Variation-Aware, Reliability-Emphasized Design and Optimization of RRAM Using SPICE Model

H. Li[1,2], Z. Jiang[1], P. Huang[2], Y. Wu[1], H.-Y. Chen[1], B. Gao[2], X. Y. Liu[2], J. F. Kang[2#], and H.-S. P. Wong[1*]
[1]Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA
[2]Institute of Microelectronics, Peking University, Beijing 100871, China
E-mail: [#]kangjf@pku.edu.cn; [*]hspwong@stanford.edu

***Abstract*—Resistive switching random access memory (RRAM) is a leading candidate for next-generation nonvolatile and storage-class memories and monolithic integration of logic with memory interleaved in multiple layers. To meet the increasing need of device-circuit-system co-design and optimization for applications from digital memory systems to brain-inspired computing systems, a SPICE model of RRAM that can reproduce essential device physics in a circuit simulation environment is required. In this work, we develop an RRAM SPICE model that can capture all the essential device characteristics such as stochastic switching behaviors, multi-level cell, switching voltage variations, and resistance distributions. The model is verified and calibrated by a variety of electrical measurements on ~10 nm RRAMs. The model is applied to explore a wide range of applications including: 1) variation-aware design; 2) reliability-emphasized design; 3) speed-power assessment; 4) array architecture optimization; and 5) neuromorphic computing. This experimentally verified design tool not only enables system design that includes the complete suite of RRAM device features, but also provides solutions for system optimization that capitalize on device/circuit interaction.**

Keywords—emerging memory, resistive switching memory, SPICE model, design tool, variability, reliability.

I. Introduction

Emerging memory technologies such as phase change memory (PCM), spin-transfer torque magnetic random access memory (STT-MRAM) and resistive switching random access memory (RRAM) have attracted great research interests in recent years [1]-[3]. These novel memory technologies promises to bring a revolution to the memory hierarchy in computer architecture [2]. Besides, brain-inspired neuromorphic computing beyond the von-Neumann architecture may also benefit from the capabilities of emerging non-volatile memories. RRAM is a leading candidate for these applications due to its low cost, low operating power, fast switching speed, demonstrated scalability and 3D integration capability [1]-[4]. Much effort have been paid to understand the underlying physics of switching behaviors and operating mechanisms of RRAM devices [5]-[8]. Based on these studies, many features such as multi-level cell (MLC), wide distributions of switching parameters, and disturb effects have been revealed to describe the complete characteristics of RRAM. Meanwhile, the system design community is developing system-level mathematical models [9]-[11] to assess the benefits of systems that uses RRAM. Although both fields have made substantial progress, in order to bring realism to the analyses, it is necessary to bridge the gap between deep device physics and in-depth device-circuit-subsystem co-design. A variety of RRAM compact models have been developed either in a phenomenological manner or based on physics [12]-[17]. However, some models are not SPICE-compatible, and therefore cannot be implemented in industry-standard simulators such as HSPICE. And some may not include part of the essential device features such as intrinsic variations and fluctuations. The loss of memory cell 'information' may lead to inaccurate assessment of system performance and reliability. Thus, a SPICE compact model that includes all the essential features for accurate system design is needed. In this work, we use Verilog-A to develop a physics-based and experimentally verified RRAM SPICE model covering all the major features from observations (**Table I**). The model includes both MLC switching behaviors and intrinsic device variations. Then, we demonstrate the application of this model for a wide range of applications from variation-aware design to speed-power assessment and reliability-emphasized design, from array architecture optimization to neuromorphic computing.

TABLE I
MODEL FEATURE REQUIREMENTS FOR DESIGN EXPLORATIONS

	Dynamic switching	Multi-level states	Intrinsic variations	Disturb effects
Memory array operation	Required	Required	Required	Required
MLC design		Required	Required	
Variation -aware design	Required		Required	
Reliability assessment	Required			Required
Speed-power assessment	Required		Required	
Architecture optimization		Required	Required	
Neuromorphic computing	Required	Required	Required	

II. Requirements for RRAM Models

RRAM is an MIM-structure device that can be switched between low resistance state (LRS) and high resistance state (HRS) under external voltage control. For bipolar RRAM devices, the switching voltage (SV) for the SET process (from HRS to LRS) is positive and SV for RESET (from LRS to HRS) is negative. Compact models [12]-[17] serve as the interface between device properties and RRAM system design. The requirements of model features for various design targets are summarized in **Table I**. To enable model implementation in circuit simulators, SPICE compatibility is essential. In the contrast to the simplified mathematical models, device compact models should be able to reproduce the inherent switching behaviors. To capture transient responses during array

