True random number generator based on switching probability of volatile Ge_XSe_{1-X} ovonic threshold switching selectors

Z. Chai[1,2], P. Freitas[1], W. Zhang*[1], J. F. Zhang[1], J. Marsland[1]

[1]School of Engineering, Liverpool John Moores University, UK, [2]Xi'an Jiaotong University, China

*Corresponding author, email: w.zhang@ljmu.ac.uk

Abstract

Ge_xSe_{1-x} Ovonic Threshold Switching (OTS) selector is a promising candidate to suppress the sneak current paths in resistive switching memory arrays. A novel method is developed to quantitatively characterize the variations in the threshold voltage (Vth), the hold voltage (Vhd), and the switching probability dependence on the OTS operation conditions. The time-to-switch-on/off (t_{on}/t_{off}) at a constant V_{OTS} follows the Weibull distribution, based on which the dependence of switching probability on pulse waveform, bias, and time can be extracted and extrapolated. Based on this analysis, a novel technique for true random number generator (TRNG) application is proposed. The inherent variability in OTS threshold voltage results in a bimodal distribution of on/off states which can be easily converted into digital bits. The experimental evaluation shows that the proposed TRNG enables the generation of high-quality random bits that passed 12 tests in the NIST statistical test suite without complex external circuits for post-processing.

1. Introduction

True random number generators (TRNGs) are essential in many applications such as communication systems, statistical sampling, computer simulation and cryptography systems. Unlike the software-based pseudo-RNGs, hardware-based true RNGs use local physical phenomena to produce truly random outputs, which cannot be replicated or predicted externally, and are particularly critical in hardware security applications. Existing CMOS-based TRNGs based on randomness sources such as thermal noise, random telegraph noise (RTN), current fluctuation in oxide after soft breakdown, time-dependent oxide breakdown (TDDB), etc. suffer from complex circuitry for randomness extraction [1]. Emerging memory devices are promising for applications such as storage class memory and neuromorphic computing. Several TRNGs based on emerging memory devices utilized RTN, cycle-to-cycle variability or the stochastic delay time of switching as the randomness source, in which the randomness extraction can be greatly simplified by reading the random "0" or "1" bits directly from the on/off state [2-3]. However, many emerging devices suffer from limited endurance or throughput which hinders their TRNG application [4-5].

Selector devices are critical in resistive-switching random-access memory arrays for suppressing the sneak path currents, which limit the maximum array size and reliability (Fig.1a). Chalcogenide material based ovonic threshold switching (OTS) selectors have high on-state drive current (> 10 MA/cm^2), good half-bias non-linearity, fast switching speed, and excellent endurance when compared with other selectors [6-9]. Particularly, the endurance of Ge_xSe_{1-x} based OTS device is more than 10^{10} cycles by using a simple recovery scheme. After composition optimization, Se-Ge-As-Te OTS has achieved an excellent endurance of more than 10^{11} cycles without recovery operation. The statistical analysis of the OTS switching voltage and switching time and their correlation is critically important for choosing suitable operation conditions for the 1S1R structure. Recent efforts in tuning the material composition and process to improve the performance makes it necessary to develop an experimental characterization method for quantitatively evaluating the switching probability and its dependence on the operation conditions, which is the essential information required for designing the 1S1R array.

We developed a novel method to record the switch-on and -off events either at a constant bias or during the rise/fall pulse edge. The time-to-switch-on/-off (t_{on}/t_{off}) at a constant bias is found following the Weibull distribution. The statistical switching probability is dependent on both the pulse bias and time. This stochastic nature of OTS switching can be exploited to implement TRNGs. The volatile nature of OTS makes reset operations unnecessary, simplifying the operation conditions and improving the generation frequency. We demonstrate a scalable TRNG using GeSe-based OTS selector with volatile switching. The 10,000-bit random stream generated from a single device without complex extraction circuitry demonstrates good randomness by passing the National Institute of Standards and Technology (NIST) statistical tests and the machine learning prediction test based on long short-term memory (LSTM) architecture, proving that OTS selector is a promising candidate for TRNG application.

