Utilizing *I-V* Non-Linearity and Analog State Variations in ReRAM-Based Security Primitives

G.C. Adam[1,3*&], H. Nili[1*†], J. Kim[2], B. D. Hoskins[1], O. Kavehei[2], D.B. Strukov[1]
[1] UC Santa Barbara, Santa Barbara, CA 93106, USA
[2] RMIT University, Melbourne, Victoria 3000, Australia
[3] National Institute for R&D in Microtechnologies, Bucharest 077190, Romania
[&]gina_adam@engineering.ucsb.edu, [†]hnili@ece.ucsb.edu

***Abstract*— The underlying variability in the ReRAM device operation, while undesired in many applications, can be advantageous for hardware security primitives. ReRAM devices also come with the advantage of having non-linear multi-state operation. By comparison with previous reported ReRAM PUFs, which utilized spatial variations in the devices' binary ON/OFF states, we proposed to use sneak path currents and device / network nonlinearity as its main source of randomness to implement robust, reconfigurable and dense security primitives. In particular, in this work we present an in-depth discussion of how device non-linearity is affected by the read bias and the thermal stresses applied to the ReRAM crossbar. For the experimental demonstration, we used a three-dimensional stack of two 10x10 Al_2O_3/TiO_{2-x} –based ReRAM crossbars with good uniformity for the memristors in both crossbar layers. The results highlight the utility of device non-linearity to extract more complex and more reliable one-way functions from relatively small ReRAM crossbar arrays.**

Keywords—ReRAM, PUF, security, metal-oxide memristor, passive crossbar circuits

I. INTRODUCTION

Hardware-based security primitives have recently attracted significant attention due to their potential to be physically embedded with their cryptographic data thanks to unique and unpredictable structural features. Physically Unclonable Functions (PUFs) are an example of such cryptographic primitives and represent the hardware implementation of a mathematical one-way function that maps an input (challenge) to an ideally unique and unpredictable output (response). PUFs put to use the otherwise disadvantageous device variabilities that naturally arise during the manufacturing process [1-2]. A PUF should ideally be unclonable against a wide range of adversarial attacks [3-4], but also have a stable and reliable operation. Unfortunately, PUF implementations using traditional CMOS technology have been shown to be susceptible to a range of modeling attacks related to their need to reprogram the volatile SRAM memory [5].

ReRAM crossbar arrays are a promising emerging technology for PUF implementation, due to their simple and relatively low-cost manufacturing, scalability, compatibility with CMOS monolithic integration and most importantly, their stochastic process-induced variations in their *I-V* characteristics. ReRAM devices are based on a mixed electronic-ionic transport, utilizing amorphous materials as the switching medium. These amorphous materials are prone to have compositional inhomogeneities, defects, grain boundaries, etc., which together with the inherent imperfections of the electrodes all can lead to variability in the device behavior [6-7].

Previously reported ReRAM-based PUFs utilize the spatial (device-to-device) variations of high and/or low resistance states (HRS/LRS) to construct a unique key [8-12]. In this approach, the ReRAM crossbar is reduced to a static resistive network, needing a large crossbar size for a high reliability. Since the average distributions of HRS and LRS are typically constant across the array, such method is not suitable for implementing multiple different PUF instances.

We proposed to use ReRAM device nonlinearity as the source of intrinsic variations to implement PUFs with increased security metrics [13] (Fig. 1a). Our approach utilizes a crossbar with non-linear devices that can be tuned to a resistive state anywhere in their dynamic range. The variations in the individual device conductances and their respective non-linearities are shown to be ubiquitous toolbox with a high degree of complexity for constructing robust PUF architectures that are difficult to model and predict [14].

This work focuses on the impact of read bias and of the thermal stress to the device non-linearity variation across the ReRAM crossbar. The results show that by using device non-linearity variability as a source of randomness, more complex and more reliable one-way functions can be obtained from relatively small ReRAM crossbar arrays. Since in this approach the ReRAM state can be tuned to any resistance in the range, it is important to choose a median array state that can satisfy both the requirement of high network nonlinearity and the one of minimum power requirements. Higher resistance states typically satisfy both these requirements.