This work is supported in part by the member companies of the Stanford Non-Volatile Memory Technology Research Initiative (NMTRI) affiliate program, and Systems on Nanoscale Information Fabrics (SONIC) Center, one of six centers of Semiconductor Technology Advanced Research Network (STARnet), a Semiconductor Research Corporation (SRC) program sponsored by Microelectronics Advanced Research Corporation (MARCO) and Defense Advanced Research Projects Agency (DARPA), and the NCN-NEEDS program, which is funded by the National Science Foundation, contract 1227020-EEC, and by the Semiconductor Research Corporation. This work is also supported in part by 973 Program (2011CBA00600) and NSFC Program (61334007, 61404006). Z. Jiang is additionally supported by the M. Stanley Rundel Fellowship. H.-Y Chen is additionally supported by the Intel PhD Fellowship. H. T. Li is supported in part by the Stanford UGVR Program during the summer of 2014.

TABLE II

COMPARISON OF RRAM COMPACT MODELS

	SPICE Compatibility	Filament Dimension	HRS Variations	LRS Variations	SV Statistics	MLC Design	Current Fluctuations	Parasitic /Overshoot	Disturb Effects
Sheridan 11 [12]	Compatible	1-D	Not included	Not included	Not included	Partially supported	Not included	Not included	Not included
Ielmini 11 [13]	Not compatible	2-D	Not included	Not included	Not included	Supported	Not included	Partially included	Not included
Degraeve 12 [14]	Not compatible	2-D	Not included	Not included	Not included	Supported	Not included	Included	Not included
Huang 13 [15]	Not compatible	2-D	Not included	Not included	Not included	Supported	Not included	Included	Included
Jiang 14 [16]	Compatible	1-D	Included	Included	Not included	Partially supported	Included	Not included	Partially included
This work	Compatible	2-D	Included	Included	Included	Supported	Included	Included	Included

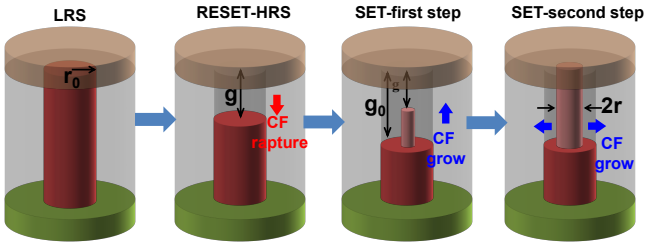

Fig. 1. Schematic of modeling conductive filament (CF) evolution processes during SET and RESET operations on RRAM cell. Key physical variables are gap distance (g) and CF radius (r).

operations, current fluctuations and parasitic effects should be included in the model. For MLC design, multi-level states with statistical distributions should be supported. For variation-aware design, a complete framework for the statistical distributions of LRS resistance (R_{LRS}), HRS resistance (R_{HRS}) and SV for SET/RESET is required. A model will not be able to reflect realistic conditions without including switching parameter distributions. For reliability assessment, the stochastic switching behavior and read/write disturb effects on unselected memory cells for a memory array should be supported to assess the status of memory cells during long-term programming. Speed-power assessment of RRAM circuits can be performed by directly simulating and monitoring array operations. At the system level, array architecture optimization and neuromorphic system design also require some key model features such as MLC, intrinsic variations, and stochastic switching. Aimed at supporting a broad portfolio of design explorations for RRAM technology, we develop a SPICE-compatible compact model of RRAM covering all the critical features [5]-[8], [12]-[14]. A comparison of RRAM compact models in terms of supported features is given in **Table II**.

III. SPICE MODEL OF RRAM

A. Model Core: Filament Evolution

It has been widely accepted that for metal-oxide RRAM, resistive switching is due to the formation and rupture of conductive filament (CF) [1], [5], [15]. The switching characteristics are strongly correlated with CF geometry, which is determined by the generation and recombination of oxygen vacancies (V_O) in the oxide layer. Based on this physical picture, we model the RRAM switching behavior as the CF evolution processes during SET and RESET, as shown in **Fig. 1**. The key control variables are the tunneling gap distance (g) and the CF radius (r), which describe a 2-D filament in the

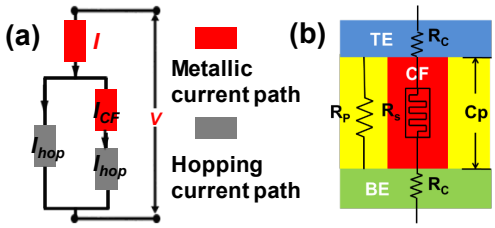

Fig. 2. (a) Schematic of conduction paths in RRAM. (b) Parasitic elements of the MIM-structure RRAM.