2. Devices and Experiments

Amorphous Ge_xSe_{1-x} films are prepared by room temperature physical vapor deposition (PVD). TiN/GeSe/TiN selector devices were integrated in a 300 mm process flow, using a pillar (TiN) bottom electrode which defines the device size down to 50 nm. A Ge_xSe_{1-x} chalcogenide films control from 20 nm down to 5 nm thickness was achieved and passivated with a low-temperature BEOL process scheme, as shown in Fig.1c. Three different waveforms have been developed in this work: (1) A triangle pulse to record the I-V, Vth and Vhd during switching, as shown in Fig. 1b. (2) A constant bias square pulse to record the time-to-switch-on, as shown in Fig. 2a. (3) A more complex waveform consisting a triangular segment to firstly switch on the device, followed by a constant bias segment to record the switch-off events, as shown in Fig. 2c. In this work, the device size is 65 nm, the GexSe1-x thickness is 10 nm and x = 0.4. The fast I-V characterization in this work is carried out with a Keysight B1530A Waveform Generator/Fast Measurement Unit (WGFMU) embedded in a B1500A semiconductor analyzer.
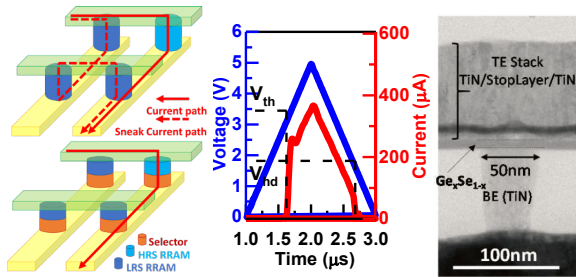

Fig.1. (a) Schematic of selector-embedded 1S1R array for suppressing the sneak current path; (b) A typical I-V of a triangular switching pulse. (c) TEM of the OTS selector used in this work.

3. Switching probability

Square pulses are usually used to program the memory and selector devices. In Fig. 2b&2d, an OTS device is switched on and off by a square pulse for 100 cycles and the I-t waveforms are measured and plotted together. It is observed that the time-to-switch-on (t_{on}) recorded at the constant pulse top bias of 2.8 V spreads over a wide time range and that the t_{on} follows the Weibull distribution, as shown in Fig. 3a. Further data points can be extrapolated using the Weibull plot, as also shown in Fig.3a. It is found that t_{on} spreads over 5 orders of magnitude, and t_{on} gradually decreases as V_{OTS} increases. t_{on} at different biases follows the Weibull distribution with the same slope, shifting towards shorter t_{on} at higher V_{OTS}. The dependence of t_{on} on V_{OTS} is clearly shown in Fig. 3b. A

99.7% switching-on probability, equivalent to the 3σ of the normal distribution (t_{on},3σ), can be obtained via a linear extrapolation. The pulse duration that ensures this switch-on probability decreases linearly against bias in the log-linear scale. An increase of bias by 0.2V will lead to a $t_{on,99.7\%}$ decrease by nearly two orders. V_{OTS} of 3.5 V is needed in order to achieve a t_{on} of 10 ns with 99.7% of probability by a linear extrapolation. Similar results can also be observed during switch-off and with triangular pulses (not shown). This supports that the t_{on}/t_{off} measurement at constant biases is a reliable method to evaluate the switching probability and conditions.
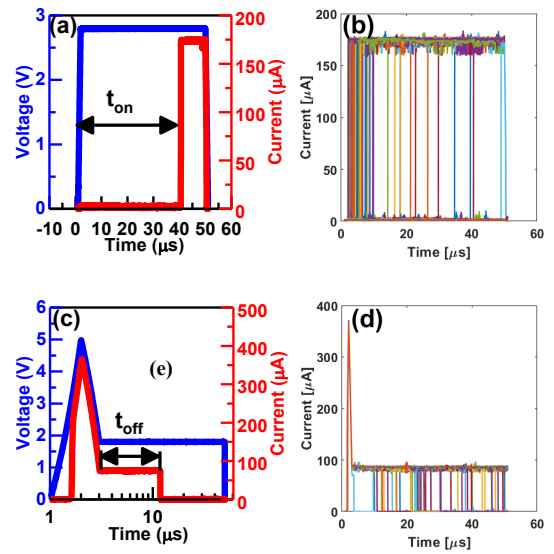

