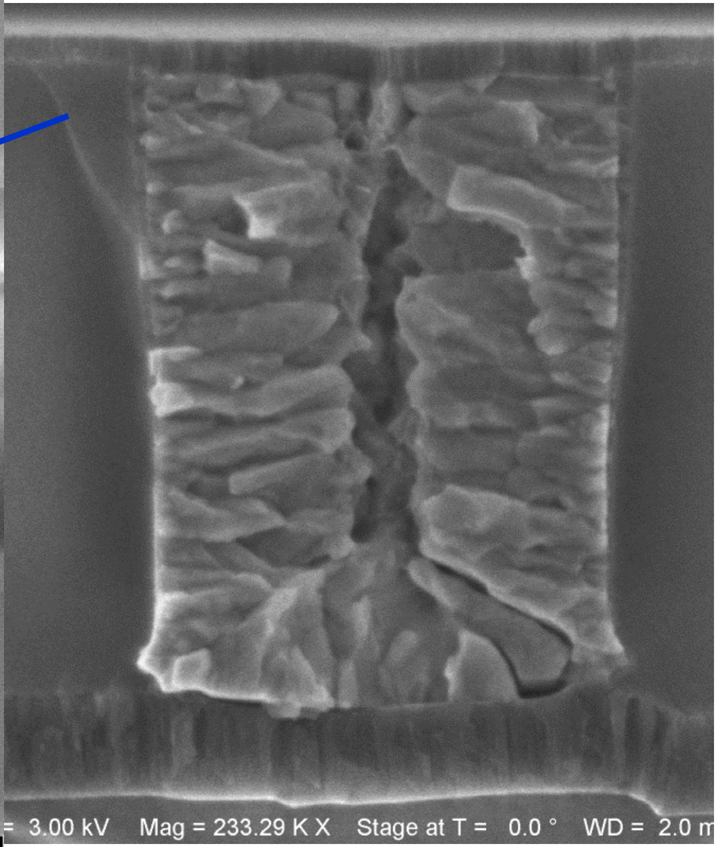


Final Structure



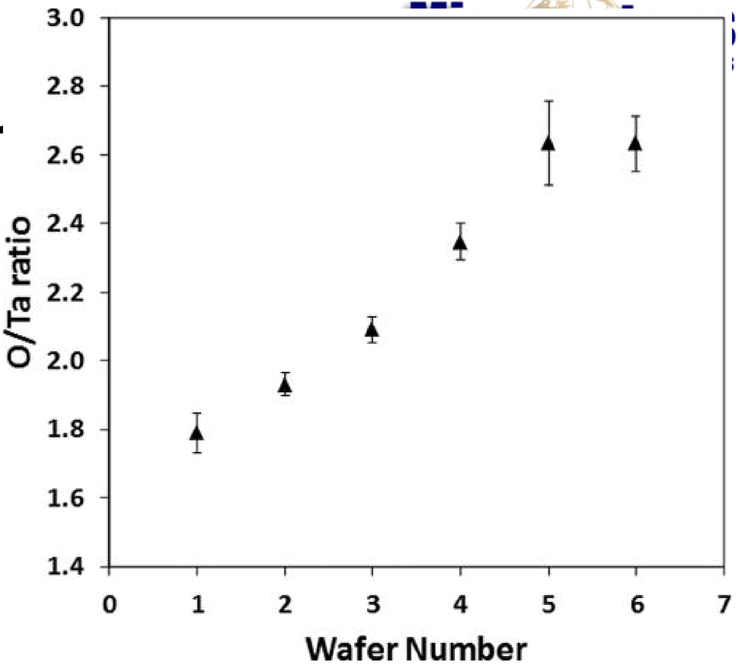
Important to have extremely flat surface under bit

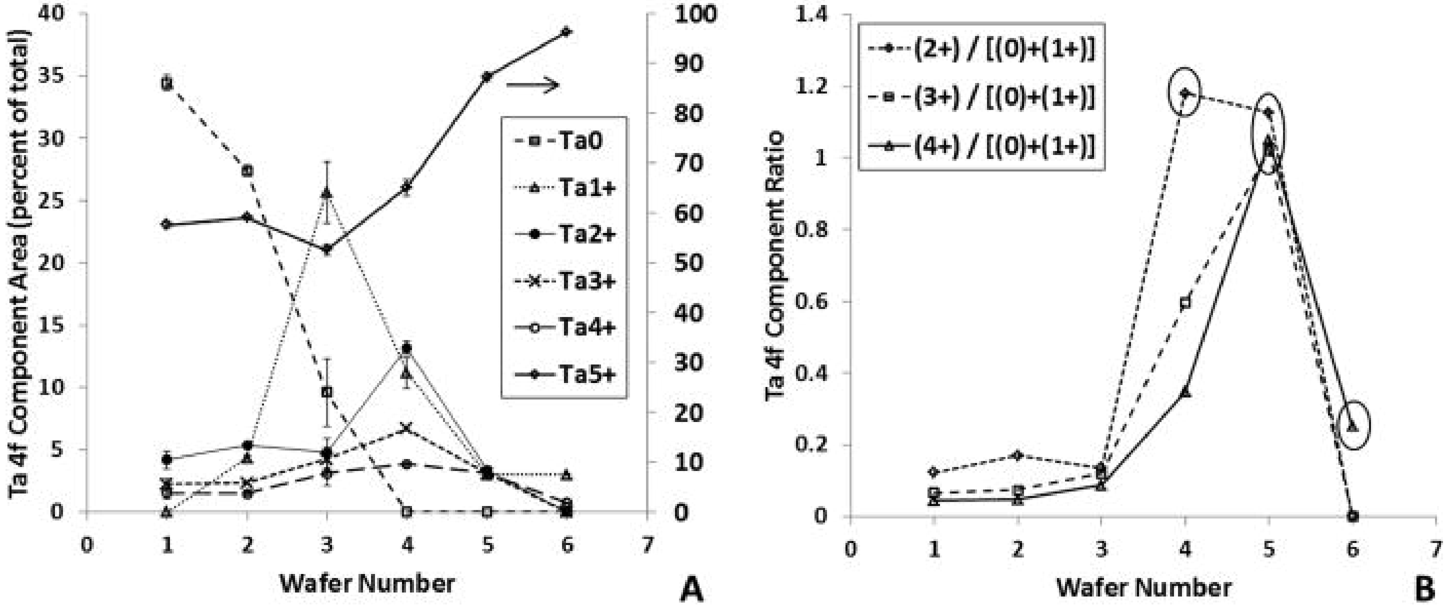
Polished TiN Surface



Beyond Stoichior

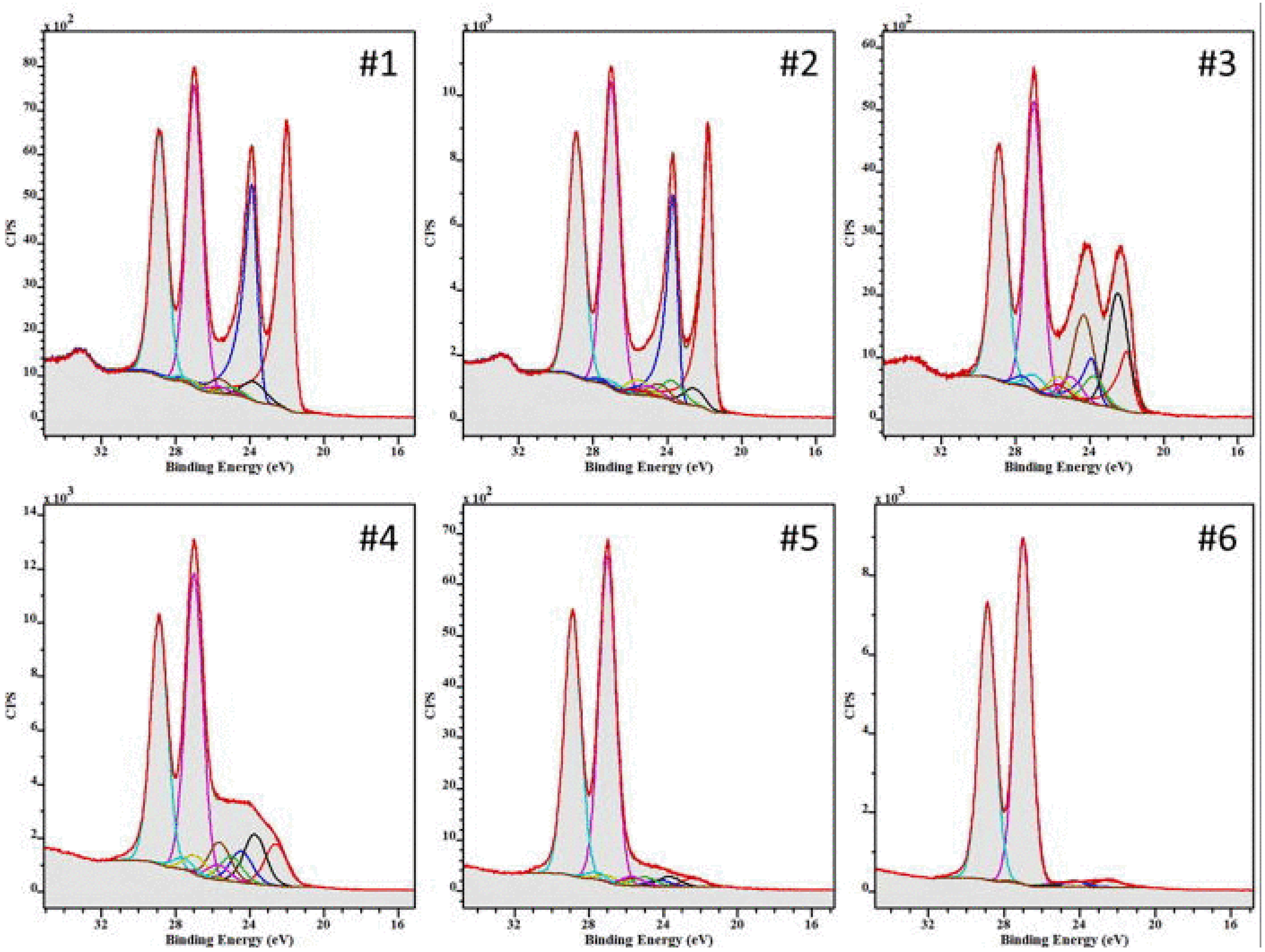
- Goal: Assess properties that make a “good” memristor
- Stoichiometry doesn’t tell the entire story
- Deconvolving XPS spectra provides Ta valence makeup





Brumbach et al, JVST 2014

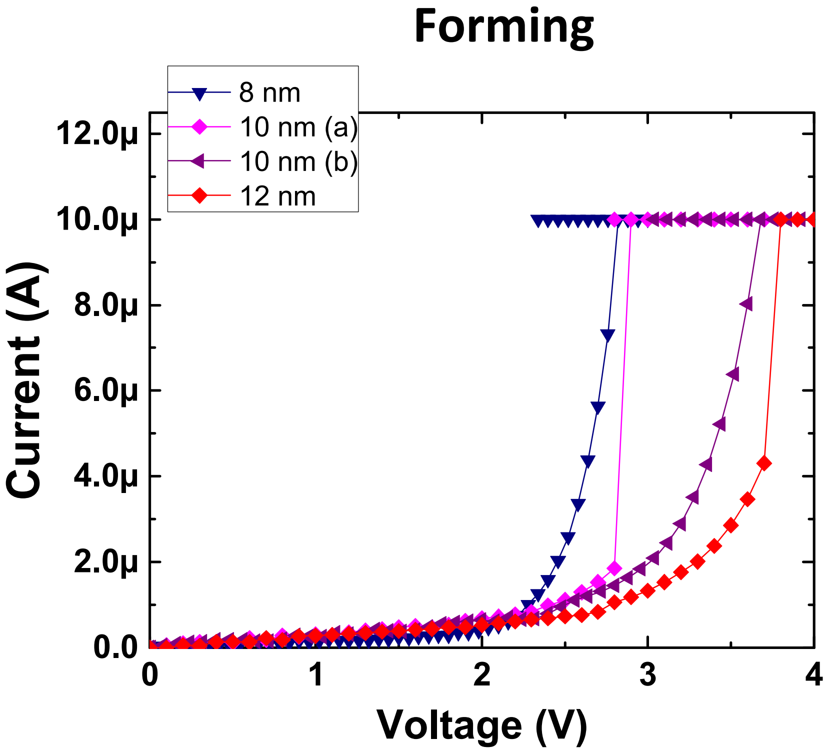
XPS: Properties of a Switching TMO

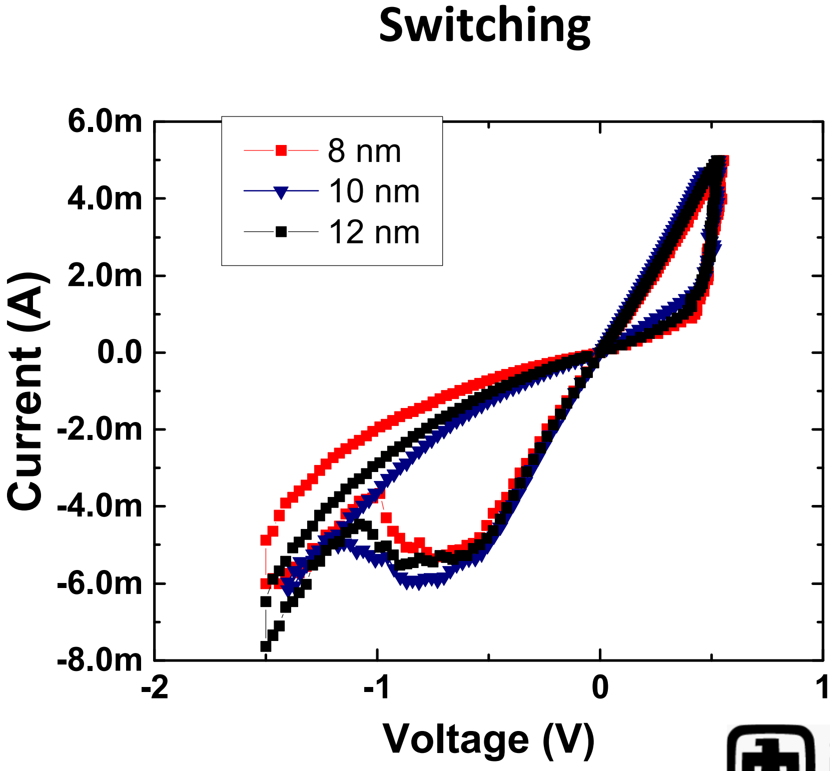


Brumbach et al, JVST 2014

Forming Process

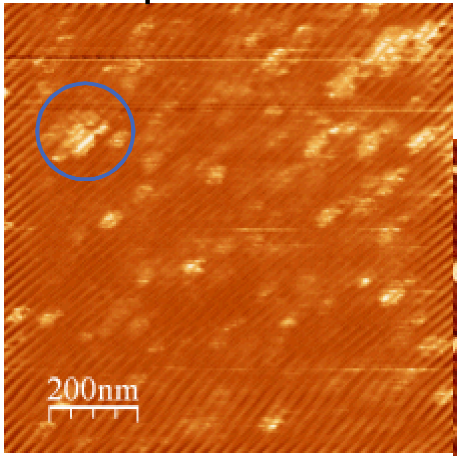
- Roughly depends on film thickness (still varies)
- Macroscopic (wafer scale) and nanoscopic variations in thickness



