Robust FPGA based True Random Number Generator utilizing Oscillatory Metastability in Transition Effect Ring Oscillators

MICHAL VARCHOLA, MILOŠ DRUTAROVSKÝ
Department of Electronics and Multimedia Communications
Technical uniersity of Košice,
Park Komenského 13, 041 20 Košice
SLOVAK REPUBLIC
michal@varchola.com, milos.drutarovsky@tuke.sk

MAREK REPKA
Institute of Computer Science and Mathematics
Faculty of Electrical Engineering and Information Technology
Ilkovičova 3, 812 19, Bratislava
SLOVAK REPUBLIC
marek.repka@stuba.sk

Abstract: - We present novel architecture for True Random Number Generator based on Transition Effect Ring Oscillators. The TRNG consists of 16 transition effect oscillators and also malfunction detector. The malfunction detector can evaluate each TRNG bit right after it was generated. The evaluation utilizes described TRNG mathematical model. Our designed True Random Number Generator is implemented in Actel Fusion FPGA and passes the NIST 800-22 test suite for randomness.

Key-Words: - TRNG, Oscillatory Metastability, Transition effect, FPGA, Randomness

1 Introduction

Tremendous growth of communication systems was significant during the last decades. Heavy research efforts in this area have been, and still are, in the development systems that are more reliable, faster and more power efficient. Simultaneously, substantial research has been carried out to find suitable security solutions in order to prevent cyber-attacks and leakage of the confidential or secret information. Modern cryptography [1] provides powerful techniques for a successful application of security services. Almost every cryptographic system contains a Random Number Generator (RNG) in order to provide random values for underlying algorithms. Random numbers are required e.g. as session keys, signature parameters, temporary keys, challenges or in zero knowledge protocols, and therefore they should meet very strict requirements – they should be unpredictable, uniformly distributed on their range and independent [2]. A RNG of insufficient quality can weaken an otherwise strong cryptographic system as was shown e.g. in smart card attack in [3].

We can divide RNGs into two main subgroups [2]: Pseudo RNG (PRNG) and True RNG (TRNG). The output of PRNG is mathematically defined and all entropy is given by the (preferably random) seed. On the other hand, the entropy of TRNG is increased by each generated bit and its output cannot be described in a deterministic way. TRNGs employ a physical phenomena and are inseparable part of modern security equipments. There are several sources of entropy: non-physical (e.g. access time of hard drive, keystrokes, computer mouse movements) and physical (thermal noise, nuclear decay). Practically, the most useful for embedded equipment is the electronic noise, which exhibits in various electronic platforms such as ASICs (Application Specific Integrated Circuits), MCUs (MicroControllers) and FPGAs (Field Programmable Gate Arrays).

Employing the suitable randomness sources in an FPGA is still a challenging research task that recent papers included in [2] underline. The most popular randomness sources in FPGAs are: a time delay instability of logic components, a time

ISBN: 978-1-61804-271-2

instability of Phase Locked Loop (PLL) clock signals, and an analogue properties of the logic gates (e.g. metastability). The time delay instability of the logic components causes e.g. a jitter of Ring Oscillators' (ROs) output, what was analyzed in [4], [5], or[6]. The most straightforward RO-based TRNG design suitable for the FPGAs was proposed in [7]. Note, that its reliability is being heavily discussed by the World TRNG community and it is still remaining to be an unanswered question [8]. However, it is recently turning out that the ROs are unreliable as a source of randomness due to a low entropy [9] and due to high dependence on external or internal deterministic perturbations [10], [11]. Time instability of the PLL clock signals as a reliable randomness source was firstly proposed in [12]. Despite of the thorough reliability of the design, such TRNG consumes rare FPGA components – PLLs that are even not available in all FPGAs. The metastability has been analyzed for decades, especially in terms of reliable synchronization circuits synthesis [13], [14], or [15]. As the digital circuit is extremely sensitive to noise of a circuit during the metastable state, this phenomenon was suggested as a randomness source even in FPGAs e.g. in [19], or [23], but the majority of the metastability-based TRNG designs are aimed at the ASIC technology or custom designs [16, 17,18], while synthesis in the FPGAs is considered to be rather awkward [19]. This is caused by the inability to implement efficient stabilization feedback mechanism required for reliable operation of traditional metastable structures in pure digital FPGA circuits in contrary to the custom (analog) designs where such implementations are (at least in principle) straightforward.

The aim of this paper is to introduce a complete reliable TRNG with a malfunction detector embedded in FPGA that is based on recently proposed Transient Effect Ring Oscillator (TERO) element that utilizes oscillatory metastability [20]. TERO structure and extraction mechanism belongs to the class of ROs with even number of inverting elements firstly proposed for TRNG generation in [21] and independently in [20], later analyzed also in [22]. Authors of [21] propose to use general ROs with even number of inverting elements while [20] concentrates to a very efficient TERO structure with the minimal number of two inverting elements that can be very efficiently implemented in one CLB block of Xilinx FPGA. Such method is capable to extract very small internal circuit random variations, what was underlined by a simple mathematical model and practical comparison of TERO elements and classical RO based element in Xilinx FPGA [20]. Such model is improved within this paper by assuming more realistic model of noise. Moreover, advantages of TERO structure are compared with similar FPGA TRNG designs utilizing oscillatory metastability that were proposed up to now. Proposed complete TRNG embedded in Acel FPGA incorporates also an optional malfunction detection mechanism, which can evaluate circumstances under which the random bit was generated. Furthermore, plenty of realized experiments underline the robustness of the entire TRNG where a variation of the operating condition in a very wide range (even far outside of the recommended Actel FPGA operating condition range, particularly the core power voltage and the environment temperature) does not affect the quality of the resulting random numbers at all.