Fig. 2. (a) t_{on} measurement waveform; (b) Typical I-t waveforms of switch-on events at 2.8 V during 100 cycles; (c) t_{off} measurement waveform; (d) Typical I-t waveforms of switch-off events at 1.9 V during 100 cycles. The duration of all constant biases is 50 μs. The rise/fall time is 1 μs for both the triangle and square pulses. The triangle pulse amplitude is 5 V. The sampling time of the current measurement is 10 ns/point.
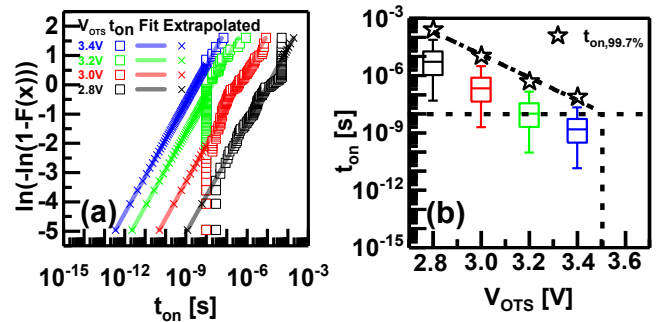

Fig. 3. (a) Weibull plot of measured data (□), linear fitting (——) and extrapolation beyond measurement resolution and pulse duration (×) of t_{on} at different biases. (b) Boxplot of t_{on} at different biases and the extracted t_{on} to ensure the probability of switching-on reaches 99.7% (☆) as obtained from the Weibull distribution. V_{OTS} of 3.5 V is needed to achieve a t_{on} of 10 ns with 99.7% of probability by a linear extrapolation.

4. True random number generator

Since the OTS's switching response shows statistical variability under a constant voltage bias, 100 repetitive square pulses are applied to an OTS selector with amplitude of 2.7V for a duration of 50 μs. Note that t_{on} cannot be recorded below 10 ns due to instrumental limitation. As shown in Fig.4a, around 50% of t_{on} fall below 1 μs, suggesting that under this condition (2.7V, 1μs), the OTS device has a balanced probability to be switched on, which can become a good source of random "0" and "1" generation. Fig.4b shows the schematic of the bitstream generation waveform for the TRNG. Current is measured at the end of each pulse. Since device might be switched on immediately (1st pulse), or not switched on (2nd pulse), or switched on after some time during the pulse (3rd pulse), such stochasticity is used as the source of the randomness generation to convert it to a bit stream '101'.
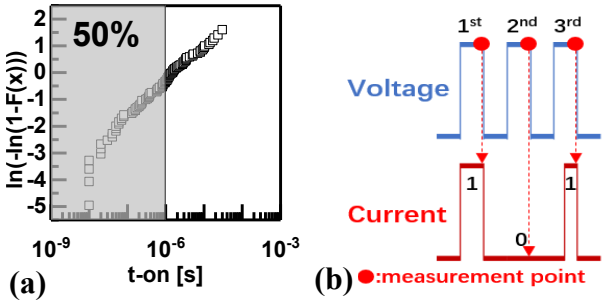

Fig.4. (a) Weibull plot of t_{on} measured from the 100 pulses with Vpulse = 2.7 V and tpulse = 50 μs. ~50% of the t-on falls below 1 μs. (b) Schematic of the bitstream generation waveform. Current is measured at the end of each pulse. Device might be switched on immediately (1st pulse), or not switched on (2nd pulse), or switched on after some time during the pulse (3rd pulse).

