A Calibration-Free In-Memory True Random Number Generator Using Voltage-Controlled MRAM

Jiyue Yang[1]*, Di Wu[1,2]*, Albert Lee[1], Seyed Armin Razavi[1], Puneet Gupta[1], Kang L. Wang[1], Sudhakar Pamarti[1]
Electrical and Computer Engineering, University of California, Los Angeles[1], and
Inston, Inc.[2], Los Angeles, CA, 90095, USA
Email: jyang669@ucla.edu. *equal contribution.

Abstract—In this paper, we propose an in-memory True Random Number Generator (TRNG) using Voltage-Controlled MRAM that doesn't require calibration of the writing pulse's width and amplitude. Previous solution using Spin Transfer Torque (STT) MRAM requires calibration for every MTJ, thus making the multi-row random number generation inside the memory impossible. We also propose a 100% relative throughput digital bias correction circuit that doesn't degrade bit rate. The VC- MTJs are fabricated in CMOS BEOL compatible process with an 80 nm diameter and high TMR ratio of 160%. MRAM array circuits and bias correction circuits are fabricated in 65 nm CMOS technology and wire-bonded with the VC-MTJ devices. Multiple VC-MTJs are tested and shown to pass all NIST randomness tests.

Keywords— Cryptography, hardware security, MRAM, true random number generator, integrated circuits.

I. Introduction

True Random number generators (TRNG) are key components in cryptography applications. With the advent of the quantum computing, many of the traditional cryptography algorithms may be impaired or broken in a reasonable number of tries. To prevent security problems in the post-quantum cryptography, much longer keys are required [1]. This requires the TRNG hardware to have higher throughput and lower energy cost. Previous works have demonstrated hardware random number generators in CMOS technology using inverter's metastability [2], jitter in a ring oscillator [3] and transistor's oxide breakdown [4]. Most of them require dedicated circuit with large area and power-consuming post processing circuits to remove bias. The memory based TRNG reduces the cost of energy and area by reusing the memory array to generate random numbers. Previous works have explored TRNG based on STT-MRAM [5][6], which exploits metastability in a current controlled spin-transfer-torque (STT) MRAM. However, the switching probability of the TRNG is highly sensitive to the amplitude and duration of the critical current. Given inevitable device variability, extensive calibration may be required to find qualified devices. Besides, the STT MRAM suffers from large energy consumption and limited endurance due to large write current.

To overcome those issues, we propose an in-memory TRNG using Voltage-Controller MRAM that does not require calibration of the write pulse. It improves energy consumption and endurance by having 50× larger resistance area (RA) product than STT-MRAM. Furthermore, a new bias correction digital circuit that has 100% relative throughput is proposed to ensure high speed and robust randomness under potential magnetic field interference.

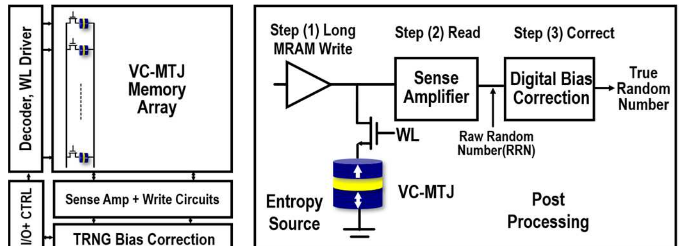


Fig. 1. *(a) System architecture (b) Block diagram of TRNG operation*

II. In-Memory TRNG System Architecture

The in-memory TRNG based on VC-MRAM uses the same MRAM array for both the storage and true random number generation. Figure.1(a) shows the architecture of the in-memory TRNG system. When used as a memory, the wordline driver turns on the access transistors on one row that is decoded using the address input. The access circuit reads or write the bitcells on the same wordline. Fig.1 b) shows the operation of the system under the in-memory TRNG mode. The wordline drivers of multiple rows inside the array can turn on at the same time. A long voltage pulse (~10ns) is applied between each bitline and source line pair. The VC-MTJs generate the random numbers in parallel and store the outputs in each bitcell. The sense amplifiers then read the raw random numbers from the bitcells row by row and pass them to the post-processing circuits to remove any potential bias. Since the same array auxiliary circuits are used in the TRNG mode and memory mode, a significant area save is achieved.