oxide switching layer. During the SET process, the growth rates of CF in length and in radius are described by the two-step process [15]:

$$dg/dt = af\exp\left(-(E_a - \alpha_a ZeE)/k_BT\right) \quad (1)$$

$$dr/dt = \left(\Delta r + \Delta r^2/2r\right) f\exp\left(-(E_a - \alpha_a ZeE)/k_BT\right) \quad (2)$$

where a is the adjacent V_O distance, f is the vibration frequency, E_a is the activation energy of V_O, α_a is an enhancement factor, E is the electrical filed, T is the temperature and k_B is the Boltzmann constant. During the RESET process, the rupture rate of CF is determined by the slower one of two physical processes, namely, O^{2-} release from electrode and V_O/O^{2-} recombination [15]:

$$dg/dt = af\exp\left(-\left(E_i - \gamma ZeV\right)/k_BT\right) \quad (3)$$

$$dg/dt = af\exp\left(-E_h/k_BT\right)\sinh\left(\alpha_h ZeE/k_BT\right) \quad (4)$$

where E_i and E_h are electrode/oxide interface barrier and O^{2-} hopping barrier, respectively. γ and α_h are enhancement factors. During the switching, the local CF temperature plays an important role [6] to accelerate the temperature-dependent processes incorporated in equations (1)-(4). Joule-heating alters the local CF temperature and is described by:

$$T = T_0 + IVR_{th} \quad (5)$$

where T_0 is the ambient temperature and R_{th} is the thermal resistance of the CF.

The conduction of RRAM cell is modeled based on two dominant mechanisms [7]: hopping current paths and metallic conduction paths, as shown in **Fig. 2(a)**. The I-V characteristics associated with g and r can be calculated as:

$$I_{hop} = I_0\left(\pi r^2/4\right)\exp\left(-g/g_T\right)\sinh\left(V_{gap}/V_T\right) \quad (6)$$

$$I_{CF} = \pi r^2 V_{CF}/4\rho\left(g_0 - g\right) \quad (7)$$

Next, we take into account the parasitic effects originating from electrode capacitance (C_p), contact resistance (R_C) and

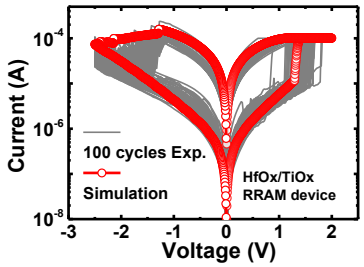

Fig. 3. Measured (100 cycles background) and simulated I-V characteristics of HfO_x/TiO_x bi-layer RRAM device.

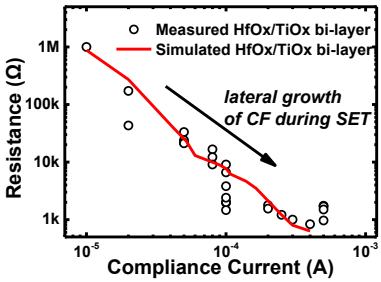

Fig. 4. Measured and simulated R_{LRS} under different SET compliance current.

leakage paths (R_p) in the oxide layer. The simplified parasitic model containing the resistive switching component is shown in **Fig. 2(b)**. Thus, transient responses can be captured.

B. Modeling Intrinsic Variability

Variability is one of the intrinsic characteristics of RRAM [5], [18]. As compared to other emerging memories, RRAM exhibits wide statistical distributions of the switching parameters such as HRS resistance (R_{HRS}), LRS resistance (R_{LRS}) and SV [5]. These device-level variations have a significant impact on system performance and reliability, and thus should be properly treated in the device model. Recent studies reveal that variations of R_{LRS} and R_{HRS} result from the fluctuations in the CF radius [18] and the tunneling gap distance [5], respectively. In addition, it is found that the stochastic nature of the ion migration barriers may lead to the variations of switching voltage [18]. Based on these physical pictures, the variations of g and r are incorporated by the following relations:

$$g = \int (dg/dt + \delta g \times \chi(t))dt \quad (8)$$

$$r = \int (dr/dt + \delta r \times \chi(t))dt \quad (9)$$

where $\chi(t)$ is zero-mean Gaussian sequence with a root mean square of unity. δ_g and δ_r represent variation amplitude to be determined based on device measurement data. Finally, energy barrier variations are included using a Monte-Carlo approach to reproduce cycle-to-cycle SV variations during SET/RESET:

$$g = S(E_a + \delta_{Ea}) \quad (10)$$

$$g = R(E_i + \delta_{Ei}, E_h + \delta_{Eh}) \quad (11)$$

C. Experimental Verification and Calibration

TiN/HfO_x/TiO_x/Pt bi-layer RRAM devices of ~10 nm feature sizes were fabricated and the detailed fabrication process was reported in [19]. To verify and calibrate the developed model, electrical measurements are performed using the Agilent 4156C parameter analyzer. The model is implemented in HSPICE simulations for comparison. **Fig. 3** shows measured and simulated I-V characteristics of RRAM.