To examine the randomness, a sequence of 10,000 square pulses with the width of 1 μs and amplitude of 2.7 V is applied to this OTS device. Current measurement is carried out at the end of each pulse to check whether the device has been switched on. Fig. 5a shows the current measured at the initial 1000 pulses. The low current state refers to the cycles in which OTS remains off, while the high current state refers to the cycles in which OTS has been switched on. The two cases, i.e. switch-on and remain-off, distribute randomly. For the statistical analysis of the entire 10,000 current states, a 2-D kernel density of the current at (n+1)th pulse as a function of the current at nth pulse, i.e. time-lag-plot, is plotted in Fig. 5b to visualize the current states and transitions between them. The balanced color indicates that the two states are indeed statistically balanced. Fig. 5c also shows a bimodal distribution, with an abrupt separation probability at 50%.
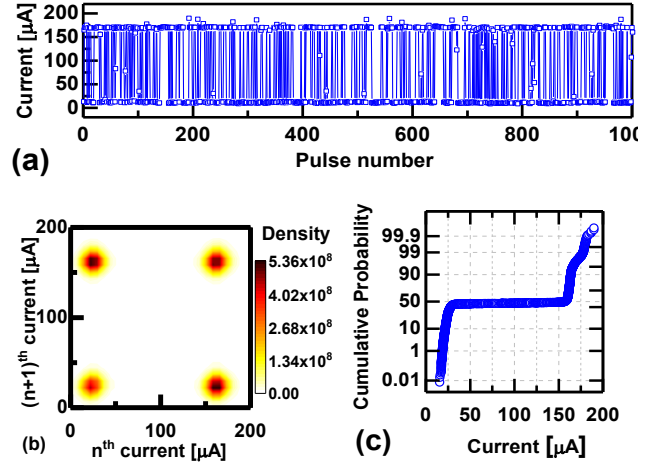

Fig.5. (a) Demonstration of the current measured at the initial 1000 pulses using V_{pulse} = 2.7 V and t_{pulse} = 1 μs. (b) 2-D kernel density of the time lag plot of the entire 10,000 obtained current states. (c) Cumulative distribution plot (CDF) of the 10,000 current states showing a bimodal distribution of states abruptly separated at ~50%

The entire 10,000-bit "0" and "1" sequence is generated by comparing the measured current with a criteria level of 100 μA. "1" will be generated for the switch-on case, and otherwise "0" will be generated, as demonstrated with the initial 20 pulses in Fig. 6a. The generated random bits are presented as a 2-D image in Fig. 6b for clearer illustration. Black color represents bit 0 and white is bit 1. Almost equal number black and white bits are randomly distributed without any regular pattern.
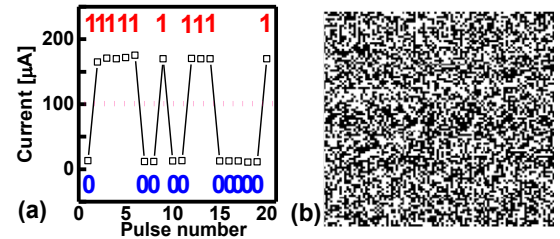

Fig.6. (a) Demonstration of the "0" and "1" bits generated by the initial 20 pulses. 100 μA is used as the criteria (dash line). (b) A 100×100 2-D image representation of the generated 10,000-bit random stream. Black colour represents "0" and white is "1".

The randomness of the bits produced was further evaluated by the National Institute of Standards and Technology (NIST) Test Suite, a statistical package to evaluate the randomness of binary sequences. Table I summarized the test result for the 10,000 random bits generated by the OTS-TRNG. Each test calculates a p-value, and p>0.001 and success proportion ≥ 9/10 are considered good performance. Note that some of the tests consist of several individual tests and we report the smallest p-value out of them. The generated random sequences have passed 12 NIST tests, supporting an excellent randomness performance of the proposed TRNG. The randomness is also evaluated with machine learning test using long

short-term memory (LSTM) architecture, a powerful tool for handling long-range dependencies in general-purpose sequence modeling tasks [10]. The LSTM consists of two hidden layers, with 32 neurons in each layer. A prediction rate of 49.63% has been achieved, further supporting the unpredictability of the TRNG's bit-sequence.