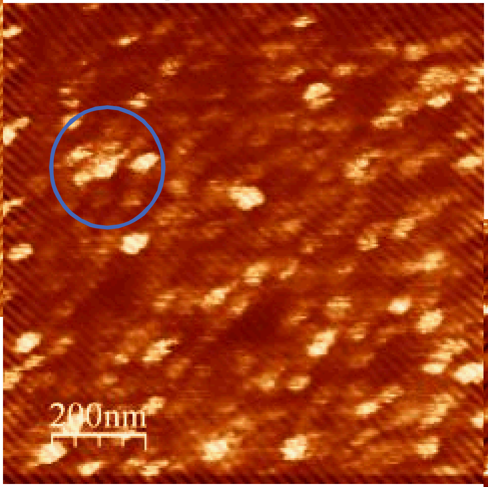


“Hot Spot” Formation

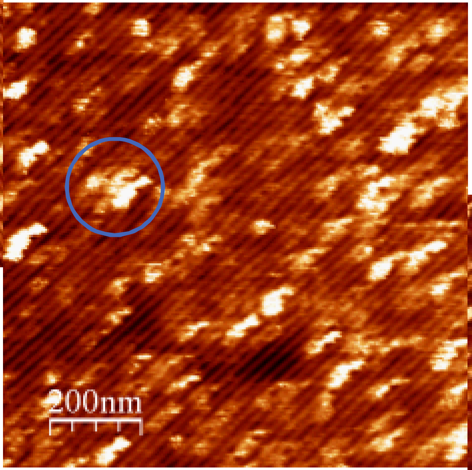
V_{tip}=2.4 V



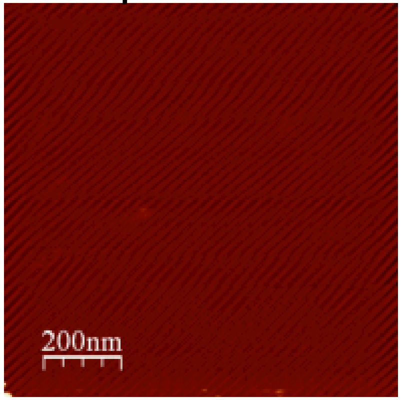
V_{tip}=2.6 V



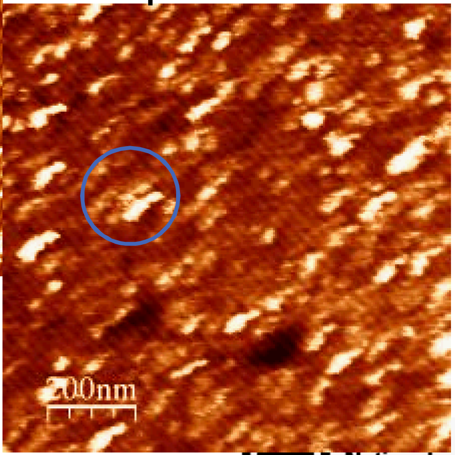
V_{tip}=2.8 V



V_{tip}=2.0 V



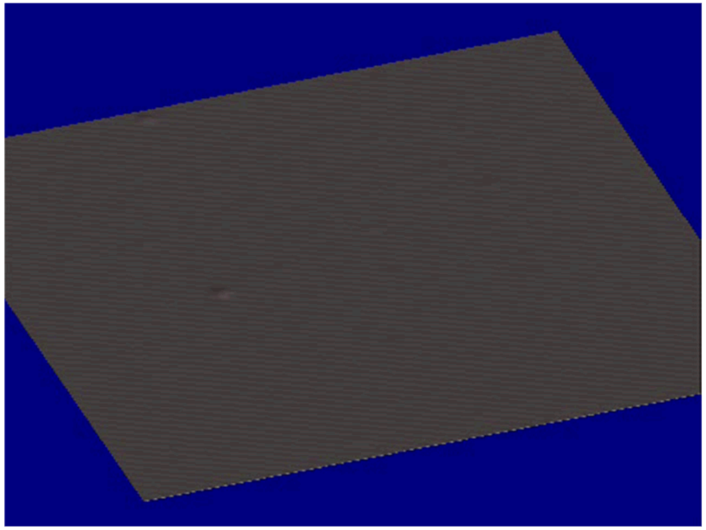
V_{tip}=3.0 V



Hot Spot Evolution

C-AFM Current Map Movie (2D)

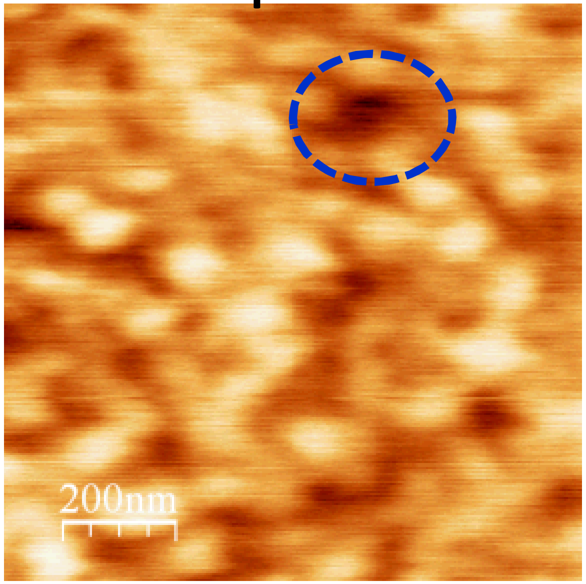
C-AFM Current Map Movie (3D)



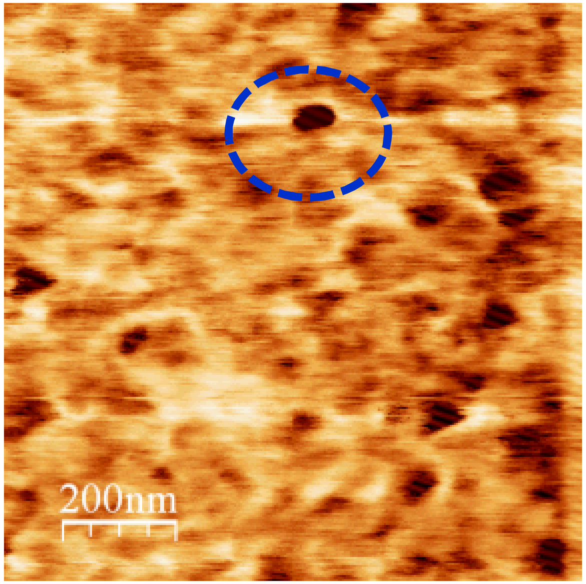
Topography versus Conductivity

- Prominent hot spot appears to depend on geometry
- Other hot spots do not necessarily correlate with geometric defect

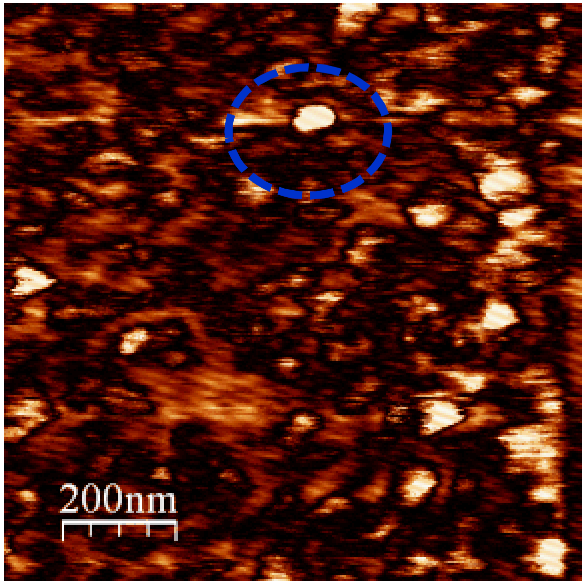
Topo



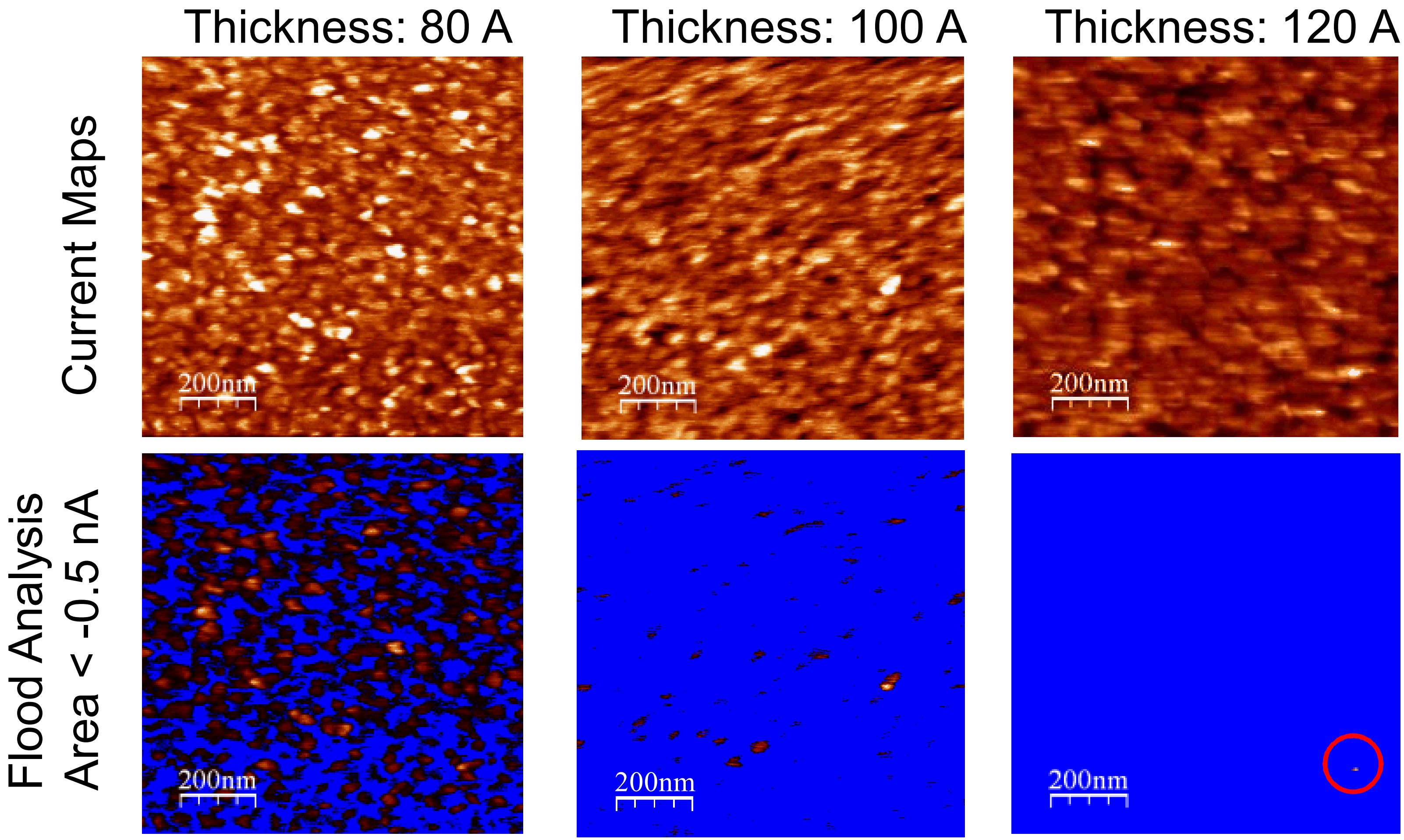
Current



Merged (And)