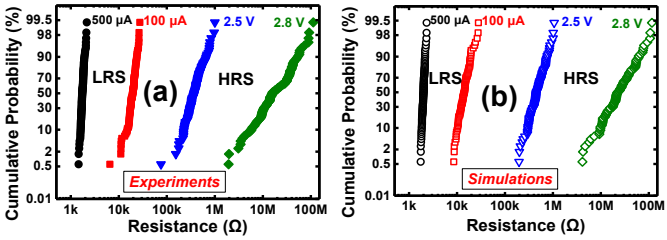

Fig. 5. (a) Measured and (b) simulated complete MLC characteristics with resistance distributions due to cycle-to-cycle variations.

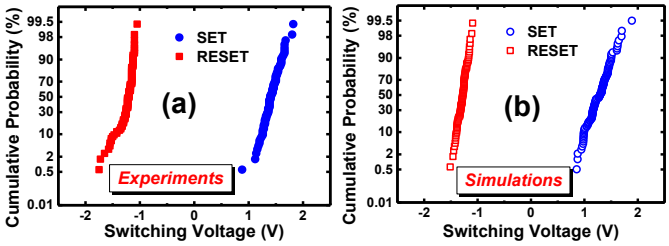

Fig. 6. (a) Measured and (b) simulated switching voltage distributions.

The model can reproduce typical 'abrupt SET' and 'gradual RESET' behaviors well. The model assumes a lateral CF growth process during SET. To verify this assumption, the dependency of R_{LRS} on compliance current during SET is measured and compared with the simulation results, as shown in **Fig. 4**. The use of a larger compliance current leads to lower R_{LRS} after SET, which implies that a wider or 'stronger' CF is formed, so consequently a lower resistance is read out. The agreement between experimental data and simulation over 4 orders of magnitude validates the 2-D filament assumption in our SPICE model. A large set of statistical experiments (100 cycles for each condition) are conducted to calibrate the model to reproduce the intrinsic variability of RRAM. As shown in **Fig. 5(a)**, varying the maximum voltage during RESET leads to different level of R_{HRS}, and different SER compliance currents result in different levels of R_{LRS}, all accompanied by resistance variations. The model reproduces MLC characteristics with variations since the CF geometry variability is incorporated into the model. The simulated statistical distributions shown in **Fig. 5(b)** agree with the experimental data well. The model supports complete MLC RRAM design and does not ignore the intrinsic parameter distributions. The stochastic properties of SV during SET/RESET can be also reproduced, as shown in **Fig. 6(a)** and **Fig. 6(b)**. The fluctuations in energy barriers lead to variation of the SET/RESET voltages at different cycles, resulting in a normal distributions for the SV. Hence, this model not only captures inherent switching behaviors, but also reproduces the statistical properties correlated with intrinsic variability of RRAM. Other essential features such as overshoot phenomenon (due to parasitic capacitance) and read/write disturb effects (sub-threshold switching) are also supported by this model. The details will be described elsewhere.

IV. Design Case Studies

After setting up the experimentally verified RRAM SPICE model, we employ the model for the study of crossbar arrays (with and without selectors) and 3D vertical RRAM (VRRAM) arrays [20]. We generate full-size array structures with interconnect wire resistance (R_{wire}) and wire capacitance (C_{wire}) at the 22 nm node (ITRS2013) for HSPICE simulations. **Fig. 7**

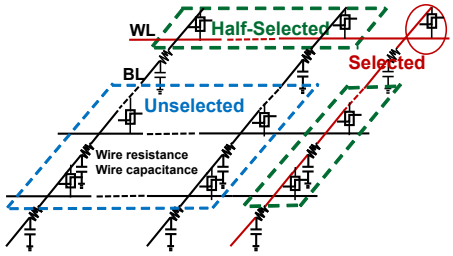

Fig. 7. Diagram of crossbar RRAM arrays with interconnect RC components.

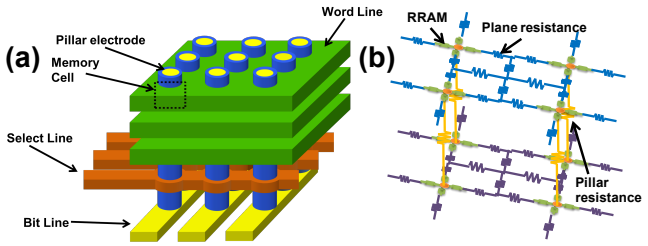

Fig. 8. (a) Schematic of 3D vertical RRAM array.
(b) Architecture of sub-circuit model for 3D VRRAM in HSPICE simulations.