In particular, in order to mitigate the effect of state variations on the PUF response, the devices need to be tuned randomly to a narrow pseudo-normal distribution (Fig. 1b). These device states need to be randomly distributed across the array so that the average column and row resistances are uniformly distributed (Fig. 1c). In this work, we investigated how the read bias affects the apparent non-linearity distribution across the ReRAM crossbar and how stable the state maps are with respect to applied thermal stresses.

* These authors contributed equally to this work.

This work was supported by the AFOSR under the MURI grant FA9550-12-1-0038, by NSF CCF-1528250 grant, by the Australian Research council under ARC DP140103448 grant and by the U.S. Department of State under the International Fulbright Science and Technology Award.

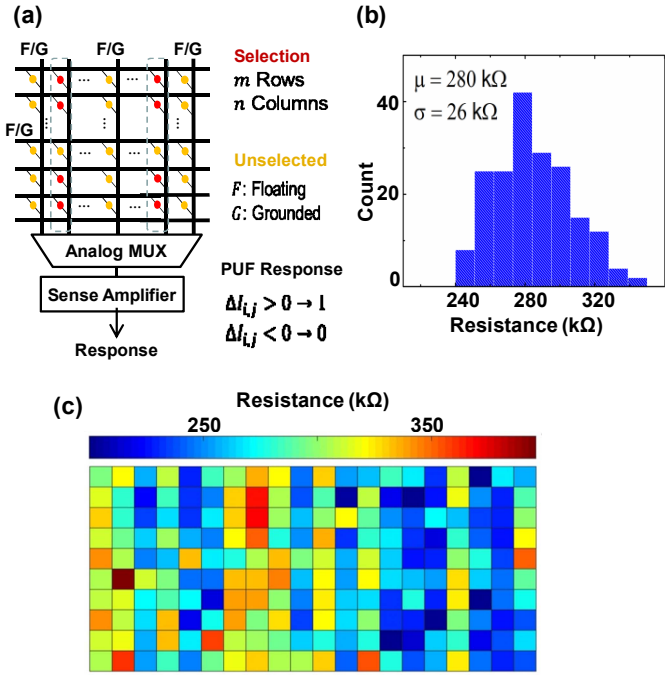

Fig. 1. Proposed ReRAM-based PUF (a) Security primitive selection scheme (b) Tuned resistance pseudo-normal distribution (R_0 = 280 ± 60 KΩ measured 200 mV) (c) Spatial distribution of resistances across the 2x10x10 ReRAM crossbar.

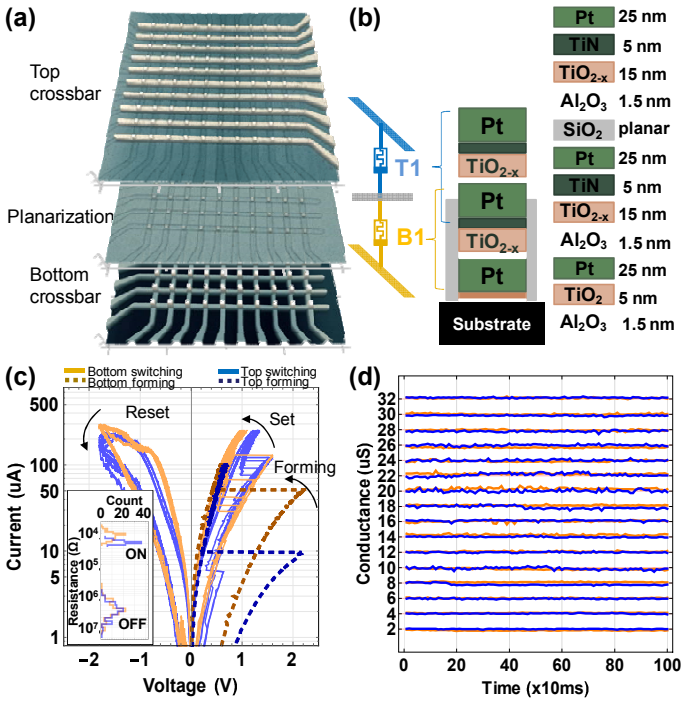

Fig. 2. Fabrication details: (a) AFM images of the two-layer crossbar, and (b) Cartoon of its cross-section showing the material layers and their corresponding thicknesses. (c) Typical I-V forming and switching characteristics for bottom and top devices (inset shows ON and OFF resistance distributions) (d) Bottom and top device analog tuning to 16 states (between 2μS and 32μS).