Hot Spot Density vs Thickness



Thickness: 80 A

Thickness: 100 A

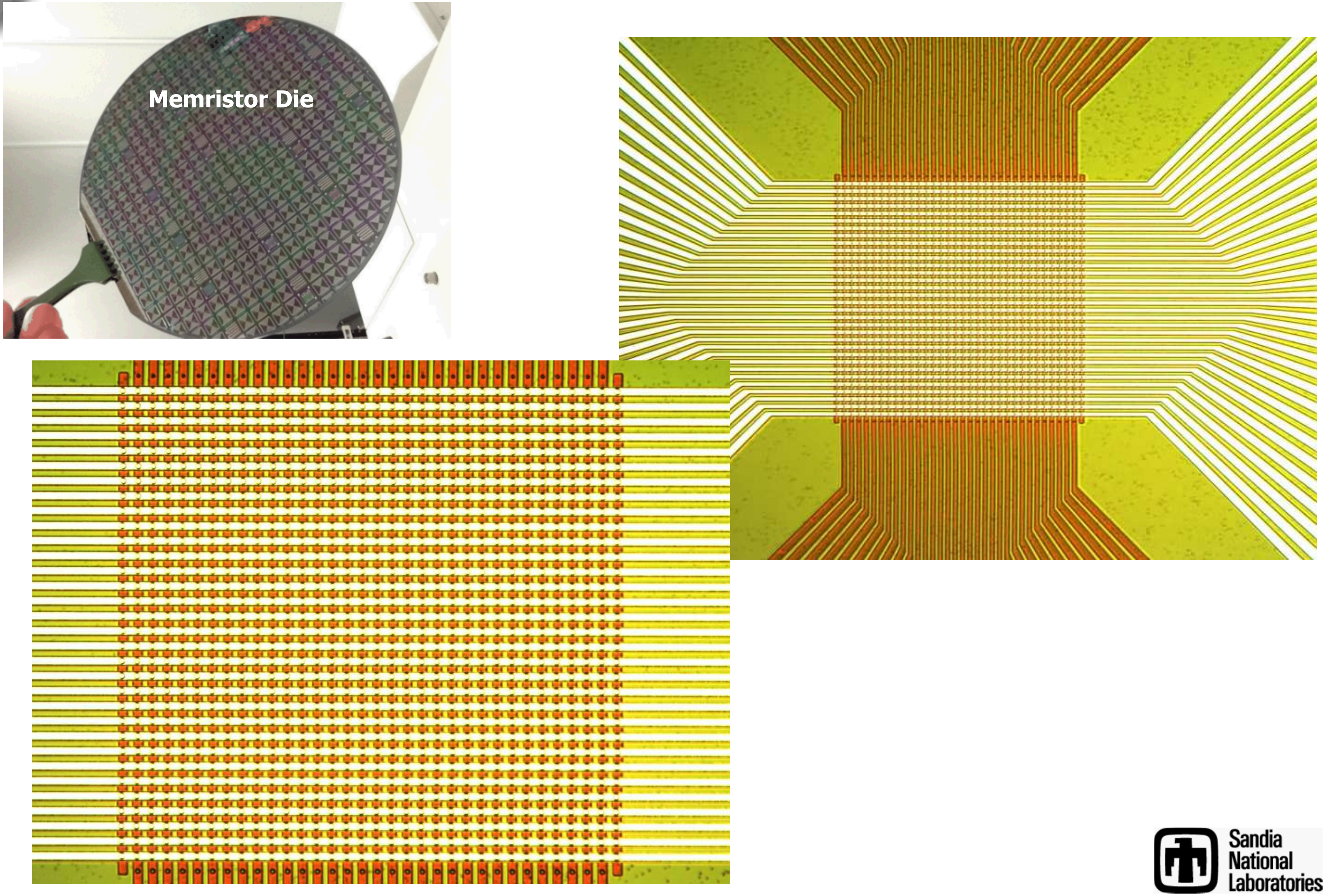
Thickness: 120 A

Current Maps

Flood Analysis
Area < -0.5 nA

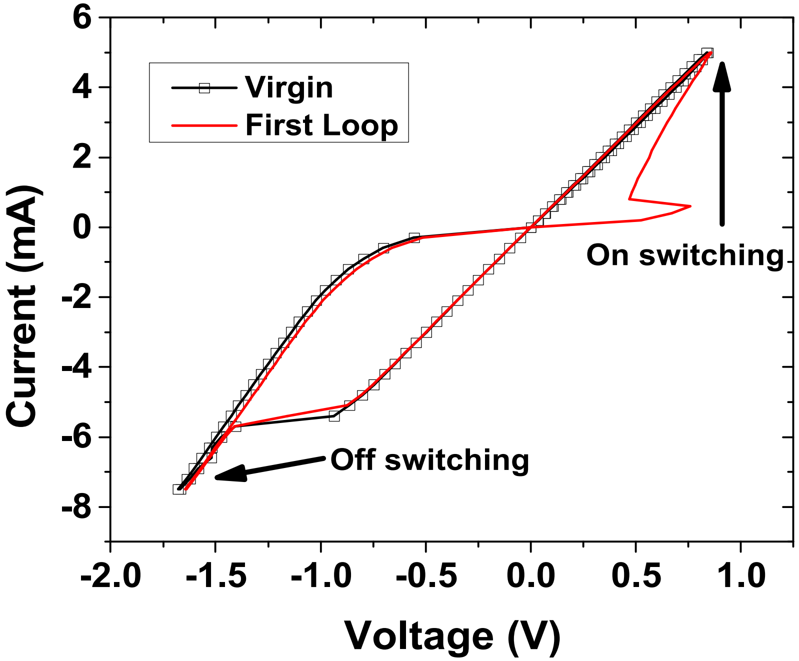
Sample Voltage: -3.5 V
Same tip used for all imaging

Memristor Crossbar Die



Basic Device Performance

- Typical devices form at very low currents
- Appear “forming free” in current sweep mode
- Do not need a high voltage transistor!!
  - Unlike flash/SONOS
- Can be tailored by stoichiometry

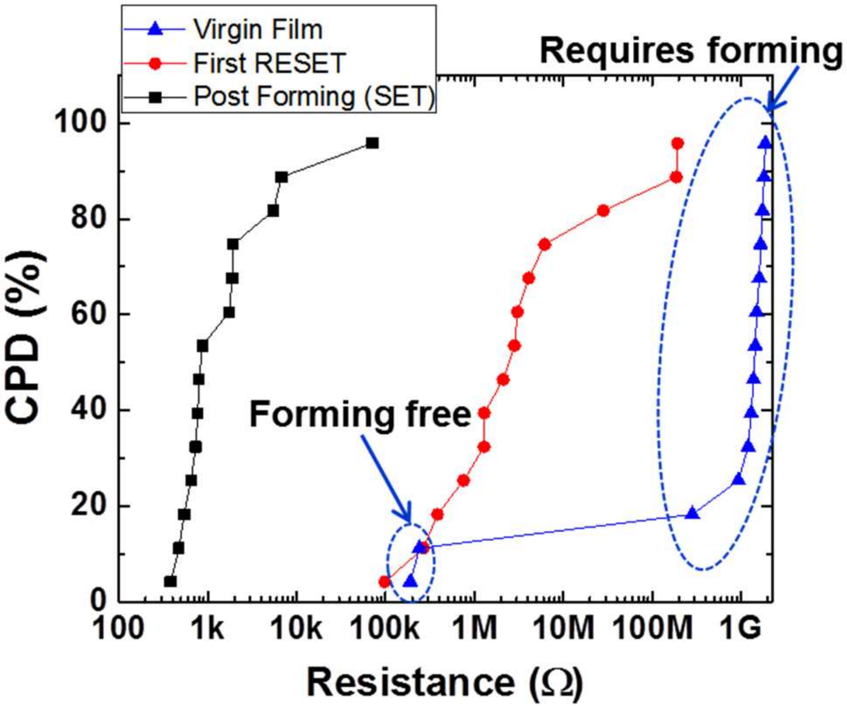


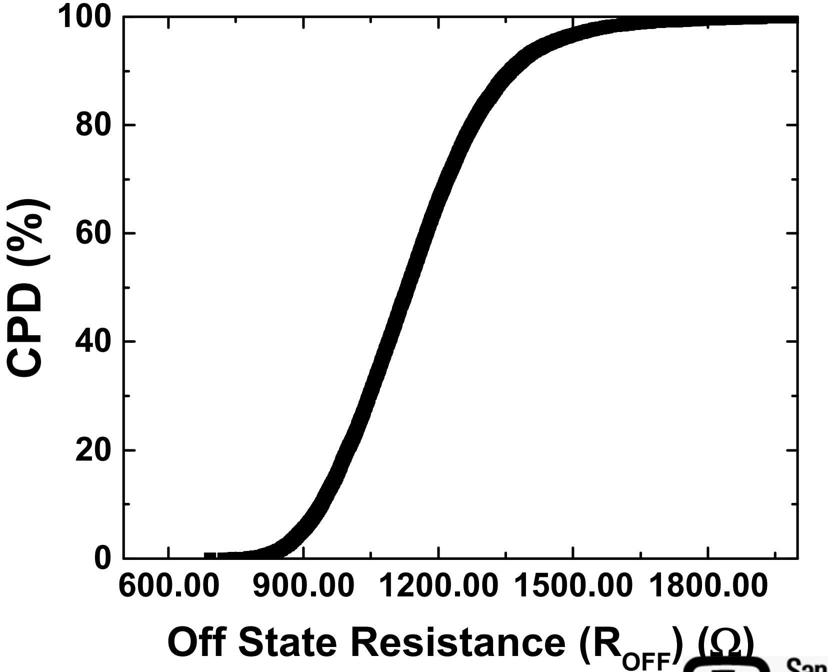
Outline

- Neuromorphic Computing with ReRAM
- Development of a CMOS/ReRAM Process
- Characterization of Device Behavior
- Modeling the Source of Intradevice Variation
- Conclusion and Future Work

Variability

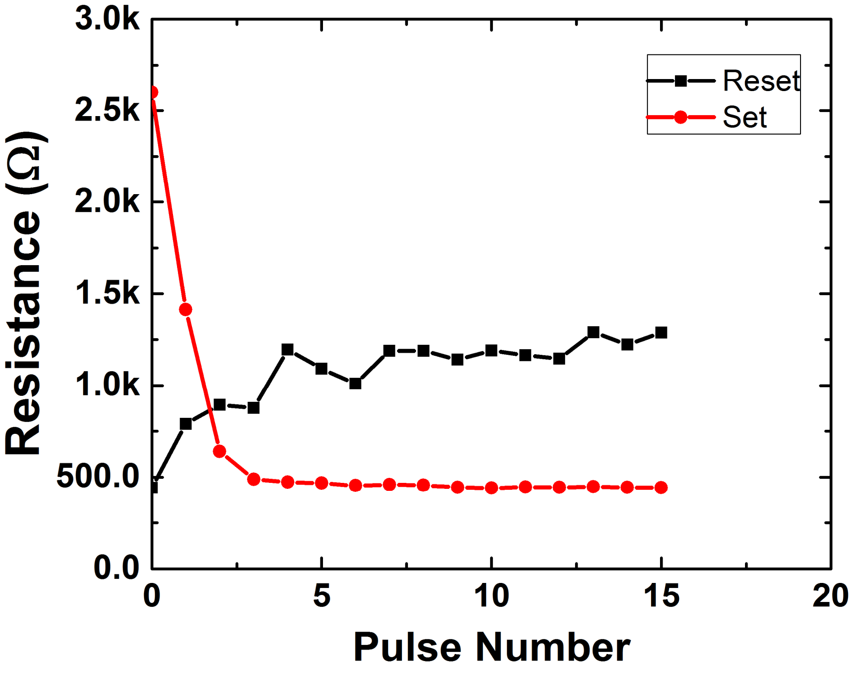
- Between different devices
  - Manufac
- Cycle to cycle
  - Fundamental physical attribute





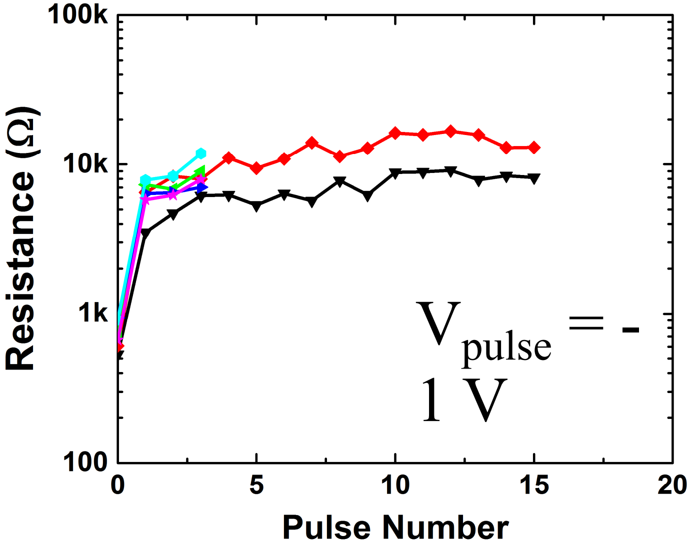
Set and Reset Transition

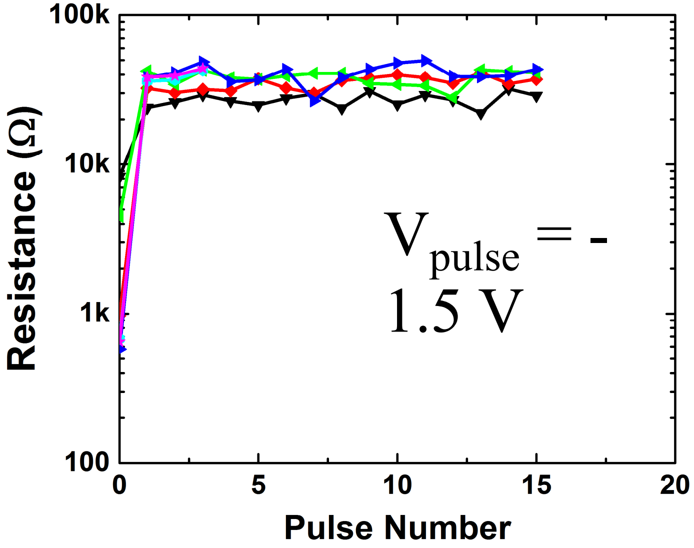
- Repeated pulsing can gradually change resistance
  - – SET transition more abrupt



Transitions to Varying Resistance Levels

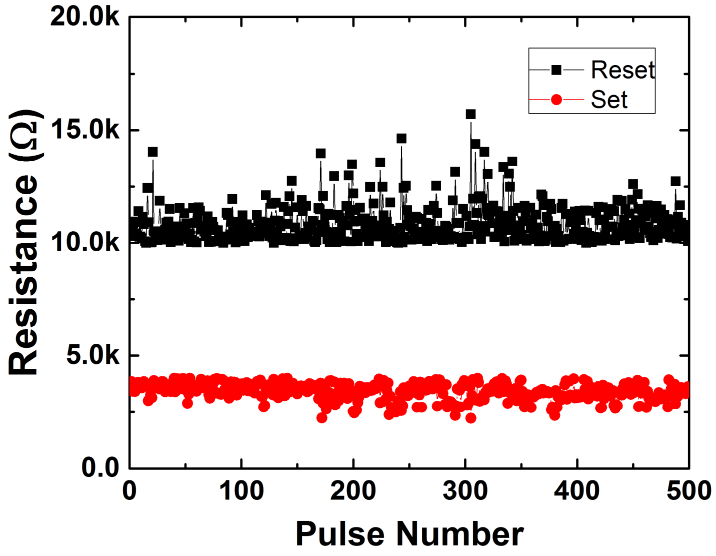
- Resistance level can begin to saturate for different voltage levels
  - Repeated pulsing may change resistance very gradually
  - Transitions occur from multiple starting resistance values
- Resistance usually falls into a typical band within the first three pulses