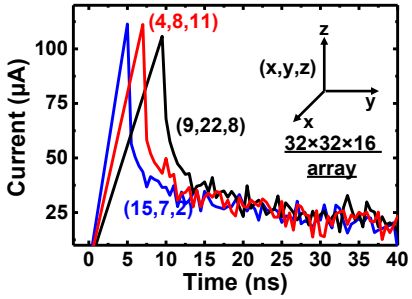

Fig. 9. Transient current responses during random-access programing of a 16-kb 3D VRRAM array. The cell location is indicated by (x,y,z) in the legend.

shows the diagram of crossbar RRAM arrays formed of word lines (WL), bit lines (BL) and RRAM cells at cross-points. A worst-case scenario is considered where the selected cell is located at the farthest corner in the array. The write schemes follow the widely used 'V/2' and 'V/3' schemes [9], and voltage-driver sensing scheme is used for read operation [9]. The model is also employed in 3D VRRAM simulations. **Fig. 8(a)** shows the schematic of 3D VRRAM array [20]. We use the sub-circuit model [20] to construct the full-size 3D VRRAM arrays in HSPICE simulations, where the plane WLs and vertical pillars are modeled as resistor network shown in **Fig. 8(b)**. Plane-to-plane, pillar-to-pillar and RRAM capacitors are all considered in array structure. An experimentally validated write/read scheme is used in the simulations [21].

A. Variation-Aware Design

First, we present the application of the developed model in variation-aware design, which includes the impact of intrinsic variability of RRAM devices for circuit design and assessment. Random-access write operation is simulated in a 16-kb (32SLs × 32BLs × 16 layers) 3D VRRAM array, as shown in **Fig. 9**. Pulse operation (4 V, 50-ns width, 5-ns pulse edge) is used to program RRAM cells located at different address in the 16-kb memory block. Cells at different address suffer from different RC delay. And the variability of device filament growth during the nanoseconds-long programming time window can lead to significant RESET current fluctuations. Simplified models may not be able to monitor the dynamic status of RRAM cells

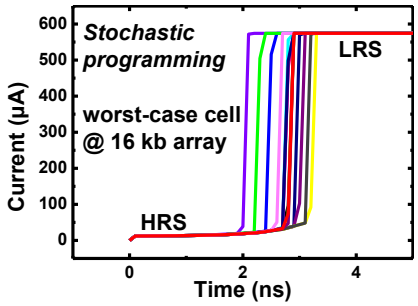

Fig. 10. Stochastic programming on the worst-case cell in a 16-kb crossbar RRAM array.

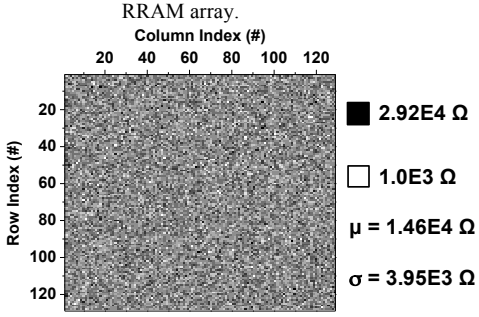

Fig. 11. Random data pattern of a 16-kb (128×128) RRAM array.

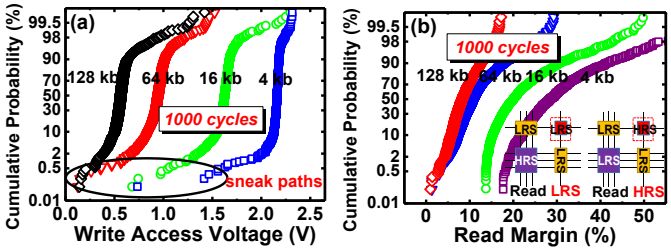

Fig. 12. Statistical distributions of (a) write access voltage and (b) read margin with array data pattern randomness. Inset of (b): For the worst cases of reading LRS and HRS, the unselected cells are HRS and LRS, respectively.

during programming. To design RRAM arrays with better stability, current fluctuations [22] during write operations must be taken into consideration.

We then focus on the stochastic programming behavior of one single cell. A 16-kb (128WLs × 128BLs) crossbar RRAM array is simulated, and the 20-cycle SET operations on the worst-case cell are shown in **Fig. 10**. It is noticed that the SET time varies from cycle to cycle. Hence, for fast-speed operation of RRAM circuits, the stochastic properties of switching voltages should be carefully considered to leave sufficient margin for successful write operations. It is known that crossbar RRAM arrays suffer from sneak path problems. Considering device resistance variations, a random data pattern is generated as shown in **Fig. 11**, which is closer to real chip conditions. This implies that it is not sufficient to simply calculate the sneak currents and write/read margin using simple worst-case scenario, since the resistance map that results from specific data pattern may lead to complex sneak path configurations. The statistical distributions of write access voltage (V_{access}) and read margin (RM) are further simulated from 1000 random data patterns of different array sizes, as shown in **Fig. 12**. V_{access} is the voltage drop on the farthest selected cell in the array, and RM is defined as the output voltage swing ratio upon reading LRS and HRS. During the write operations, the random data patterns that result from resistance variations have a significant impact on the sneak