II. Device Fabrication and Characterization

Two fully passive TiO_{2-x} 10×10 memristor crossbars with an active device area of ~350-nm×350-nm were monolithically integrated in a conventional (horizontal) configuration using an in-situ low temperature reactive sputtering deposition and a precise planarization step. The middle electrodes are shared between the bottom and top layers (Fig. 2a and b). These crossbars have a similar fabrication process and electrical parameters as the one described in [15].

The detailed electrical characterization reveals very good uniformity and analog programmability of the fabricated 3D crossbar circuits. All electrical testing was performed at room temperature, unless otherwise specified, with an Agilent B1500A semiconductor device parameter analyzer and an Agilent B1530A waveform generator / fast measurement unit.

The I-V characteristics of the bottom and their corresponding top devices measured in a floating configuration in a small 2×2 portion of the crossbar show good similarity in device behavior between the bottom and top devices. An ON/OFF ratio of ~10 was observed when a conservative reset voltage of -1.8 V and a set current of 300 μA were used (Fig. 2c). The ON/OFF ratio can be further increased to ~100 when a more aggressive reset voltage of -2.4 V is used (Fig. 2c inset).

The devices in both layers show good analog tunability (Fig. 2d) to 16 clearly distinguishable states equally spaced in the 2 μS – 32 μS range. The devices were tuned to 1% precision using the tuning algorithm presented in [16] with 500 μS long voltage pulses of maximum amplitudes ±2.6 V and a step of ±0.01 V.

The variations in effective voltage switching thresholds are sufficiently low for a precise tuning of the devices within array, while still very substantial to be utilized in the considered application – see below. The device I-V is nonlinear, especially at higher resistance states (Fig. 2d).

III. Device Non-Linearity in PUF Operation

We investigated the device non-linearity for all the 200 devices in the fabricated stacked crossbar, together with its variation to the read voltage and the applied temperature. The nonlinearity factor was calculated as:

$$NLF = |1 - G_0/G(V_B)| \times 100\% \quad (1)$$

In Fig. 1a, all the 200 devices were tuned to R_0 = 280 ± 60 KΩ at 200 mV, using the write-verify algorithm. The resistances of these devices show a Gaussian distribution. Their non-linearity factors have a narrower distribution at lower read voltages (300 mV) by comparison with higher read voltages (600 mV).

In Fig. 3a, one can notice how at higher voltages, the nonlinearity changes the conductance map in its totality. Then the spatial distribution of the device nonlinearity over the crossbar is calculated according to Eq. 1. The maps show, as expected, higher nonlinearity at higher voltages as compared to lower biases. The relative nonlinearities of the devices at higher voltages are very different despite the fact the resistance state are so close. These variations in device non-linearity are due to individual device variations which are heavily

predicated on the nanometric parameters such as local interfacial states, roughness and doping variations (vacancies, passivation centers, etc.). Therefore even knowing the nominal linear resistances and a working model of average device nonlinearity does not lead total knowledge about the dynamics. This in effect means that nonlinearity can be used both as an additional space of entropy and as a tool for tuning the security by trading off higher operational power.

The larger entropy space at higher voltages will result in more robust adaptive circuits or security primitives. De facto, the operation at higher bias means operating a different crossbar both in terms of absolute conductance value and relative distribution of the map. So the same instance and distribution can be re-purposed for different functions (e.g. unique security primitives) by just tuning the operational bias.

As it can be observed from Fig. 3b, the variability at higher biases means a wider distribution of currents (both absolute and relative measures). Assuming we have a detection limit of 50 nA, more current differentials at 600 mV are above the detection limit than the differential at 200 mV. This means less chance for errors and consequently, this leads to a higher confidence margin for different "differential" operations.