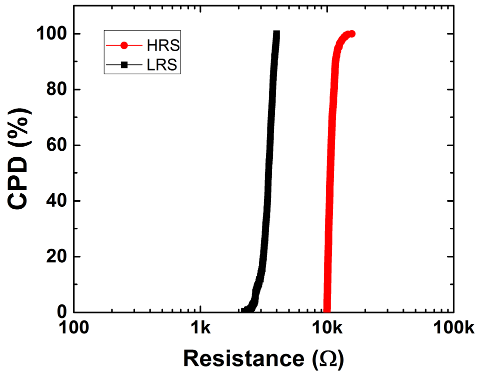


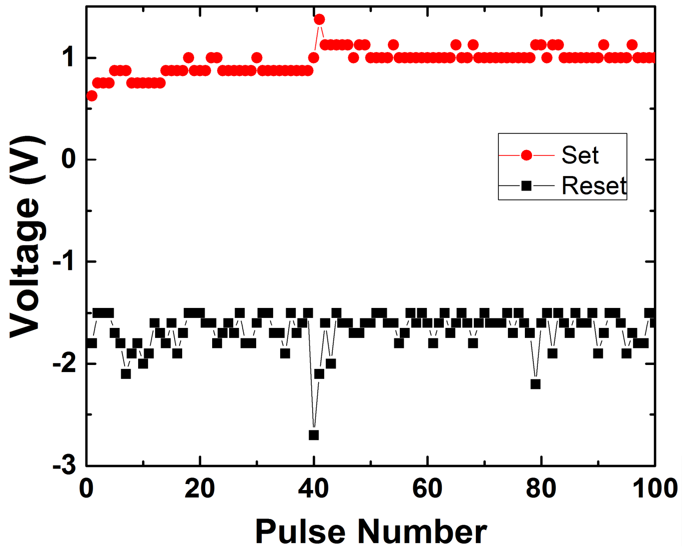


Closed Loop Cycling

- Continues pulsing until a threshold resistance value is passed
  - Allows for tighter resistance values
  - Helps with initial characterization of parts to determine ideal pulse heights

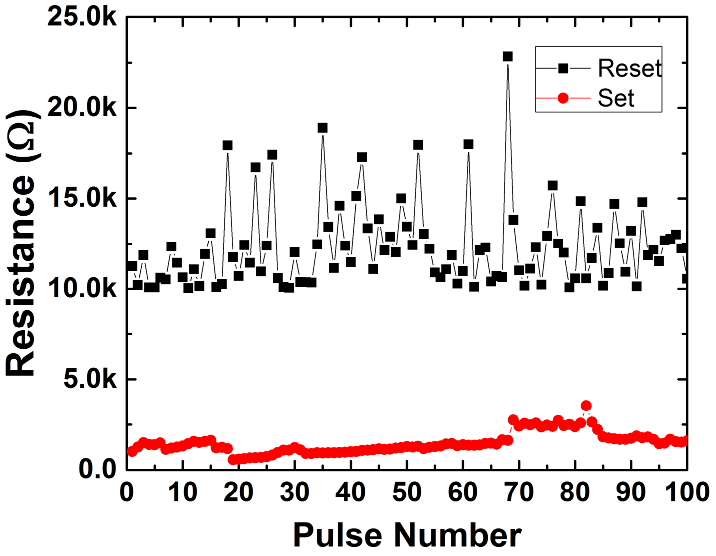


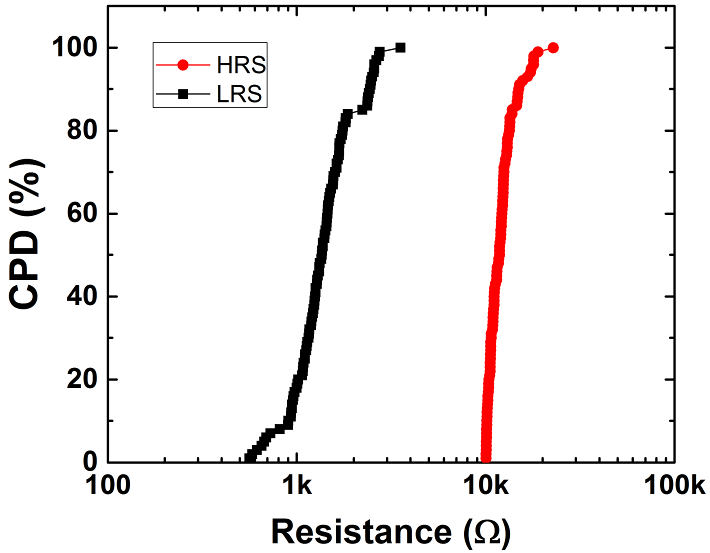




Open Loop Cycling (Set)

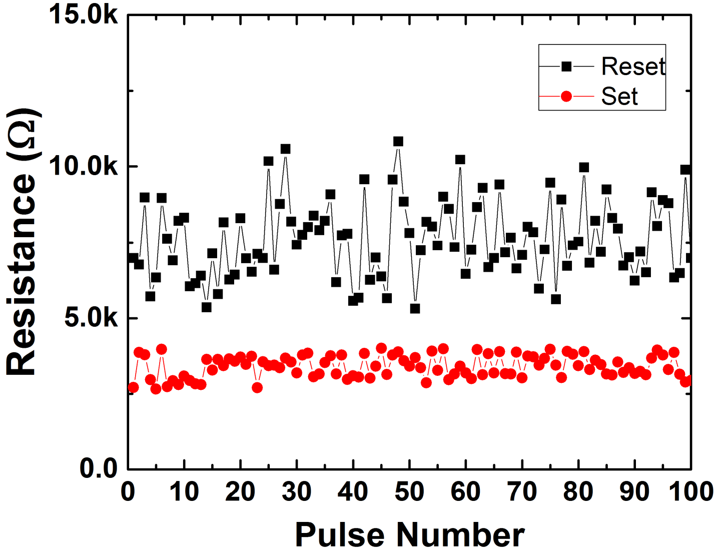
- Based off the closed loops tests, a value of 1 V is used for an open loop test of the Set function
- The CPD for the LRS state has a larger spread of resistance values

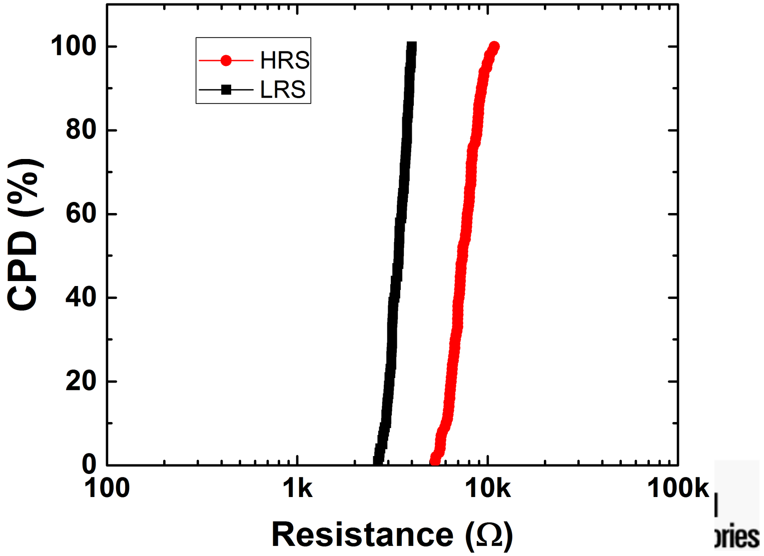




Open Loops Cycling (Reset)

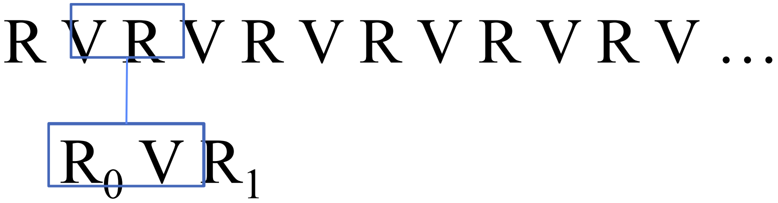
- For the reset open loop, a value of -2 V was used
  - More spread and lower resistances than the closed loop
- Feedback from closed loop tests can continue to improve open loop results





Megasample collection method

- Use HP test board
  - Capable of reading or pulsing a given row-column pair
  - Voltage pulse time set to constant 2us
  - Read voltage level = 0.25V
  - 10 samples averaged per read
- Interleave voltage pulses and resistance readings: R V R V R V …
- No attempt to force a particular starting resistance
- Careful control of voltage pulses to keep R within operating range
- 10M reads in working dataset
  - Construction of table-based model
- Sliding window over readings
- Table has two axes: R_0 and V
- Cell size: 100Ω x 0.1V
- Cell contains change in resistance: $D=R_1-R_0$
- Bilinear distribution of each D to four nearest cells

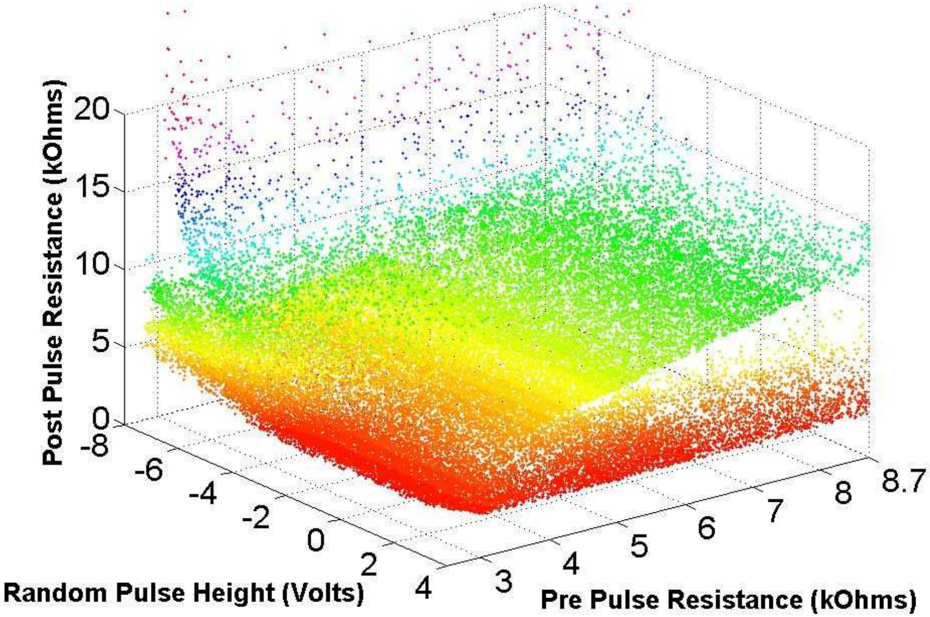


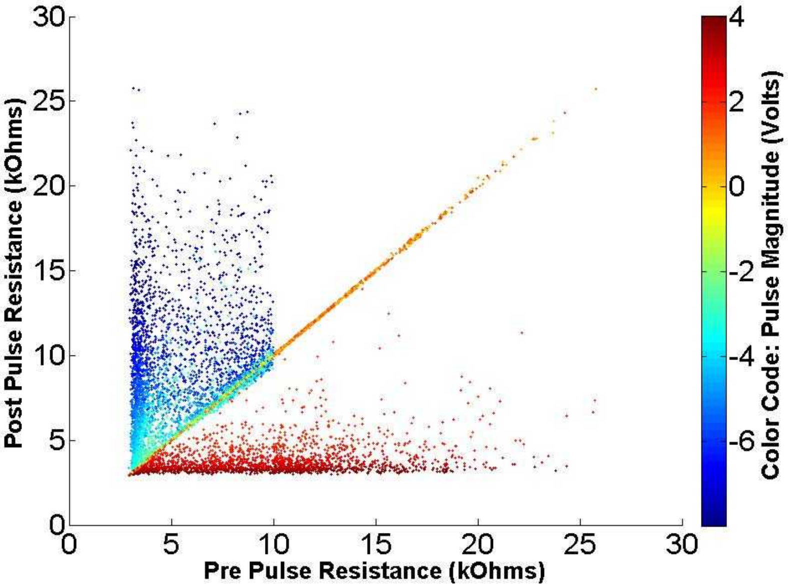
R V R V R V R V R V R V R V …

R_0 V R_1

Array Test Board

- **Random pulse technique**



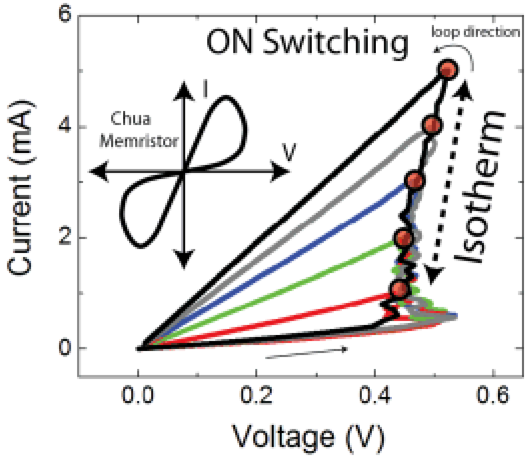


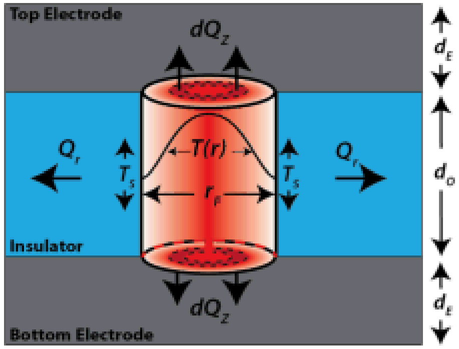
Outline

- Neuromorphic Computing with ReRAM
- Development of a CMOS/ReRAM Process
- Characterization of Device Behavior
- **Modeling the Source of Intradevice Variation**
- Conclusion and Future Work