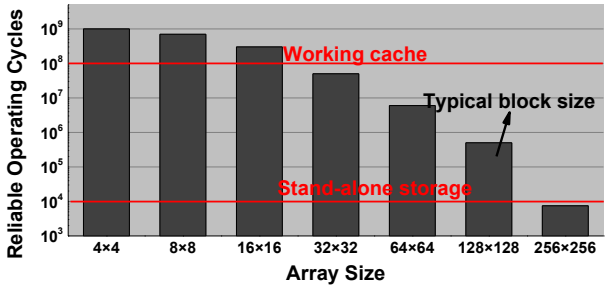

Fig. 13. Reliable operating cycles as a function of array size.

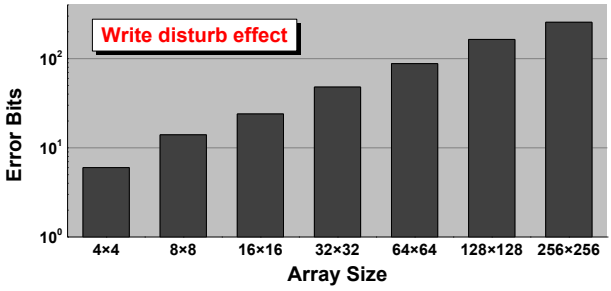

Fig. 14. Error bits after write disturb as a function of array size.

path configuration and the IR drop along the selected signal lines. Thus, both V_{access} and RM show wide spreads. The spread of read margin involves both R_{LRS} and R_{HRS} variations since the worst-case data patterns for reading LRS and HRS are different, as indicated in the inset of **Fig. 12(b)**. The wide distributions of V_{access} and RM present challenges for the design of peripheral I/O circuits, and therefore, must be considered for device-circuit co-design. Improving the robustness of peripheral circuitry may help reduce the write failure probability that results from the device-level variations. To meet the reliability specifications (ppm and ppb level), the device-level tuning of memory cells should accompany the circuit-level design and system-level optimization.

B. Reliability-Emphasized Design

Reliability is one of the major concerns for RRAM technology. For some applications such as working cache, it is critical to design the system emphasizing reliability metrics. Here we assess write disturb effects on crossbar RRAM arrays as memory blocks ranging from 4×4 to 256×256. 'Disturb' on RRAM originates from accumulated sub-threshold switching which exhibits a stochastic behavior. Under V/2 bias scheme, those half-selected cells (located along selected WL/BL) will be biased with around half V_{dd}. To guarantee a successful write operation on the farthest selected cell, a higher V_{dd} may be required to compensate for the interconnect IR drop. In this case, half-selected cells face even worse disturb issues. This can bring severe reliability challenges for circuit design. Successive pulse programming on memory arrays may disturb the status of the half-selected cells and generate error bits in the memory block. By simulating consecutive pulse operations, the reliable operating cycles for the memory block can be obtained, as shown in **Fig. 13**. The decreasing trend of reliable operating cycles with larger array size is due to the increase in array V_{dd} necessary for a successful write operation. For larger arrays, half-selected cells are under more voltage stress in V/2 scheme. This is a trade-off between reliability and sneak path problems. As is well-known, higher R_{LRS} and R_{HRS} benefit overall

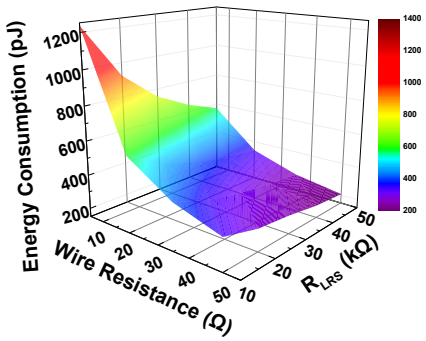

Fig. 15. Write energy consumption as a function of R_{LRS} and interconnect R_{wire}.