2 Transition Effect Ring Oscillator and Related TRNG Designs based on Oscillatory Effects and Metastability

TERO element and randomness extraction mechanism described in the next section is based on a combination of transient oscillatory metastability and behaviour of bistable flip-flops. Similar sources of randomness, but in different context were already used independently in some previous TRNG designs. This section summarizes shortly existing principles, compares their basic features and highlights main advantages of recently proposed structures that use "differential" feature of multiple transient oscillating waves in the specially designed and suitably stimulated RO loop with even number of inverting elements.

Authors in [23] use two independent free-running short length ROs that were dynamically switched on (in the so-called oscillation phase) and off (in the so-called resolving phase) by switching elements. In the resolving phase, two inverting elements of independent ROs are interconnected in such a way that they create a bistable memory element. Because of positive feedback, the outputs of inverters can resolve by flipping to a consistent final logic state. This final state represents one random output bit. The randomness comes from a combination of the drift and the jitter of two independent free-running ROs during the oscillation phase and possibly also from a metastability effect during the resolving phase.

Design [19] proposes a generic Meta-RO architecture that uses digital elements initially forced into individual near-metastable states

ISBN: 978-1-61804-271-2

(initialization phase) to be next reconfigured to an free oscillating RO structure. The output random bit is got by sampling RO output signal during stabilized oscillation phase, once transition process from near-metastable state disappears. Meta-RO design was intended for implementation in ASICs, but authors claim that it is also feasible in Xilinx FPGAs. Although this design is worth consideration, we believe assumption that the bi-stable structure reaches indeed a metastable state cannot be confirmed by measuring the internal signals (having analogue behaviour) that are routed outside the device via standard logic input/output pins, as it is presented in [19].

In complete TRNG design we can use the following features of TERO element embedded in FPGA:

• Sufficiently higher entropy rate than previous RO based designs,
• Lower sensitivity on global interference and working conditions than previous RO based designs,
• Ability to extract reliably very small intrinsic noise generated by FPGA logic elements,
• Clear description of the simplified mathematical model describing basic circuit behaviour,
• Ability to restart the element before each random bit generation period in order to utilize the stateless entropy concept [BGL+06],
• Usage of least number of logic elements all implemented in the single block of logic to minimize signal
paths, minimize interference, minimize resources utilization, and decrease the power consumption,
• Element structure has simple place and route strategy and clear recommendations on how to synthesize the structure in the target FPGA.
• Ability for several entropy elements operating independently and in parallel in order to place them into the same FPGA for enhancing statistical parameters and/or increasing the bit-rate,
• Inner testability feature in order to detect instantly when the entropy source is out of order and/or has weak statistical properties, the inner testability feature should be implemented as a simple circuitry tailored on the particular principle of the randomness source.

The complete TRNG architecture have to compensate natural dependence of TERO element on the distribution of ”analogue” parameters of internal FPGA resources that cause deviation form perfect symmetry of the TERO element embedded in practical FPGA. The TRNG architecture based on a set of TERO elements presented in the next sections provides highly robust solution and is demonstrated for Actel FPGA operated even out of recommended FPGA working conditions. Although proposed TRNG architecture is deeply tested only in Acel FPGA, it principle is quite general and can be easily extended to the Xilinx or Altera FPGAs.

3 TERO Element Implementation and Analysis

3.1 The TERO Structure

TERO structure used for experiments in Xilinx Spartan 3E FPGA [20] incorporates two XORs, two ANDs and several control signals. We found out that structure can be simpler when purposes of XORs (forcing oscillatory metastability) and ANDs (forcing the same initial conditions) are merged into NANDs. Feedback paths are lengthened by pair of two inverters. TERO with the simplified control adopted for Actel Fusion FPGAs is depicted in Fig.1. Shape of control waveform and consequent output waveforms are given in Fig.2. New TERO structure has two phases of operation: the reset phase and the oscillation phase. The ctrl='0' activates the former and the '0' to '1' ctrl transition activates the latter. The purpose of reset phase is to force the same starting electrical conditions for each generating of a random bit in order to fulfil the stateless entropy concept [24]. All parasite capacitances are charged or discharged to the same level during this phase. That means longer duration of reset phase can ensure better convergence to the same starting conditions. Accordingly, newly proposed TERO structure has the advantage of several times longer reset phase in comparison to its predecessor [20]. The reset phase is altered by the oscillation phase. The '0' to '1' ctrl transition causes disturbance of steady conditions forced by the reset phase and TERO begins to oscillate due to lengthened feedback paths. Generated oscillations will disappear after a while due to shortening or enlarging of duty-cycle of generated signals. The shortening or enlarging is caused by unbalanced TERO that incorporates even number of inverting elements in the loop. We will assume just shortening in the next text for the simplification. This behaviour is denoted as an oscillatory metastability of a bistable structures in [13]. The random bit is extracted as a LSB of number of oscillations. The detail of custom implementation of 16 TERO structure in Actel Fusion M1AFS600 FPGA is shown in Fig.3. All TERO channel are placed as close as possible. We can do so since we did not observe no interlocking and correlation between neighbouring TEROs. We use the same place topology for each TERO.