Thermal Model

Isothermal signatures identified in IV data, which encouraged a thermal model





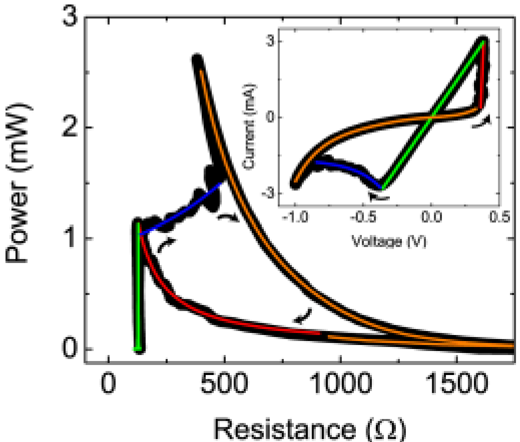
Approximation of heat equation for cylindrical filament gives temperature profile within the filament:

$$T(r) = T_{RT} + \frac{\sigma V^2 d_E}{2k_E d_O}\left[1 + \frac{k_E}{k_F}\,\frac{r_F^2 - 2r^2}{4\,d_E d_O}\right]$$

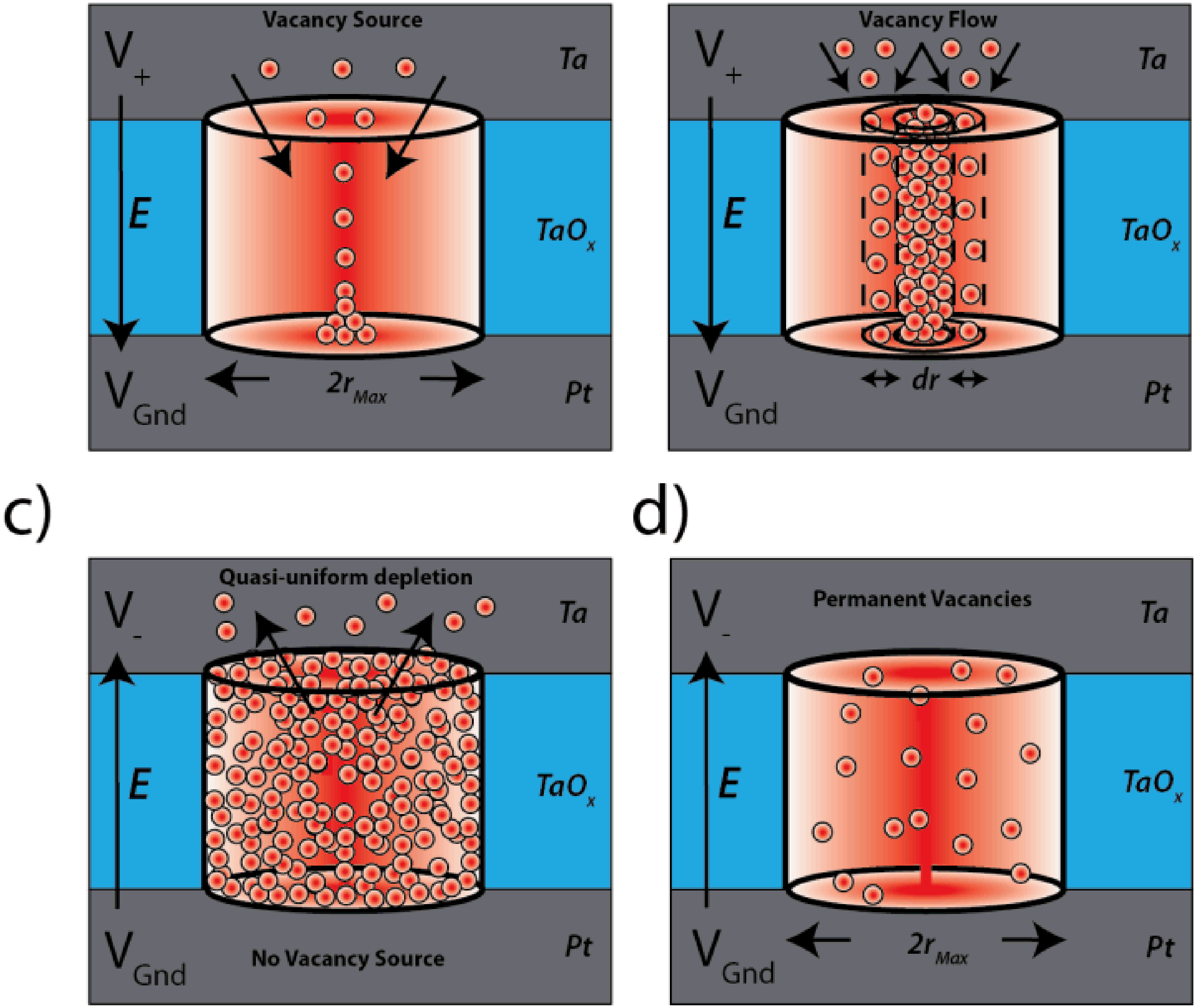
Temperature can be transformed into electrical parameters of power and resistance:

$$P = \frac{\Delta T/R}{\dfrac{d_E\sigma}{2k_E d_O} - \dfrac{r_F^2}{8L_{WF}T_{crit}d_O^2}}$$

Quantitative agreement w/data



Thermal Model



Mickel et al, Adv Mat (2014)

Two State Variables

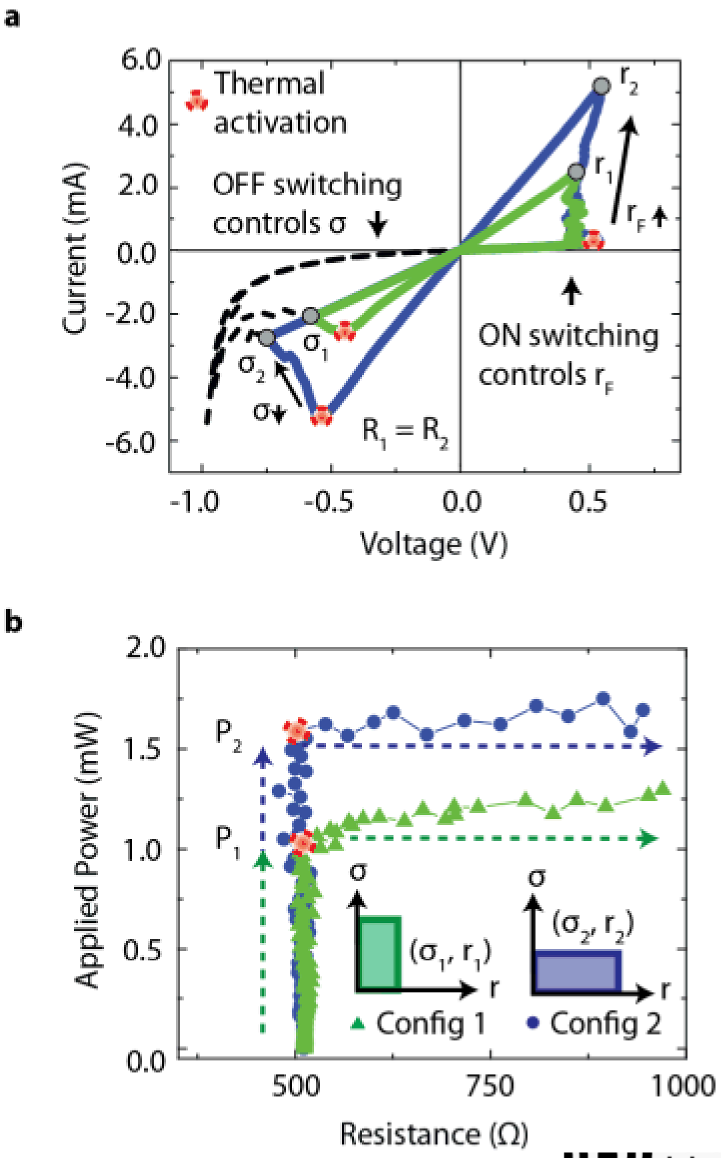
Independent control of σ and r means degenerate resistances can be set.

$$R = \frac{d_O}{\sigma \pi r_F^2}$$

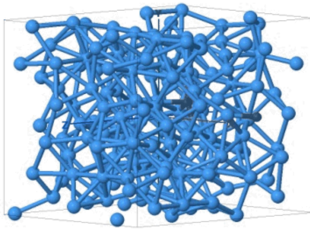
However, degenerate resistance states will activate at distinct applied powers

$$P = \frac{\Delta T / R}{\frac{d_E \sigma}{2 k_E d_O} - \frac{r_F^2}{8 L_{WF} T_{crit} d_O^2}}$$

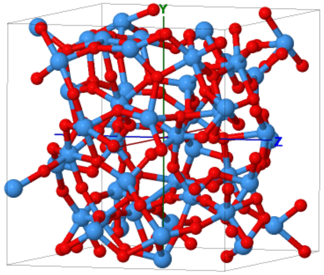
This additional parameter gives another “dimension” in which to encode analog or digital information



Ta_2O_x Atomistic Structure Set

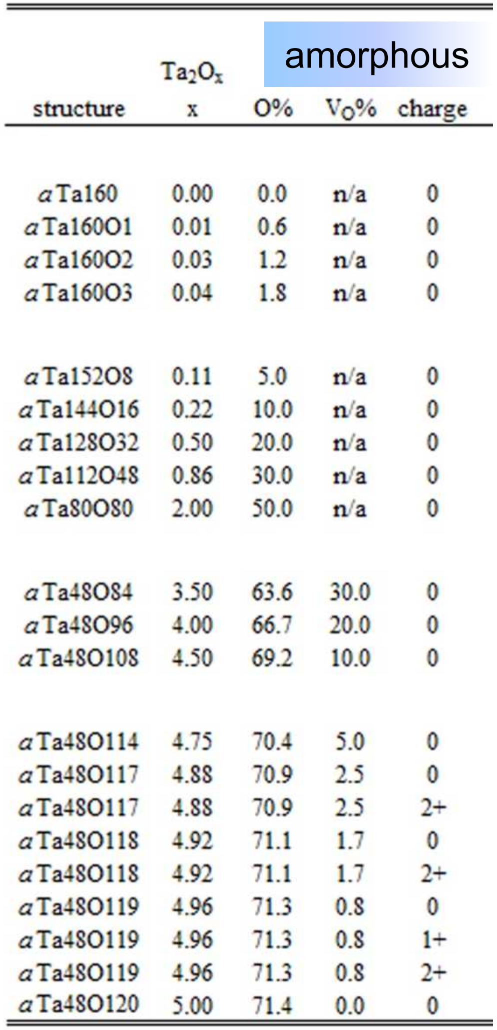


a-Ta160O2



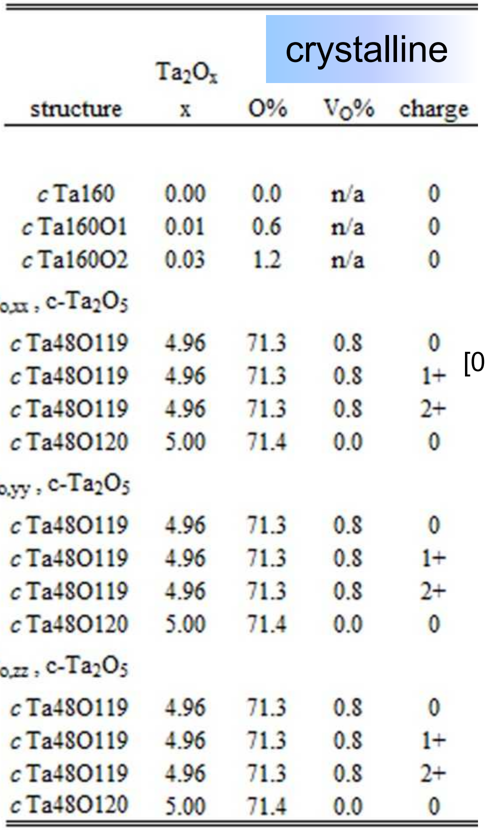
a-Ta48O120

amorphous

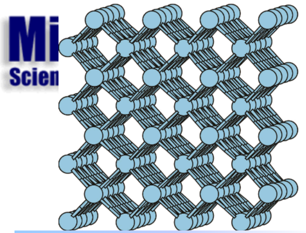


structure	Ta_2O_x x	O%	V_O%	charge
a Ta160	0.00	0.0	n/a	0
a Ta160O1	0.01	0.6	n/a	0
a Ta160O2	0.03	1.2	n/a	0
a Ta160O3	0.04	1.8	n/a	0
a Ta152O8	0.11	5.0	n/a	0
a Ta144O16	0.22	10.0	n/a	0
a Ta128O32	0.50	20.0	n/a	0
a Ta112O48	0.86	30.0	n/a	0
a Ta80O80	2.00	50.0	n/a	0
a Ta48O84	3.50	63.6	30.0	0
a Ta48O96	4.00	66.7	20.0	0
a Ta48O108	4.50	69.2	10.0	0
a Ta48O114	4.75	70.4	5.0	0
a Ta48O117	4.88	70.9	2.5	0
a Ta48O117	4.88	70.9	2.5	2+
a Ta48O118	4.92	71.1	1.7	0
a Ta48O118	4.92	71.1	1.7	2+
a Ta48O119	4.96	71.3	0.8	0
a Ta48O119	4.96	71.3	0.8	1+
a Ta48O119	4.96	71.3	0.8	2+
a Ta48O120	5.00	71.4	0.0	0