To demonstrate our idea, the MRAM access circuit and the bias correction circuit are designed and fabricated in 65nm CMOS technology. A column of VC-MTJ devices are fabricated on another die and connected with CMOS chip by wire bonds. The rest of the paper is organized as this: section III introduce the background of the VC-MTJ devices, section IV introduces the mechanism of random number generation, section V and VI introduce the circuit design of array access and bias correction circuits, and section VII presents the measurement results.

III. Background of Voltage-Controlled MTJ

Voltage-control MRAM has been proposed as a promising candidate to replace STT MRAM to dramatically improve

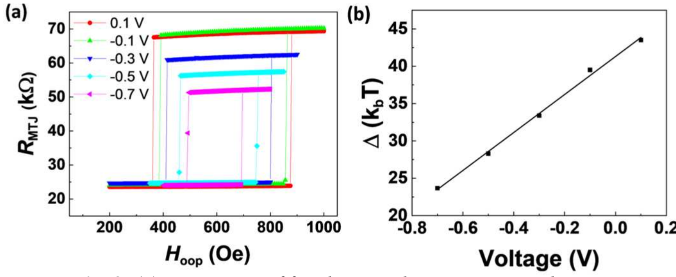


Fig. 2. *(a) R-H curves of free layer with various gate voltages. (b)Voltage dependence of thermal stability factor (Δ).*

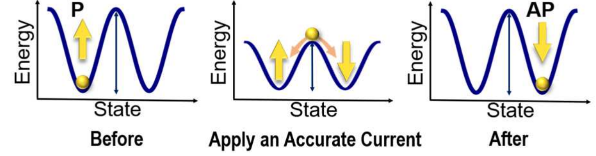
write energy and latency [7]. The STT MRAM suffers from long write time (~20ns) and high write energy (~pJ). These become the limitations in the application of high-speed and low-power cache that needs frequent and low power write operations. Voltage-controlled MRAM dramatically reduce the switching time (~1ns) and write energy(~fJ). During write, a voltage or electric field lowers the interfacial perpendicular anisotropy (PMA) of the free layer, thus reducing the writing energy barrier. An in-plane bias field causes the free layer to start precession and moves between P and AP state until the damping force aligns the free layer's magnetization to the in-plane direction. To enhance the PMA effect, the MgO barrier is made much thicker than the STT MRAM and results in much large resistance area (RA) product and lower write current.

In this work, we have fabricated an array of VC-MTJs on an 8-inch silicon wafer. The film stack consists of a CoFeB free layer, CoFeB/W/Co fixed layer, and MgO tunneling barrier. The MTJ pillars are patterned to the diameter of 80nm after 400°C annealing. Due to the thicker MgO layer, the RA of 200 Ω*um^2 is achieved, which is ~50x larger than the STT-MRAM [8]. An external magnetic bias field is given to use as a processional axis and compensate for the fixed layer's stray field. Hysteresis loop for the free layer at various bias voltages are shown in Fig.2(a). Anti-parallel state's resistance of 65K ohm and TMR of 160% are achieved. The voltage-controller magnetic anisotropy (VCMA) can be extrapolated from the thermal stability vs voltage curve [9], shown in Fig.2(b). The VCMA coefficient of ~41.5fJ/Vm is measured and corresponds to a write voltage of 1.4V.

IV. Using VC-MRAM as Calibration-Free TRNG

The VC-MTJ has the unique property of converging to metastability asymptotically without the requirement of any calibration, making it a perfect solution of generating true random numbers inside the memory. A comparison of the random number generation mechanism between STT-MRAM and VC-MRAM is shown in Fig.3. STT-MRAM rely on a critical current that has a pre-determined amplitude and pulse width to achieve metastability. High resolution timing control of the write pulse for each device during the operation or calibration ahead of the operation is needed to achieve high entropy. In contrary, VC-MRAM does not require calibration before or during the operation. When a voltage is applied, the free layer's magnetization precesses along the in-plane axis.