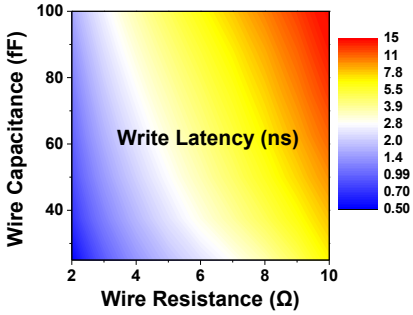

Fig. 16. Write latency as a function of interconnect R_{wire} and C_{wire}.

reliability since sneak paths can be alleviated [23]. The design options here can be either tuning device properties for better disturb immunity below SV, or choosing an optimal memory block configuration. For instance, for high-performance, close-to-CPU applications such as working cache, the unit memory block size should be reduced. For stand-alone data storage applications which typically sees less programming cycles, memory block size can be increased to reduce area overhead and communication latency among blocks. Disturb on half-selected cells takes place among a group of adjacent cells. Therefore, after write disturb there will be some error bits in the array, as shown in **Fig. 14**. The error bit number increases for larger arrays. The results of **Fig. 13** and **Fig. 14** clearly shows that RRAM device characteristics must improve significantly before RRAM can be used in practical situations where large block sizes and small error bit rate are needed. In addition, embedding error correction codes in memory blocks and implementing refreshing schemes are needed for memory arrays for meeting the requirements of various system applications.

C. Speed-Power Assessment

The developed model can be used for assessing circuit performance by directly simulating write/read transient operations in HSPICE. As shown in **Fig. 15**, the write energy of a 16-kb array is simulated as a function of both R_{wire} and R_{LRS}. With increasing R_{wire}, the energy consumption of the array decreases. Also, increasing R_{LRS} can reduce energy consumption due to smaller leakage current of the whole array. Hence, higher cell resistance is beneficial and should be one of the targets for device engineering and tuning. However, it is also shown that for a larger R_{wire}, we may lose the benefit of increasing R_{LRS}. Since larger R_{wire} leads to larger interconnect IR drop, it is not desired and should be optimized by interconnect engineering. **Fig. 16** is an assessment of interconnect scaling impact on 16-kb RRAM circuit

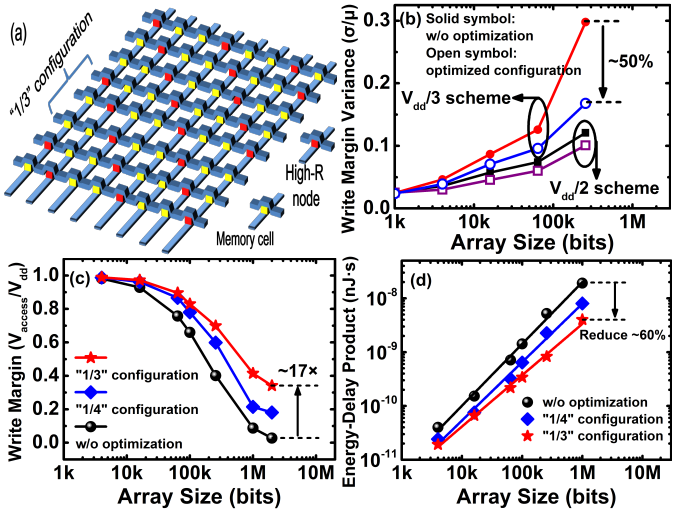

Fig. 17. (a) An optimized crossbar configuration. (b) Comparison of variation-immunity with and without array optimization under different bias schemes. (c) Write functionality and (d) EDP assessments with and without

performance. Although the device switching speed can reach sub-ns level [1], the interconnect RC delay sets a major limitation to the overall programming speed of RRAM circuits.

D. Array Architecture Optimization

The SPICE model can be used to evaluate new array architectures. Since conventional crossbar structure suffers from severe sneak path issues, here we study an alternative array configuration. In this configuration, uniformly distributed insulating nodes are built in the array, as illustrated in **Fig. 17(a)**, to reduce sneak-path leakage. **Fig. 17(b)** shows that the alternative configuration is effective in reducing write margin (WM) variance (σ/μ) especially for the V/3 bias scheme, since the distributed insulating network can 'block' sneak current flowing in the array. The proportion of insulating nodes as compared to memory nodes in the array can be adjusted to meet circuit specifications, and the optimized configuration method is effective for Mb-level arrays. As shown in **Fig. 17(c)** and **Fig. 17(d)**, under '1/3' configuration, the WM can be improved by 17 times for a 2-Mb array, and as a key metrics the total energy-delay-product (EDP) can be reduced by 60%. With a specific WM or EDP criterion, the maximum achievable array size can be enlarged even considering the capacity sacrifice due to insulating nodes.

E. Neuromorphic Computing Application

RRAM synaptic devices can be used to build neuromorphic visual systems [24]-[25]. To emulate the biological synapses, the devices should exhibit plasticity, i.e., the conductance can gradually change according to the input stimuli. **Fig. 18** shows the gradual resistance modulation using consecutive pulses to gradually RESET RRAM. Cycle-to-cycle variations of resistance can be observed, which originates from intrinsic variability of RRAM even below switching threshold. The model can also reproduce the resistance distributions after training processes. The simulated statistical distributions are compared with the experimental data in **Fig. 19(a)** and **Fig. 19(b)**. The tail bits observed in both experiments and simulations indicate the training failure events. The model can be further exploited to study neuromorphic computing systems.