Since the PUF security hardware that we proposed needs tight resistance distributions, it is important for the crossbar system to maintain a similar response independent of the temperature stress. In terms of the effect of the temperature on the device resistance, Fig. 4 shows that the crossbar has higher temperature swing resilience at higher biases. Firstly, the crossbar was tuned to 280 kΩ±60 kΩ as shown in Fig. 1b, after tuning having a conductance map at 200 mV as shown in Fig. 1c. The tuned crossbar was then stressed thermally to 90°C for 3 hours. After the stress heating, the conductance maps at 200 mV and at 600 mV were measured. As can be seen from Figs. 4a and 4b, there is a smaller change in rest resistance of individual states at 600 mV by comparison with the 200 mV. This smaller change at higher biases means lower susceptibility to thermally stress induced changes in operational characteristics and consequently, a more stable and reliable operation.

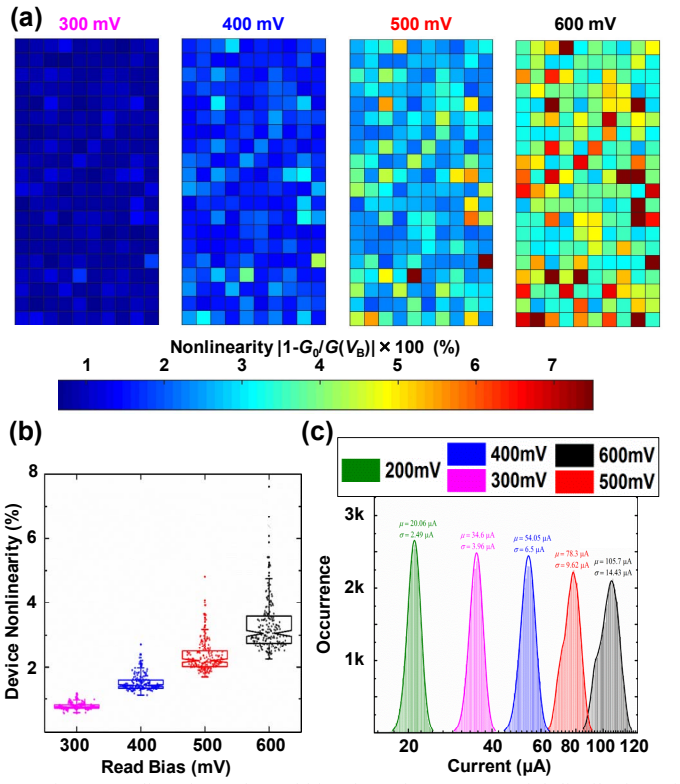

Fig. 3. Nonlinearity and Read bias dependence. (a) Spatial distribution of nonlinearity over the crossbar - calculated at the difference between conductance at V_{bias} and linear conductance at 200 mV normalized to the linear conductance. All devices are tuned to a state between 250-300 kΩ with very tight distribution.(b) Nonlinearity distribution showing a bigger relative space for higher biases; (c) Distribution of READ currents for a selection of (m = 5) rows and reading one column multiple times while the rest of the array is floating. The aggregate number of READ operations for each distribution (at each bias) is 128k. At higher biases, a wider current distribution is observed.

IV. Implications for PUF Design

A prototype PUF network was implemented in the 2x10x10 ReRAM stacked crossbar structure, by taking advantage of the variations in the devices' nonlinearity and state tuning capabilities. Each device in the stacked crossbar was tuned to the specific pre-calculated values required by the PUF using the write-verify algorithm. The details of the implementation can be found elsewhere [13]. This first experimental demonstration showed good security metrics, such as uniformity, diffuseness and uniqueness.