STT-MRAM: Metastability @ One Critical Current & Pulse Width



VC-MRAM: Asymptotically Converges to Metastability

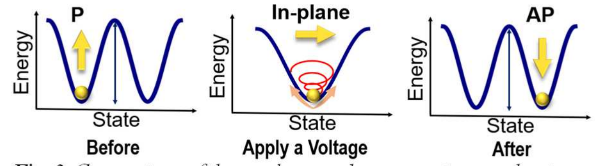


Fig. 3. *Comparison of the random number generation mechanism between STT and VC MRAM.*

It oscillates between P and AP state and asymptotically converges to the in-plane axis due to damping effect. The longer the voltage pulse is, the closer it is aligned to the in-plane direction, which corresponds to 50% probability. After the voltage pulse is removed, the free layer's magnetization is randomly switched to P or AP state under the influence of the thermal noise. The measured switching probability approaches 50% after the voltage pulse is applied 3n second, as shown in Fig.4(a).

VC-MRAM makes the in-memory true random number generation of multiple rows possible when high throughput is required. The array arrangement of the memory makes only one row of the bitcells available for read or write at a time. Since STT-MRAM requires a calibration for each individual device inside the memory, the multi-row in-memory operation is not possible to achieve. However, since VC-MTJ does not need calibration, devices in multiple rows can generate random numbers at the same time. In a multi-row operation, several wordlines turn on together, as shown in Fig.4(b). The bitcells on the same bitline share the same write pulse. A longer pulse can make sure that most of the bitcells generate high-entropy random bits under device variations. A post processing circuit can remove any potential bias during read out. The multi-row in-memory RNG can significantly increase the throughput with no extra hardware cost.

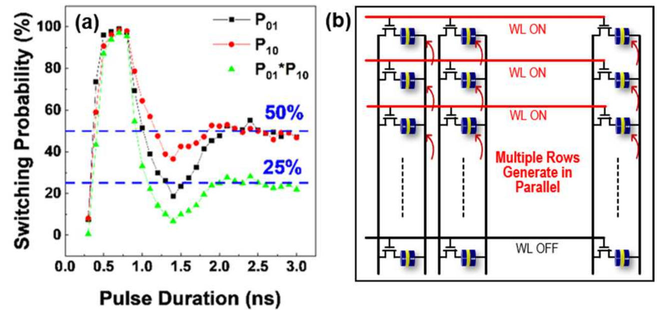


Fig. 4. *(a) VC-MTJ switching probability vs pulse duration. (a) Multi-row random number generation inside the memory.*

V. VC-MTJ Read and Write Circuitry

To support the high-speed access of the VC-MRAM array, a high-speed current sense amplifier is designed. Fig.5. shows the implementation details of the current sense