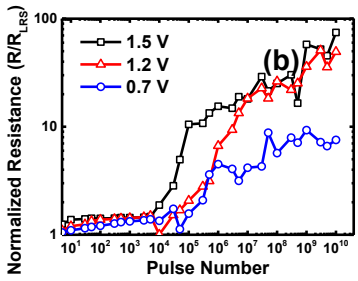

Fig. 18. Simulated gradual RESET training processes with variations using consecutive pulse operations under different pulse heights.

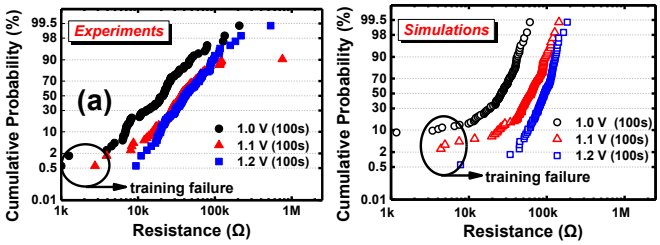

Fig. 19. (a) Measured and (b) simulated RRAM synapse resistance distributions after 100-second DC stress training.

V. Conclusion

A SPICE model of RRAM with essential features such as filament-based resistive switching, intrinsic variability, switching parameter distributions and MLC capability is developed and verified by a set of experimental data for ~10 nm device sizes. A wide range of design aspects focusing on variability, reliability, speed-power performance, array architecture, and neuromorphic computing are illustrated. The interplay between device and circuit design are highlighted. The SPICE model is a useful tool for system designers for performance assessment and optimization. This work paves the way towards device-circuit-system co-design for RRAM technology.

VI. Acknowledgement

We would like to thank G. Hills and Prof. S. Mitra for the fruitful discussions, and thank Prof. S. Yu and X. Guan for the earlier development of RRAM compact models.

References

[1] H.-S. P. Wong *et al.*, *Proc. IEEE*, vol. 100, no. 6, pp. 1951–1970, 2012.
[2] H. Li *et al.*, *Proc. DATE,* 2009, pp. 731-736.
[3] G. W. Burr *et al.*, *IBM J. Res. Dev.*, vol. 52, no. 4.5, pp. 449-464, July 2008.
[4] M. S. Ebrahimi *et al., IEEE S3S Conference*, Millbrae, CA, 2014
[5] S. Yu *et al.*, *Proc. IEDM*, 2011, pp. 413-416.
[6] U. Russo *et al.*, *Trans. Electron Devices*, vol. 52, no. 2, pp. 193-200, 2009.
[7] S. Yu *et al.*, *Appl. Phys. Lett.*, 99, 063507, 2011.
[8] B. Gao *et al.*, *Proc. IEDM,* 2011, pp. 417-420.
[9] C. Xu *et al.*, *Proc. DATE*, 2011, pp. 1-6.
[10] C. Xu *et al.*, *Proc. DAC*, 2013, pp. 1-6.
[11] D. Niu *et al., Proc. ICCAD*, 2013, pp. 17-23.
[12] P. Sheridan *et al., Nanoscale,* vol. 3, no. 9, pp. 3833-3840, 2011.
[13] D. Ielmini *et al., Proc. IEDM*, 2011, pp. 409-412.
[14] R. Degraeve *et al.*, *VLSI Technology*, 2012, pp. 75-76.
[15] P. Huang *et al., Trans. Electron Devices*, vol. 60, no. 12, pp. 4090, 2013
[16] Z. Jiang *et al.*, *SISPAD*, 2014, pp. 41-44
[17] H. Li *et al., Electron Device Lett.*, vol. 35, no. 2, pp. 211-213, 2014.
[18] N. Raghavan *et al.*, *Proc. IEDM*, 2013, pp. 554-557.
[19] Y. Wu *et al., Proc. IEDM*, 2013, pp. 550-553.
[20] H.-Y. Chen *et al., Proc. IEDM,* 2012, pp. 497-550.
[21] B. Gao *et al., Trans. Electron Devices,* vol. 61, no. 5, pp. 1377-1381, 2014.
[22] S. Yu *et al., Trans. Electron Devices*, vol. 59, no. 4, pp. 1183-1188, 2012.
[23] J. Liang *et al.*, *JETC*, vol. 9, no. 1, pp. 9:1-9:14, 2013.
[24] S. Yu *et al., Adv. Mat.*, vol. 25, no. 12, pp. 1774–1779, 2013.
[25] B. Gao *et al.*, *ACS Nano*, 8(7), pp. 6998-7004, 2014.