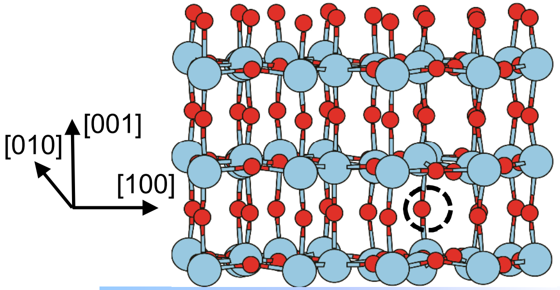
crystalline



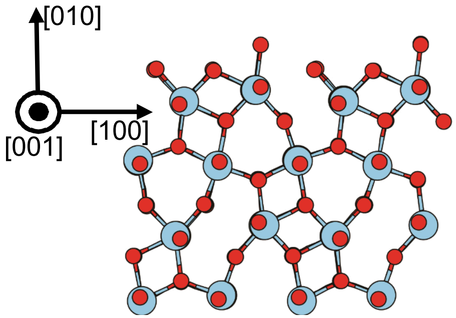
structure	Ta_2O_x x	O%	V_O%	charge
c Ta160	0.00	0.0	n/a	0
c Ta160O1	0.01	0.6	n/a	0
c Ta160O2	0.03	1.2	n/a	0
$\sigma_{o,xx}$, c-Ta_2O_5				
c Ta48O119	4.96	71.3	0.8	0
c Ta48O119	4.96	71.3	0.8	1+
c Ta48O119	4.96	71.3	0.8	2+
c Ta48O120	5.00	71.4	0.0	0
$\sigma_{o,yy}$, c-Ta_2O_5				
c Ta48O119	4.96	71.3	0.8	0
c Ta48O119	4.96	71.3	0.8	1+
c Ta48O119	4.96	71.3	0.8	2+
c Ta48O120	5.00	71.4	0.0	0
$\sigma_{o,zz}$, c-Ta_2O_5				
c Ta48O119	4.96	71.3	0.8	0
c Ta48O119	4.96	71.3	0.8	1+
c Ta48O119	4.96	71.3	0.8	2+
c Ta48O120	5.00	71.4	0.0	0



c-Ta160 (Ta)



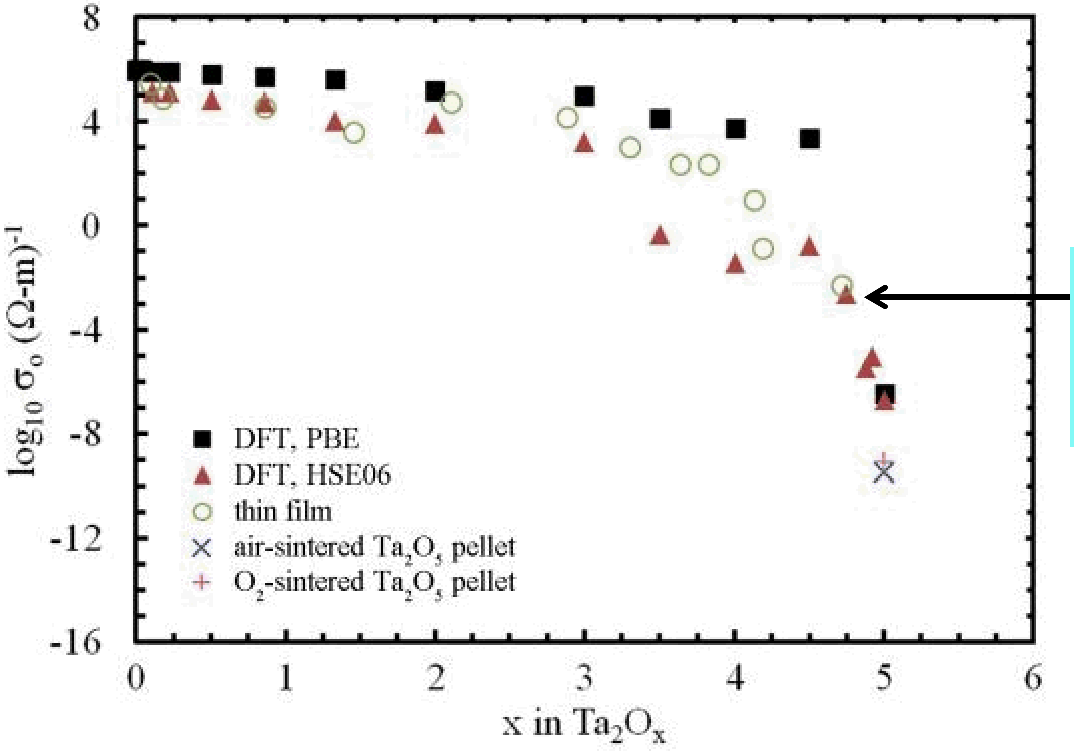
c-Ta48O119 (V_O^{2+})



c-Ta48O120 (Ta2O5)

- ►Ta_2O_x structure library generated for conductivity calculations
- ►Parameter space samples composition, phase, temperature, and charge state

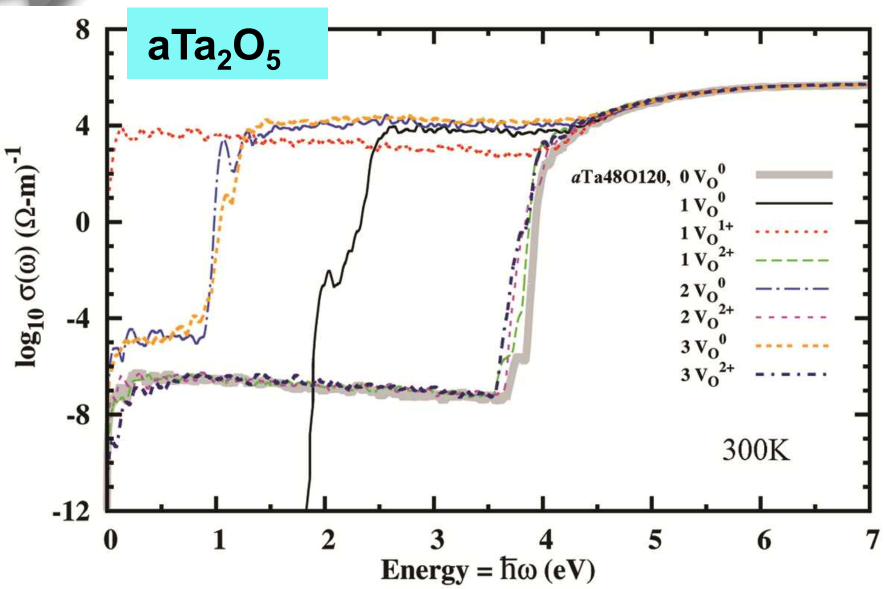
DFT vs. Experiment: *a*-Ta_2O_x σ

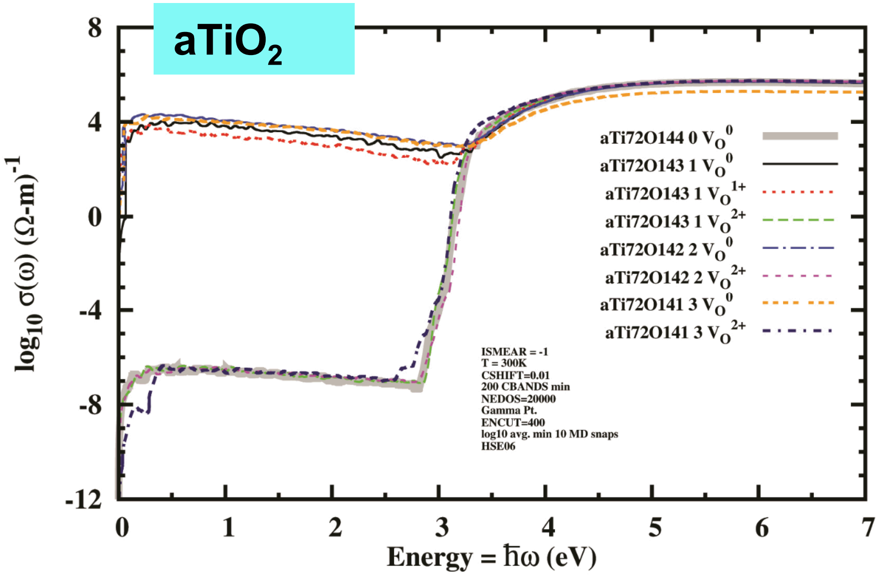


DFT outlier at x=4.75 joins trend with increased sampling ($\log_{10} \sigma_o$ decreases from 3.5 to -2.6)

- **DATA: T=300K; DFT conductivities sampled at 0.5 ps intervals on minimum12-14 (6) configuration snapshots for PBE (HSE06) functionals; MESA thin film resistivities measured with 4-pt probes; Ta_2O_5 pellets used to assess true bulk resistivity**
- **Sampling additional independently-quenched amorphous structures at each composition improves DFT trends (outlier at x=4.75 used as test case)**
- **DFT overestimation of σ_o is evident at x=5 (finite cell sizes are in part responsible)**

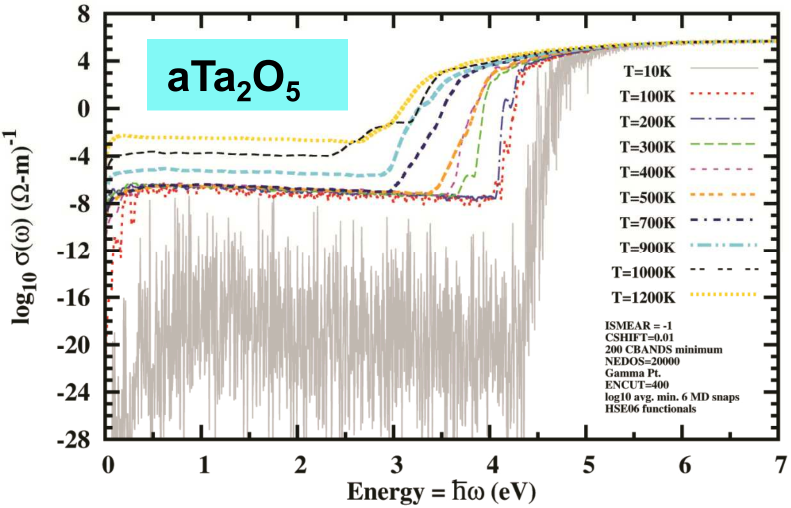
Conductivity: Oxidation State

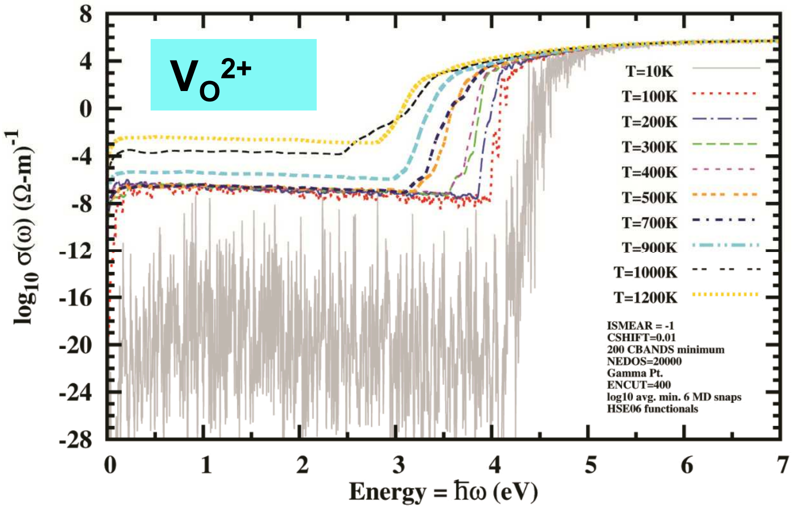




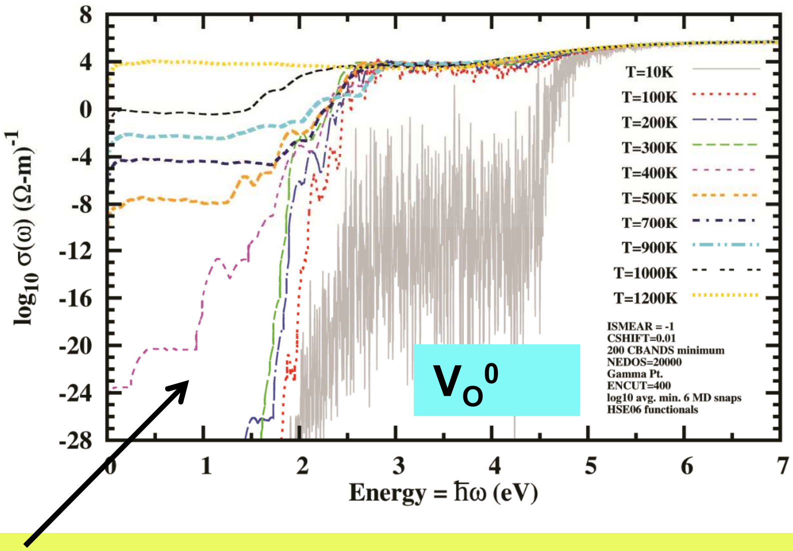
- **DATA: 300K, HSE06 functionals**
- **Rough Ta_2O_5 trend established for increasing σ_o as V_O^0 conc. increases (dopant-like behavior)**
- **Influence of V_O^n ($n=0,1^+,2^+$) oxidation state is significant; $\sigma(\omega)$ responses for all $n = 2^+$ cases essentially indistinguishable from stoichiometric oxides**
- **Possible that oxidation/reduction reactions are involved in memristor switching that effectively act as dopant deactivation/activation mechanisms for V_O^0.**