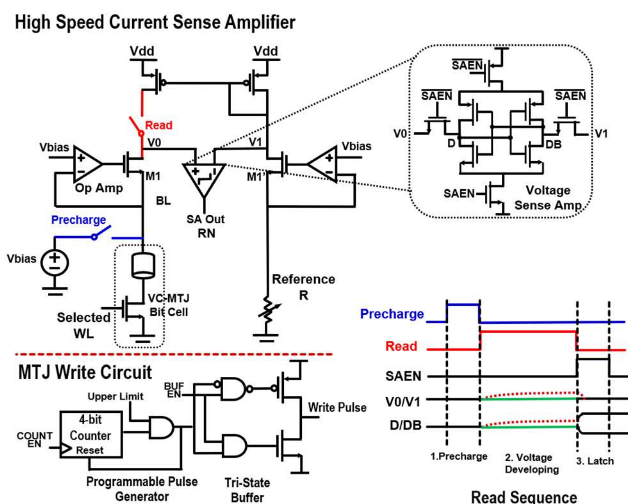


Fig. 5. *VC-MTJ read and write circuit and timing diagram.*

amplifier (CSA) and write circuits. The conventional voltage-based sense amplifier compares the voltages between the MTJ's bitline and a reference bitline. However, the speed of sensing is limited by the large RC constant of the biltine parasitic capacitance and the MTJ resistance. The CSA can achieve faster speed because the small input impedance connected to the bitline significantly lower the parallel resistance between MTJ and sense amplifier's input impedance. A feedback amplifier regulates the bitline voltage and further reduces the SA's input impedance by the loop gain. The difference between the MTJ and reference's current is amplified at the read node V0 that has much smaller capacitance. The sequence of read operation is shown in Fig.5. During the precharge phase, the bitline is pre-charged to a bias voltage. Then during the voltage-developing phase, the bitline voltage is kept constant at Vbias by the regulator. When the read switch is turned on, the current from the selected MTJ cell is passed to the read node V0. The PFET connected to V0 copies the current from a reference branch that sets the current at the middle of the MTJ's P and AP state. The currents between the MTJ and reference cell are subtracted and amplified at the end of voltage developing phase. In the latch phase, a voltage comparator compares V0 and V1 and latches the result to VDD or GND. A fast 5nsec sensing speed is achieved in the presence of large (~pF) bitline capacitance due to the small time constant at V0.

Figure 5 also shows the write circuitry for the MTJ. A programmable pulse width generator allows pulse duration to vary from 1x to 16x clock period. To reduce series resistance on the write path, a tri-state buffer is implemented to allow high impedance while the write buffer is inactive, but a very low impedance when the buffer is operational.

VI. 100% Throughput Bias Correction Circuits

Traditional Von Neumann post processing can achieve perfect bias removal but suffers from dropping a significant amount of input bits. We propose a 100% relative throughput bias corrector that can preserve the input data rate and achieves a wide correction range at the same time. Although

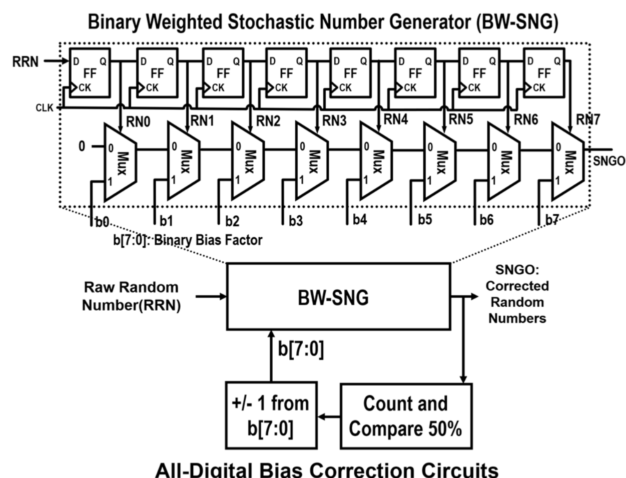
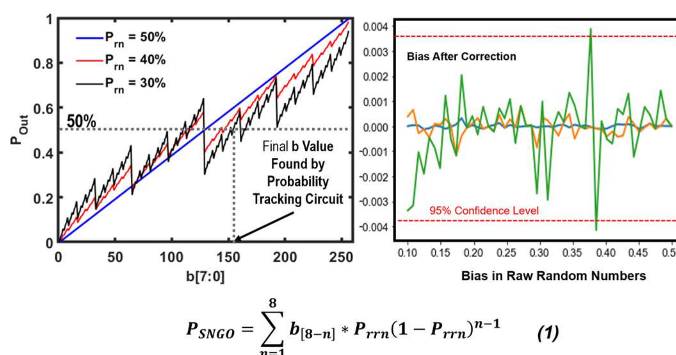


Fig. 6. *100% Throughput Bias Corrector using BW-SNG*

the ideal switching probability after the metastable state is 50%, multiple sources can cause a probability bias. For example, the stray field from the fixed layer may not be completely compensated by the external field. Malicious attackers may also apply an interference magnetic field externally to disturb the TRNG. The effective residue bias field can cause the free layer to prefer a state, and thus causing a probability bias in the random number output.