TABLE I: NIST TEST RESULT

Test	p-value	Proportion	Result
Frequency	0.350485	9/10	Pass
Block Frequency	0.350485	9/10	Pass
Runs	0.739918	10/10	Pass
Longest Run	0.534146	9/10	Pass
FFT	0.534146	10/10	Pass
Cumulative Sums	0.004301*	9/10	Pass
Linear Complexity	0.122325	10/10	Pass
Approximate entropy	0.534146	9/10	Pass
Non- Overlapping Template	0.004342*	10/10	Pass
Overlapping Template	0.002043*	10/10	Pass
Serial	0.122325*	9/10	Pass
Rank	0.949536	1/1	Pass

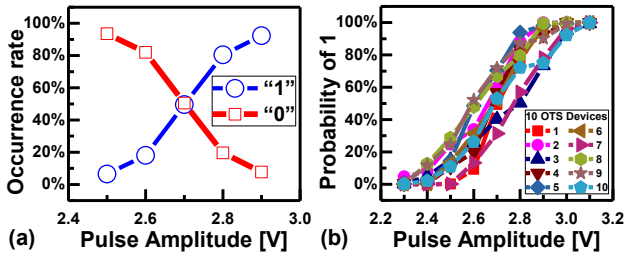*The sequence is divided into 10 streams, except the Rank test*

Fig.7. (a) Occurrence rate of "0" and "1" at different pulse amplitudes, intersecting at 2.7V. (b) Occurrence rate of "1" at different pulse amplitudes measured in 10 devices at different locations across a 300 mm wafer.

As shown in Fig. 7a, the balance between "1" and "0" is affected by the pulse amplitude. "1" appears more frequently with increasing pulse amplitude. 2.7 V is the balance point generating the balanced "0" and "1" random bits while the ratio of "0" and "1" can be modulated by simply tuning pulse amplitude. Such random switching is also observed in 10 OTS device with the same size but at different locations across a 300 mm wafer (Fig. 7b). The balance point varies between 2.6 V and 2.8 V, due to process variation across the wafer. The difficulty in finding the accurate balance point, and the sensitivity of this point to bias and process variations, is a common issue for many emerging TRNG solutions based on novel logic and memory devices such as FeFET, RRAM and magnetic tunnel junction (MTJ) [2-3]. It can be migrated by solutions at the peripheral circuitry level, for example, a bias-sweeping scanning scheme associated with a counter can be designed to determine the balance point.

The volatile switching mechanism of OTS not only makes reset operation unnecessary, but also provides excellent throughput due to its electronic-induced switching nature, which is a great advantage over memristor-TRNGs whose throughput can be bottlenecked by the relative slower atomic/ionic movement. The switching speed of OTS can be shorter than 2 ns. Furthermore, its dependence of switching probability on both voltage and time make it possible to conveniently reduce the pulse width to nanosecond level by increasing the switching voltage, while keeping the bimodal distribution balanced. It is predicted that the throughput can be conveniently boosted to 100 MHz or higher, by increasing the operation voltage for a faster switching speed.

5. Conclusions

In this paper, we characterized the switching probability of Ge_xSe_{1-x} OTS selector device, proposed and demonstrated a novel technique for TRNG application. The statistical switch-on variability following the Weibull distribution can be easily converted into digital "0"s and "1"s with high-quality randomness. The generated 10,000-bit sequence passes 12 tests in the NIST statistical test suite. The randomness is further supported by the machine-learning algorithm. Compared with the TRNGs based on non-volatile memories, the volatile nature of OTS avoids the reset operation hence improves the generation speed. OTS device has shown great potential for TRNGs with excellent reliability, tunability, and throughput.

References

[1] J. Brown, et al, VLSI Symp. Tech. Dig., 2018.
[2] S. Balatti, et al, IEEE Journal On Emerging And Selected Topics In Circuits And Systems, 2015.
[3] W.H. Choi, et al, IEDM Tech. Dig., Dec. 2014.
[4] Z. Chai, et al, IEEE Electron Device Lett., 2018.
[5] Z. Chai, et al, IEEE Trans. Electron Devices, 2017.
[6] B. Govoreanu, et al, VLSI Symp. Tech. Dig., 2017.
[7] F. Hatem, et al, IEDM Tech. Dig., 2019.
[8] D. Garbin, et al, IEDM Tech. Dig., 2019.
[9] Z. Chai, et al, VLSI Symp. Tech. Dig., 2019.
[10] J. Kim, et al, IEEE Trans. Circuits Syst. I, 2019.