Ta_2O_x Conductivity: T and Vacancies



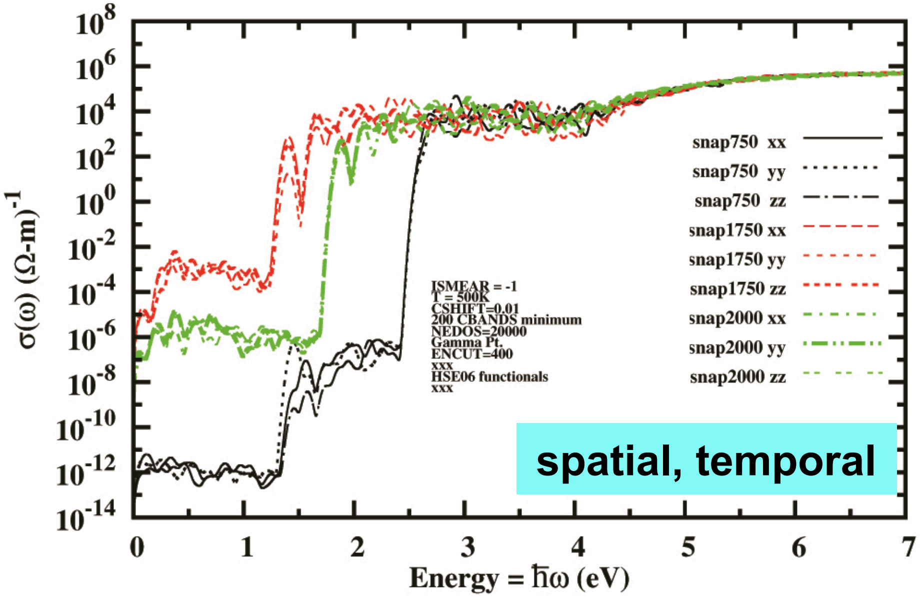


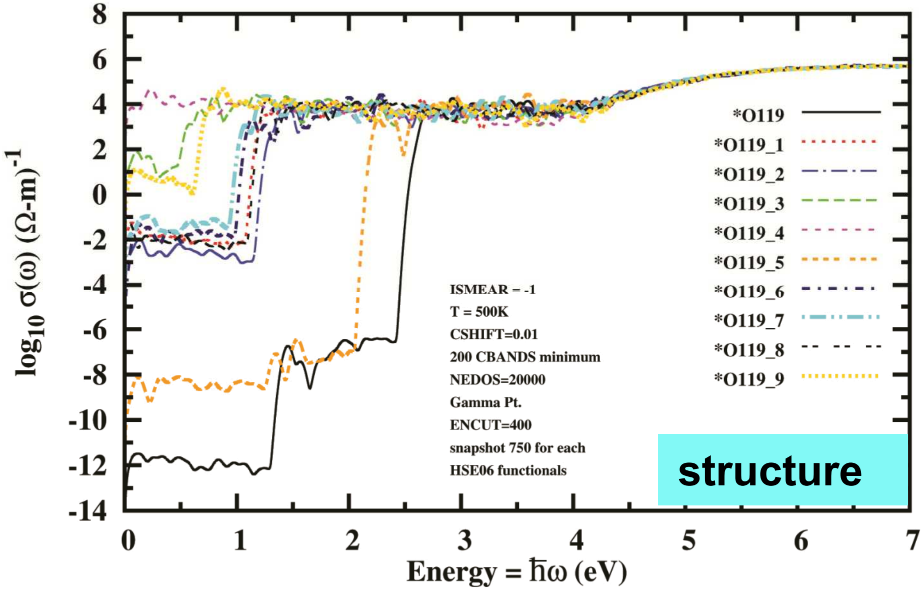
- DATA: 168-atom basis supercells, HSE06 functionals, 6 QMD configs. sampled at each of 10 to 11 temps., noisy gray curves are low T reference (10K)
- T dependence for stoich. oxide and V_O^{2+} case are very similar, while V_O^0 case is much different
- V_O^0 case exhibits stronger T-dependence of σ_o than other cases



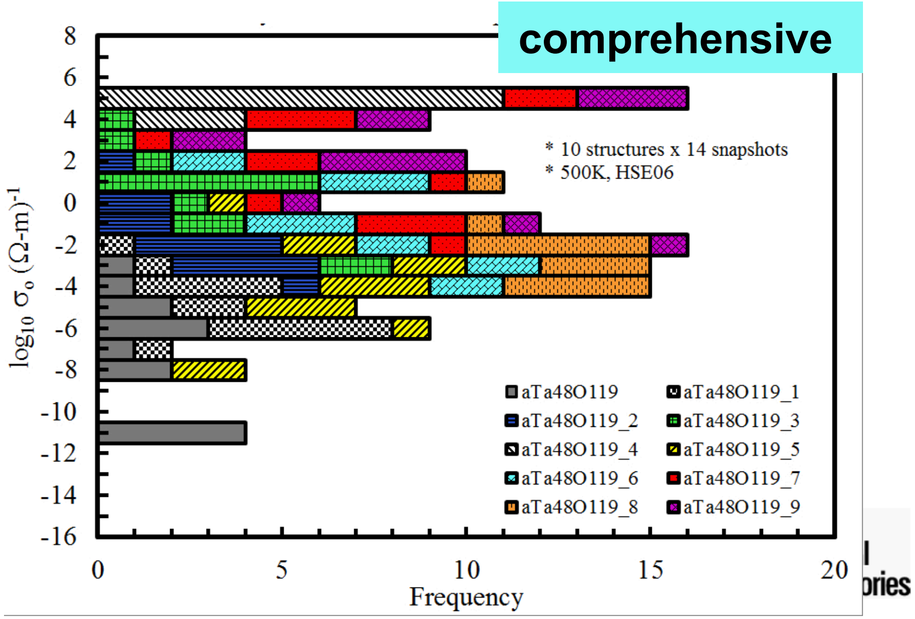
V_O^0 shows transition from zero σ_o to finite σ_o between 300 and 400 K. This is suggestive of "freeze-out" dopant behavior.

Statistical Contributions to Nanoscale σ





- DATA: 167-atom amorphous supercells containing 1 V_O^0, HSE06 functionals, T=500K
- Spatial variation: 1-2 orders of magnitude in σ_o
- Temporal variation: 9-10 orders of magnitude in σ_o
- Structural variation: 16 orders of magnitude in σ_o
- Anisotropy in σ_o relatively small effect at nanoscale



Outline

- Neuromorphic Computing with ReRAM
- Development of a CMOS/ReRAM Process
- Characterization of Device Behavior
- Modeling the Source of Intradevice Variation
- Conclusion and Future Work

Conclusions

- Neuromorphic computing is superior to traditional computing for certain applications, especially those involving pattern recognition
- Execution of certain neuromoprhic algorithms can be significantly improved using a custom analog-mode CMOS/ReRAM accelerator
- Metal Oxide ReRAM cells have relatively high cycle to cycle variability, which may significantly limit the resolution of an analog accelerator
- Possible physical origins have been studied using analytical and DFT models and two reasons hypothesized:
  - Degenerate resistance states
  - Different same-stoichiometry makeup

Acknowledgements

- This project was funded in part by Sandia's Laboratory Directed Research and Development (LDRD) Program
- Useful discussions with our collaborators at HP Labs, esp. Jianhua Yang, Yoocharn Jeon, John Paul Strachan, Si-Ty Lam, Brent Buchanan, Dick Henze, and Stan Williams

Backup Slides

Neural Hardware Simulation

- Greatest challenge is to generate row and column pulses that train memristors correctly.
- Both write noise and bin size are 6 orders of magnitude larger than weight update.
- Training curve of memristor contains highly sensitive “cliff” in one direction.

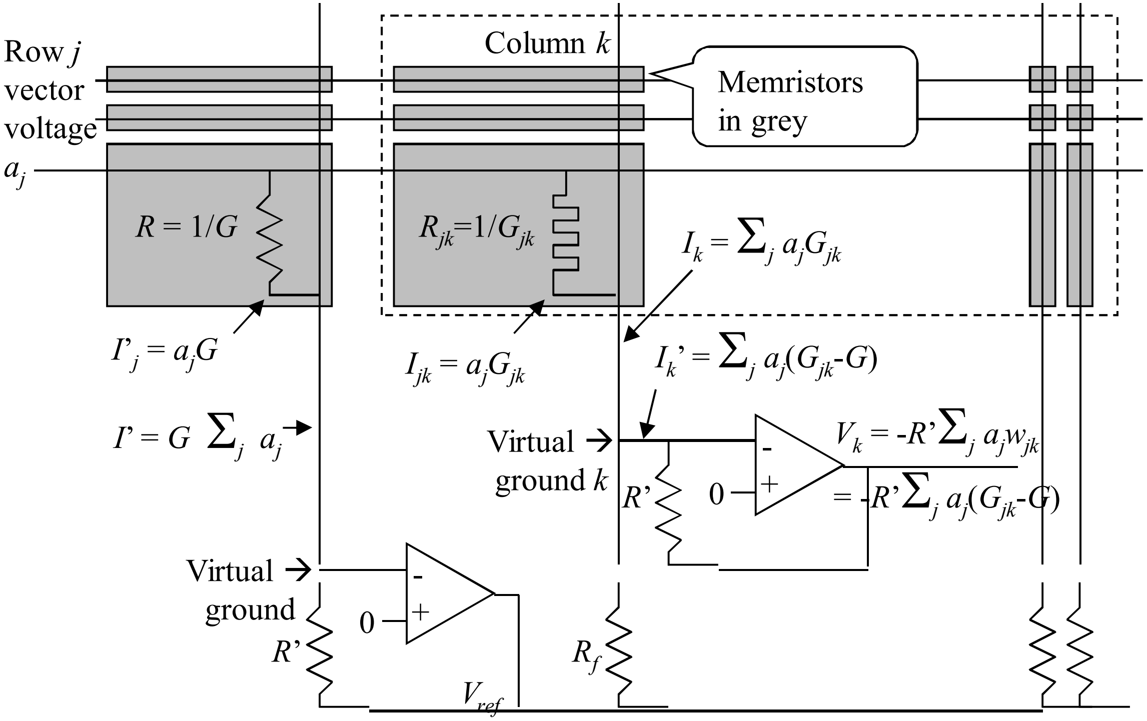
Neural Hardware Simulation

- Methods:
- Direct conversion – Compute all weights using float. Simply write resistance values into simulated crossbar without “burning” them. Works.
- Direct set – Store weights as resistances, but do all math in float. Works.
- Isolated pulse – Apply a voltage pulse to each memristor in isolation to set it. Works
- Binary – Pulse all rows and columns at once. Voltage is minimal fixed increment. Only applied if given row had error or given column had input. Diverges.
- Binary Row – Similar to binary, but only pulse row with largest error. Diverges.

Neural Hardware Simulation

- Methods (continued):
- Linear Row – Pulse only the row with largest error, but use a scaling factor to convert error to voltage. Works.
- Linear Rows – Pulse each row separately, using same method as Linear Row. Works.
- Linear – Pulse all rows and columns at once, using scaling factor to convert input or error to voltage. Diverges.
- Polynomial – Similar to Linear, but use polynomial fit of memristor data with one real exponent. Diverges.

ReRAM Neural Circuit



Emerging Memory

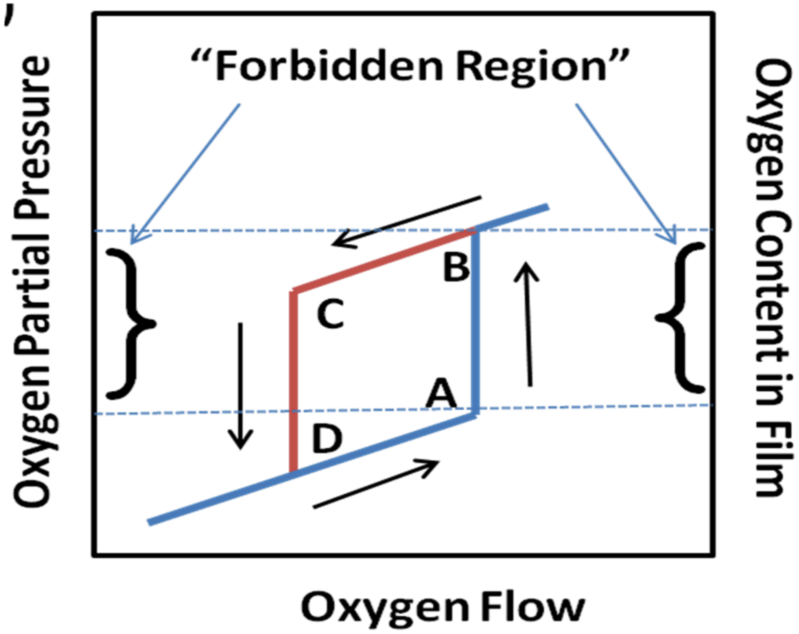
- This is a great era for emerging memory
- <u>NAND Scaling is visibly slowing</u>
  - Memory manufacturers refusing to name nodes by physical dimensions (now we have 2x and 1x nodes)
  - 16 nm retention and endurance degraded
  - 3D will quench density issues temporarily
- <u>DRAM scaling is also becoming a problem</u>
  - struggling to maintain reasonable equivalent oxide thickness
  - Dielectric for cells 30nm to 20 nm still TBD
- Opportunity: Storage Class Memory
  - Magnetic to DRAM latency gap
- New memory technologies on the horizon are rapidly maturing which can replace NAND *and* DRAM