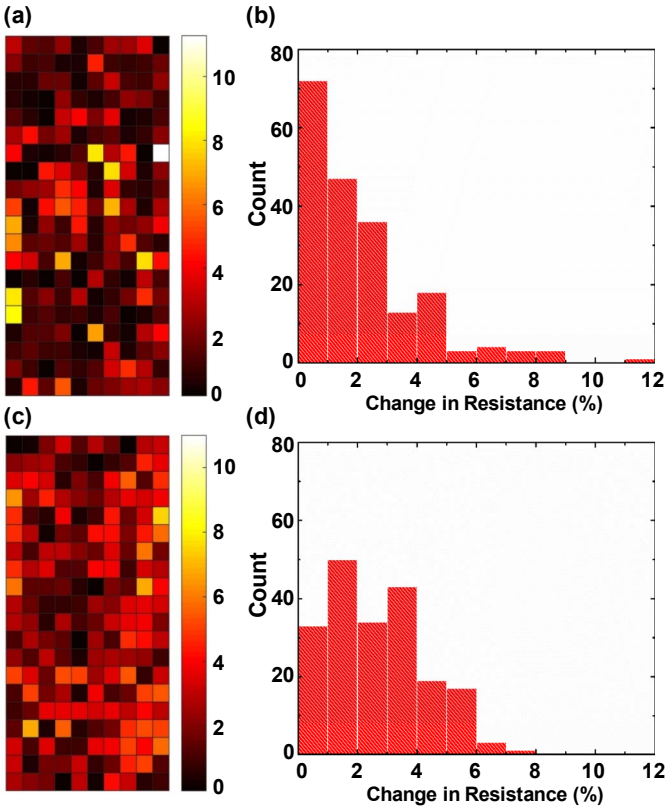

Fig. 4. Nonlinearity and resilience to temperature stresses dependence. (a-b) Absolute change in resistance map (a) and distribution (b) at 200 mV after heating the tuned stacked crossbar to 90°C for 3 hours; (c-d) Absolute change in resistance map (c) and distribution (d) at 600 mV after heating the tuned stacked crossbar to 90°C for 3 hours. The changes are measured by comparison with the initial conductance map and distribution at 200mV (Fig.1) where all devices are tuned to a state between 280±60 kΩ.

As expected, increasing by voltage bias from 200 mV to 600 mV improves the uniformity from already decent 49.5 ±6.25% to nearly ideal 50.1 ± 6.26%, while another PUF randomness metric, diffuseness is also close to ideal ~50 ± 6.25% for all voltage cases.

As discussed in the previous section, the better performance PUF at higher voltages can be attributed to the stronger nonlinearity in the device *I-V*s at higher voltages. The proposed PUF shows robust functional performance passing the NIST randomness test, while having the additional advantages of reconfigurability, high integration density of ReRAM stacked crossbars and their suitability for monolithical integration onto CMOS chips.

V. Discussion and Summary

The demonstrated utility of conductance state nonlinearity in enriching the randomness (entropy) space of ReRAM-based PUF architectures and improving the overall network resilience to environmental stress (e.g. temperature variations) highlights the promise of such higher order variations in design and implementation of hardware-intrinsic security primitives. Thus employing such variations along with the first order variations in conductance states across the ReRAM array can enable ubiquitous design for secure and robust security primitives.

The experimental results were obtained on relatively large 350×350 nm devices. However, through aggressive scaling it would be possible to achieve very dense and complex PUF primitives. Previous studies have shown similar performance metal-oxide memristors with dimensions around 8 nm to 10 nm [17, 18], though apparently utilizing slightly different nanometer scale filamentary switching mechanism [19, 20].

In summary, we have experimentally investigated the impact of device non-linearity and analog state variations on a ReRAM implemented PUF system. The three-dimensionally stacked crossbar used for the experiment showed low enough variation as required for the tuning, but high enough for the PUF entropy. Our results show that higher read biases are useful for a higher entropy space and a more stable operation with respect to thermal stresses. We believe that this work is an important step toward ReRAM-based robust implementation of the PUFs. In addition, there are many reservations for improvement, e.g. security metrics can be improved by employing multi-level responses [13].

Acknowledgment

We acknowledge useful discussions with B. Thibeault, M. Prezioso and F. Merrikh-Bayat. This work was funded through the NSF grant CCF-1528205 and through the Marie Sklodowska-Curie grant No. 705957.

All device fabrication was performed in the UCSB cleanroom facility, which is part of the NSF-funded National Nanotechnology Infrastructure Network.