The bias correction circuit is shown in Fig. 6. The core of the probability tracking circuit is a Binary Weighted Stochastic Number Generator (BW-SNG) [10]. At every cycle, a raw random number is shifted into the buffer and produce one correction output from BW-SNG, therefore maintaining the input data rate. Every 8-bit raw random numbers are used as the select signals for a chain of 8 mutiplexers. The 0-selected input signal is the output from the previous multiplexer and the 1-selected input signal is an 8-bit binary number. The output of the BW-SNG is a stochastic bit stream with probability correlated with the 8-bit binary number, b. Assuming a stream of Independent and Identically Distributed (I.I.D.) the output probability of the BW-SNG's output bit stream is represented in Equation.1. The probability tracking circuit counts the number of 1s in 256 output bits in real-time and compare the probability with 50%. Binary number b is added or reduced by 1 based on the comparison



$$P_{SNGO} = \sum_{n=1}^{8} b_{[8-n]} * P_{rrn}(1 - P_{rrn})^{n-1} \qquad (1)$$

Fig. 7. Left *(a)PSNGO vs b plot,* ***Right*** *(b) Bias after correction using 8-bit(Green), 10-bit(Yellow) and 12-bit(Blue) BW-SNG.* ***Equation.1.*** *BW-SNG's output prob.*

result until the output probability is close to 50%. The SGNO probability vs b under different RNN probabilities is shown in Fig.7(a). Starting from 0.5, the bias corrector will automatically find the closest solution of b that results in an unbiased stream. The bias after the correction is under 0.3% across 10%~50% input RRN bias range. Better bias-removing result of < 1e-5 can be achieved by using larger BW-SNG, as shown in yellow and blue curve in Fig.7(b).

VII. MEASUREMENT AND EVALUATION

The VC-MRAM based TRNG is demonstrated by having the MRAM array circuitry fabricated in 65nm CMOS technology and VC-MTJ devices fabricated on another die. Eight VC-MTJ devices are connected in a column and wire-bonded with the CMOS chip, shown in Fig.7(a).

Multiple VC-MTJs have been tested with the CMOS chip. For each VC-MTJ, we collected multiple 2M-bit random number streams for the NIST 800-22 randomness test. All VC-MTJs pass the NIST test. Table.1 shows a summary of the NIST test results. To test the probability tracking circuit, we deliberately create a probability bias in the raw random numbers under a biased external field. Fig.7(b) shows the probability of the corrected random numbers and the binary weighted number b vs correction iterations during the probability tracking process. The raw random numbers are biased at 40% at the beginning. As the binary weighted number starts to increase, the output probability increases correspondingly and reaches a stable state of 50% probability within the first 20 cycles. Several different bias fields are also tested to show that probability tracking circuit can work in a wide range of raw random number bias probabilities. Fig.7(c) shows the autocorrelation function of the three corrected random bit streams biased at different starting probabilities. They are all bounded within the threshold assuming that the bit streams are white noise with 95% confidence level.

	NIST 800-22 Test	DEVICE 1		DEVICE 2	
		Pass Rate	χ^2 of P-Value	Pass Rate	χ^2 of P-Value
1	Frequency	100%	0.018	100%	0.740
2	Block Frequency	100%	0.006	100%	0.013
3	Cumulative Sum	100%	0.035	100%	0.210
4	Runs	100%	0.534	100%	0.99
5	Longest Run	100%	0.637	100%	0.350
6	Rank	95%	0.909	100%	0.534
7	FFT	100%	0.740	100%	0.911
8	Nonoverlap Template	PASS	PASS	PASS	PASS
9	Overlap Template	100%	0.091	100%	0.637
10	Universal	95%	0.066	100%	0.350
11	Approximate Entropy	100%	0.740	95%	0.911
12	Random Excursion	PASS	PASS	PASS	PASS
13	Random Excursion Variant	PASS	PASS	PASS	PASS
14	Serial	95%	0.350	95%	0.091
15	Linear Complexity	100%	0.440	100%	0.534

1) Pass rate >90% and P-Value > 0.0001 is considered random.
2) Pass means that the tests of all subcategories are passed