Categorization of ReRAM

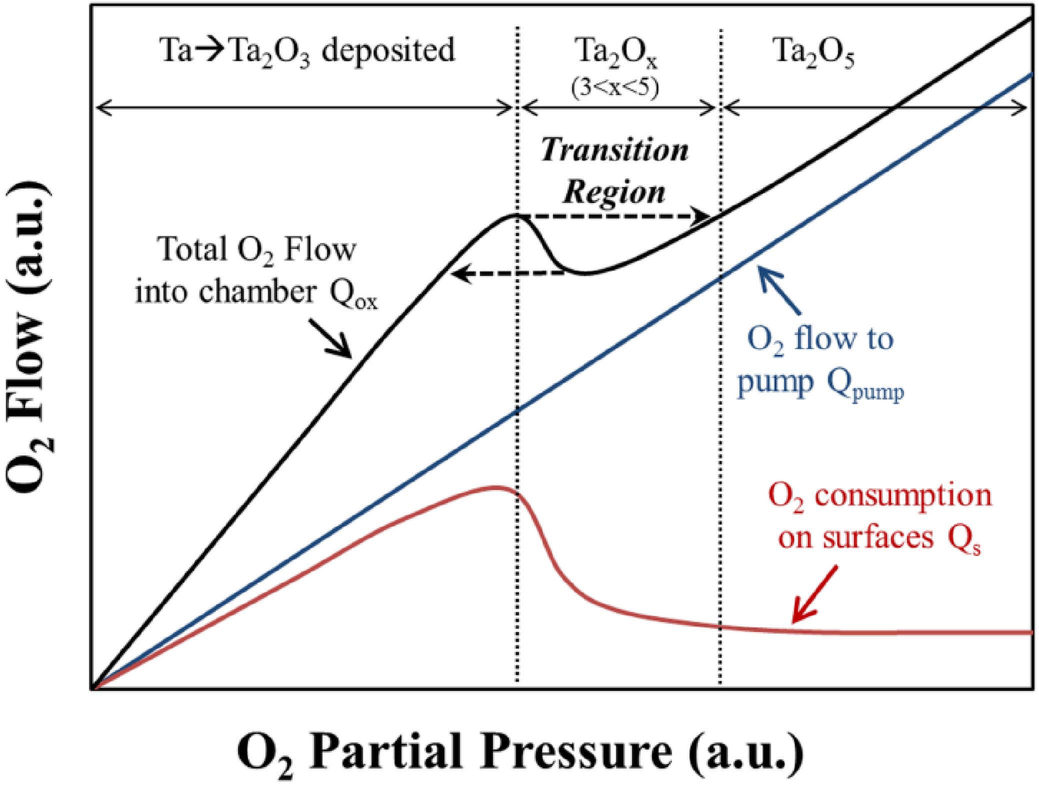
- Electrochemical Metallization Bridge (CBRAM)
  - Bipolar
  - Cation motion
  - Ag or Cu filament
- Metal Oxide: Bipolar Filamentary
  - Current independent of area
  - Anion (oxygen vacancy) motion
  - Valence change dominates
- Metal Oxide: Unipolar Filamentary
  - Current independent of area
  - Thermochemical mechanism dominates
- Metal Oxide: Bipolar Non-filamentary
  - Current depends on area
  - Anion motion near interface

Switching Film Development: Overcoming the “Forbidden Region”

- Forbidden oxygen flow-pressure region occurs due to target poisoning
- This is the region we need to be in to get ideal ReRAM stoichiometry



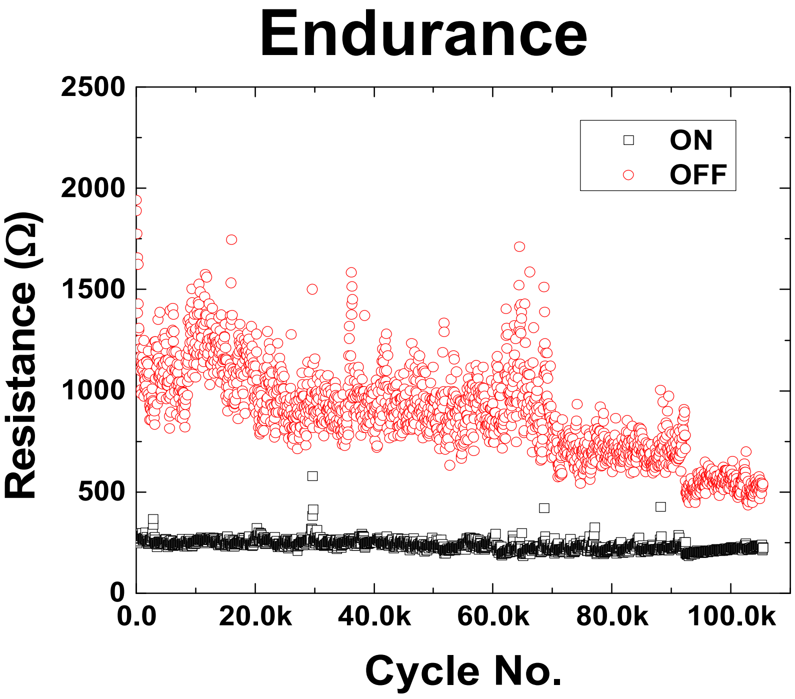
A.J. Lohn et al APL 103, 063502 (2013)

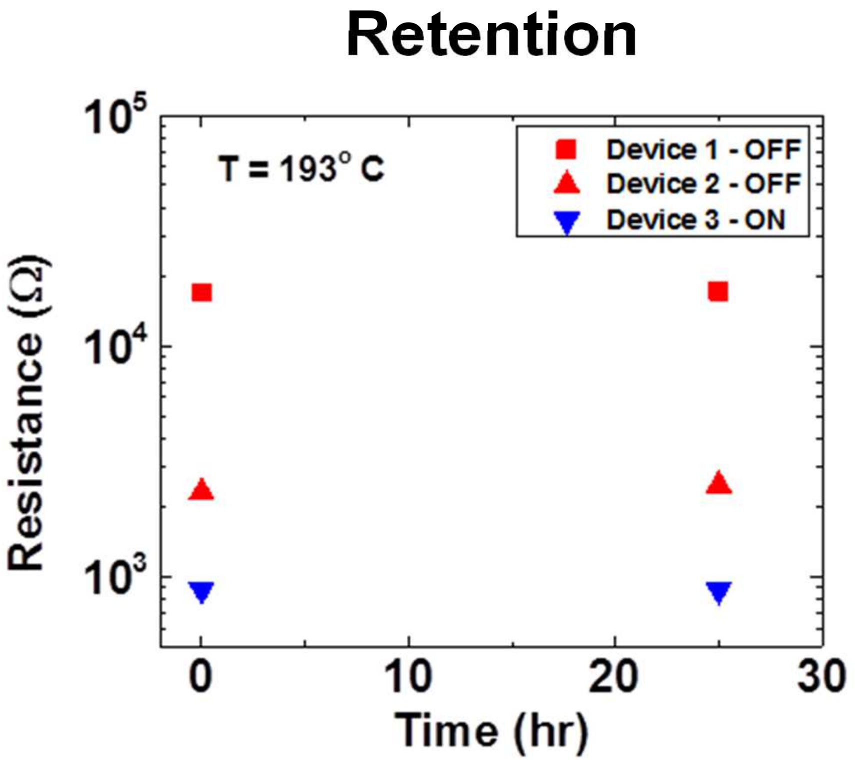


J.E. Stevens et al, accepted for publication by J. Vac Sci. Tech., 2013.

Endurance and Retention

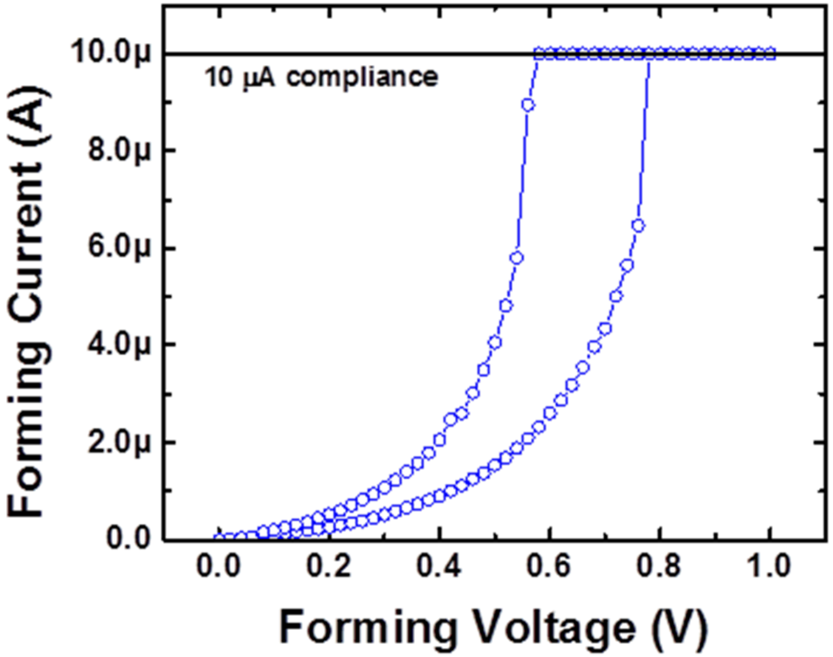
- Basic characteristics

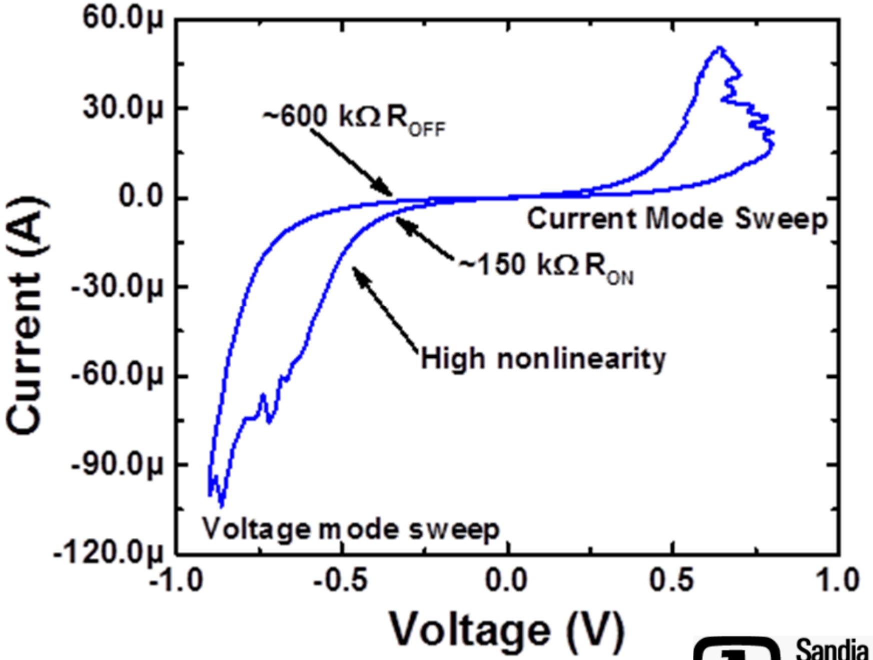




High Resistance Behavior

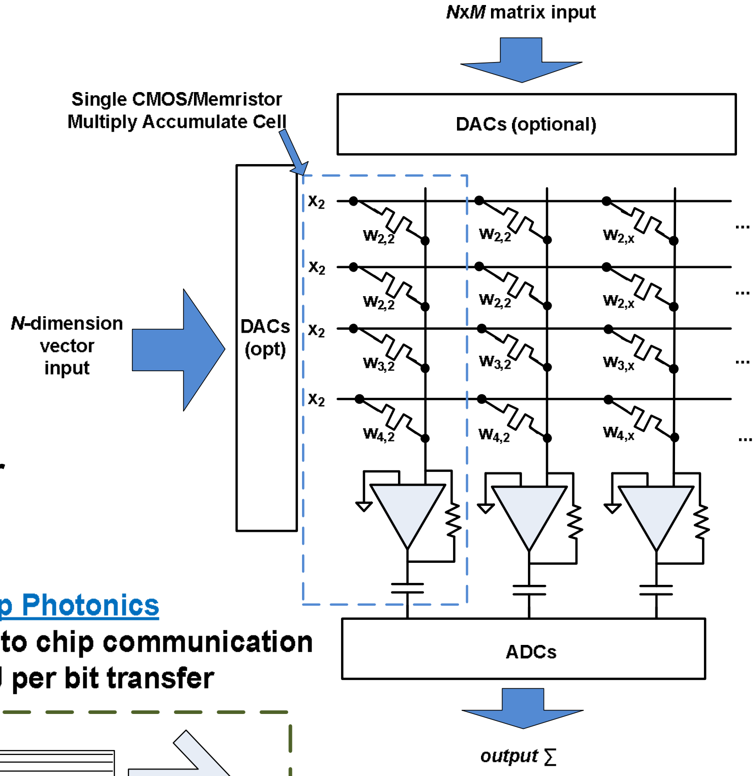
- Significant performance improvement can be achieved by careful electrical forming and control
- Power limited switching
- Very high resistance and R_{OFF}/R_{ON} possible (~100 mV read)





Analog Computing

- Vector matrix operations often comprise >> 90% of operations in pattern matching algorithms
- A monolithically integrated memristor accelerator can greatly improve power and throughput for these operations
- This could comprise a node of a future HPC system

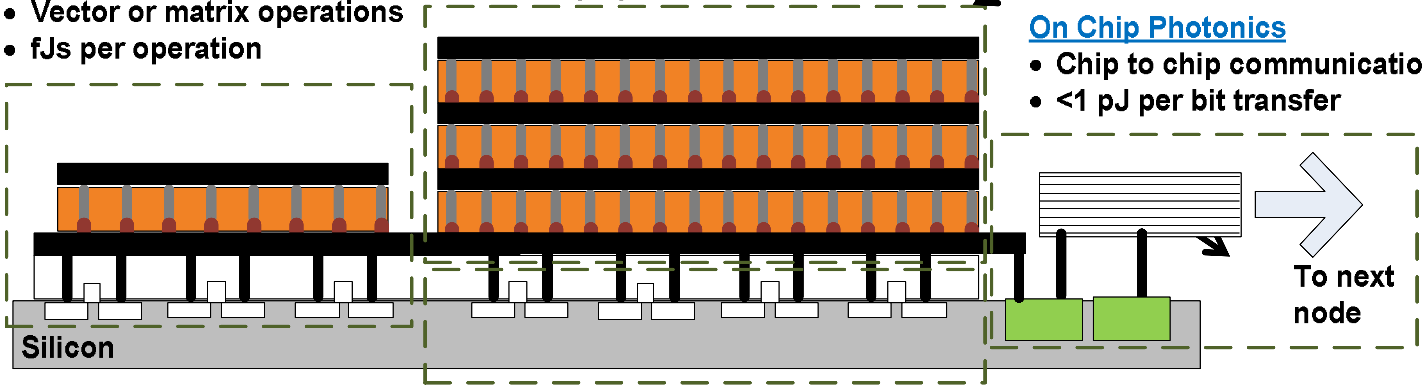


On Chip Universal Memory:
- Stacked ReRAM
- Petabit cm^{-2} Densities
- Replaces DRAM & flash
- <1 pJ per write/read

Memristor memory

On Chip Memristor Accelerator:
- Vector or matrix operations
- fJs per operation

On Chip Photonics
- Chip to chip communication
- <1 pJ per bit transfer



High Performance Logic:
- 5 nm FinFETs
- III-V on silicon