References

[1] B. Gassend *et al*., "Silicon physical random functions," in: *Proc. ACM CCS'02*, 2002, pp. 148-160.

[2] J. Guajardo *et al*., "FPGA intrinsic PUFs and their use for IP protection," in: *Proc. CHES'07*, 2007, pp. 63-80.

[3] U. Rührmair *et al.*, "Modeling attacks on physical unclonable functions," in: *ACM CCS'10*, 2010, pp. 237-249.

[4] U. Rührmair, J. Sölter, F. Sehnke, X. Xu, A. Mahmoud, V. Stoyanova, *et al.*, "PUF modeling attacks on simulated and silicon data," *IEEE Transactions on Information Forensics and Security,* vol. 8, pp. 1876-1891, 2013.

[5] M. Cortez *et al.*, "Modeling SRAM start-up behavior for Physical Unclonable Functions," in: *Proc. DFTS'12*, 2012, pp. 1-6.

[6] R. Waser, R. Dittmann, G. Staikov, and K. Szot, "Redox-based resistive switching memories- nanoionic mechanisms, prospects, and challenges", *Advanced Materials,* vol. 21, pp. 2632-2663, 2009.

[7] J. J. Yang, D. B. Strukov, and D. R. Stewart, "Memristive devices for computing", *Nature Nanotechnology,* vol. 8, pp. 13-24, 2013.

[8] P.-Y. Chen *et al.*, "Exploiting resistive cross-point array for compact design of physical unclonable function," in: *Proc. HOST'15,* 2015, pp. 26-31.

[9] R. Liu *et al.*, "Experimental characterization of physical unclonable function based on 1 Kb resistive random access memory arrays," *IEEE Electron Device Letters,* vol. 36, pp. 1380-1383, 2015.

[10] Y. Gao *et al.*, "Obfuscated challenge-response: A secure lightweight authentication mechanism for PUF-based pervasive devices," in: *Proc. PerCom'16*, 2016, pp. 1-6.

[11] Y. Gao *et al.*, "Memristive crypto primitive for building highly secure physical unclonable functions", *Scientific reports,* vol. 5, 2015.

[12] L. Gao, P.-Y. Chen, R. Liu, and S. Yu, "Physical unclonable function exploiting sneak paths in resistive cross-point array", *IEEE Transactions on Electron Devices ,* vol 63, pp. 3109-3115, 2016.

[13] H. Nili *et al.* "Highly-Secure Physically Unclonable Cryptographic Primitives Using Nonlinear Conductance and Analog State Tuning in Memristive Crossbar Arrays", *arXiv preprint arXiv:1611.07946*, 2016.

[14] J. Kim, *et al.* "A Physical unclonable function with redox-based nanoionic resistive memory", *arXiv:1611.04665,* 2016.

[15] G.C. Adam, *et al.* 3-D memristor crossbars for analog and neuromorphic computing applications. *IEEE Transactions on Electron Devices* vol. 64, pp. 312-318, 2017.

[16] F. Alibart, L. Gao, B.D. Hoskins, and D.B. Strukov, "High precision tuning of state for memristive devices by adaptable variation-tolerant algorithm", *Nanotechnology ,* vol. 23, 075201, 2012.

[17] B. Govoreanu *et al.*, "10×10 nm^2 Hf/HfO_x crossbar resistive RAM with excellent performance, reliability and low-energy operation", in: *Proc. IEDM'11*, Washington, DC, Dec. 2011, pp. 31.6.1-31.6.4.

[18] S. Pi, P. Lin, Q. Xia, "Cross point arrays of 8 nm× 8 nm memristive devices fabricated with nanoimprint lithography", *Journal of Vacuum Science & Technology B*, *Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena,* vol. 31, no. 6, 06FA02, 2013.

[19] J. P. Strachan *et al*., "Direct identification of the conducting channels in a functioning memristive device", *Advanced Materials*, vol. 22, pp. 3573-3577, 2010.

[20] D.-H. Kwon *et al*., "Atomic structure of conducting nanofilaments in TiO_2 resistive switching memory", *Nature Nanotechnology*, vol. 5, pp. 148-153, 2010.