Table 1. *NIST Results of multiple MTJs.*

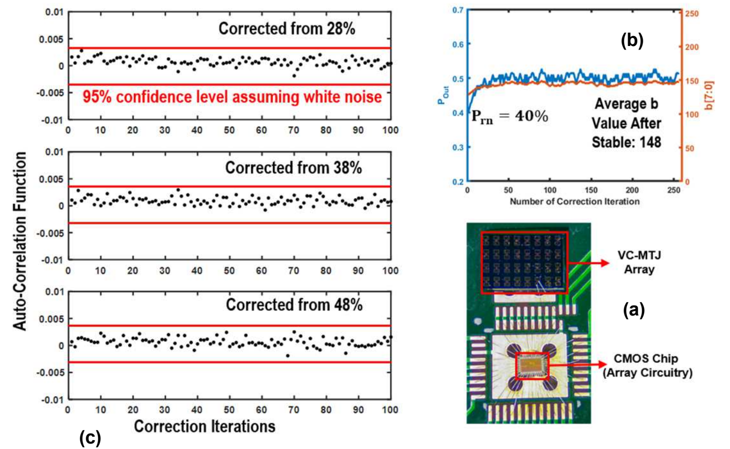


Fig. 7. *(a) CMOS and device die photo, (b)Correction output probability & b vs cycles, (a)Autocorrelation function of streams corrected from multiple*

V. CONCLUSION

In this work, we demonstrate a TRNG using a VC-MRAM that doesn't need calibration for write pulse. The dynamic properties of VC-MTJ is investigated. We also show a 100% throughput bias digital bias correction circuit that can correct from a wide input bias range. Furthermore, the potential of using VC-MRAM as a in-memory random number generator is explored as a high performance solution.

ACKNOWLEDGEMENTS

This work is in part supported by AFRL, DARPA under agreement number FA8650-18-2-7867 and NSF Translational Applications of Nanoscale Multiferroic Systems (TANMS) program. We'd like to thank H. Hosoya, Y., Nagamine, K. Tsunekawa from Canon-Anelva for MTJ fabrication.

REFERENCES

[1] Bernstein, D., Lange, T. Post-quantum cryptography. *Nature* **549,** 188–194 (2017).

[2] S. K. Mathew et al., "2.4 Gbps, 7 mW All-Digital PVT-Variation Tolerant True Random Number Generator for 45 nm CMOS High-Performance Microprocessors," in IEEE Journal of Solid-State Circuits, vol. 47, no. 11, pp. 2807-2821, Nov. 2012.

[3] K. Yang, D. Blaauw and D. Sylvester, "An All-Digital Edge Racing True Random Number Generator Robust Against PVT Variations," in IEEE Journal of Solid-State Circuits, vol. 51, no. 4, pp. 1022-1031, April 2016.

[4] N. Liu, N. Pinckney, S. Hanson, D. Sylvester and D. Blaauw, "A true random number generator using time-dependent dielectric breakdown," 2011 Symposium on VLSI Circuits - Digest of Technical Papers, Kyoto, Japan, 2011, pp. 216-217.

[5] Won Ho Choi et al., "A Magnetic Tunnel Junction based True Random Number Generator with conditional perturb and real-time output probability tracking," 2014 IEEE International Electron Devices Meeting, San Francisco, CA, USA, 2014, pp. 12.5.1-12.5.4.

[6] K. Yang et al., "A 28NM Integrated True Random Number Generator Harvesting Entropy from MRAM," 2018 IEEE Symposium on VLSI Circuits, Honolulu, HI, USA, 2018, pp. 171-172.

[7] J. G. Alzate et al., "Voltage-induced switching of nanoscale magnetic tunnel junctions," 2012 International Electron Devices Meeting, San Francisco, CA, USA, 2012, pp. 29.5.1-29.5.4.

[8] S. Sakhare et al., "Enablement of STT-MRAM as last level cache for the high performance computing domain at the 5nm node," 2018 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2018, pp. 18.3.1-18.3.4.

[9] C. Grezes, et.al., "In-plane magnetic field effect on switching voltage and thermal stability in electric-field-controlled perpendicular magnetic tunnel junctions" AIP Advances 6, 075014, 2016.

[10] P.K. Gupta and R. Kumaresan. *IEEE Transactions on Acoustics, Speech and Signal processing.*