ISBN: 978-1-61804-271-2

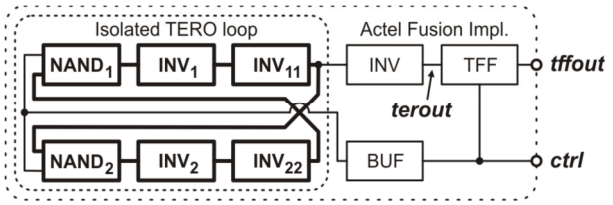

Fig.1 Practical circuit of TERO adopted for Actel Fusion FPGAs. All depicted elements are places as close as possible. Usage of BUF and INV enables such routing that signal directly connected to internal TERO loop will not be routed by long path. NANDs are used for forcing oscillatory metastability state and forcing the same initial conditions as well. INV 1, INV 2, INV 11 and INV 22 ensure sufficient feedback path extension in order to force oscillatory metastablity reliably. TFF is used for random bit extraction. TFF is cleared by ctrl='0' and therefore tffout should be sampled before falling edge of ctrl.

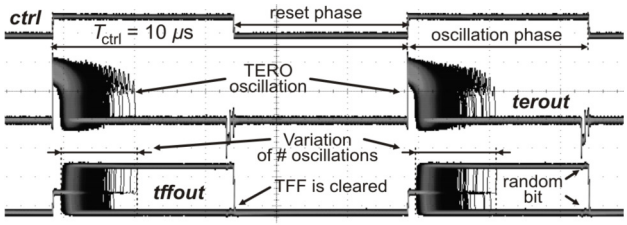

Fig.2 The new TERO operation oscilloscope screen-shot captured using infinite persistence mode and 20MHz low-pass filter on ctrl and tffout channels. The image was acquired by the Tektronix MSO 4104 oscilloscope. The rising edge of ctrl causes oscillations of the TERO loop. The number of oscillations observed varies during each ctrl period. TFF resolves whether TERO made odd or even number of oscillation periods during one ctrl period what represents one random bit (tffout signal). The same initial conditions before forcing oscillatory metastablity are ensured by ctrl='0'.

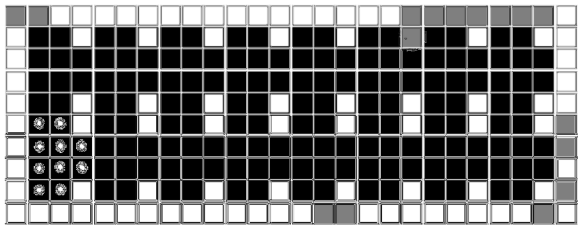

Fig.3 The TERO cluster consists of 16 TEROs. The elements belonging to single TERO channel is highlighted by dots. All TERO channels share the same custom place topology.

4 The TERO model

Although structure of new TERO is slightly changed, the mechanism of TERO randomness extraction remains unchanged and the basic model proposed in [20] is still applicable. The model is given in Fig.4 and works as follows: when a rising of ctrl appears, TERO loop begins to oscillate (Fig. 2). The mean value of TERO oscillation period is equal to a total delay of TERO loop T_T. An excited pulse of starting logic '1' level pulse length T_S is shortened during each oscillation by T_D time due to slight intrinsic non-symmetry of the loop. Excited pulse will disappear when instant logic '1' level pulse width reaches minimal possible value T_M. Asymmetry T_D is assumed to be affected by a period jitter Tij , where i and j stands for i-th T_T period and j-th Tctrl period respectively. The final number of oscillations executed for j-th T_{ctrl} is denoted as Y_{Tj} . The basic mathematical model of TERO mode is expressed as [20]:

$$T_S - T_M = \sum_{i=1}^{Y_{T_j}} \left(T_D + \Delta_{T_{ij}} \right) = T_D \cdot Y_{T_j} + \sum_{i=1}^{Y_{T_j}} \Delta_{T_{ij}} \quad .$$

Both, TS and TM can be slightly affected by intrinsic noise and so considered to be contribution to final randomness. Value of the former can be affected by actual noise conditions when circuit is entering to oscillatory metastablity phase and value of the latter can be affected by actual noise conditions when the circuit does (or does not) allow to pass last pulse. Let us denote the former as the phase of the oscillation start up and let us denote the letter as the phase of the oscillation stop. Let us denote the phase between them as phase of oscillatory trajectory. We will use alternative designation of those three phases, namely: the first, the second, and the third phase, according to their consequence in time. We will analysed them in sections that follow.

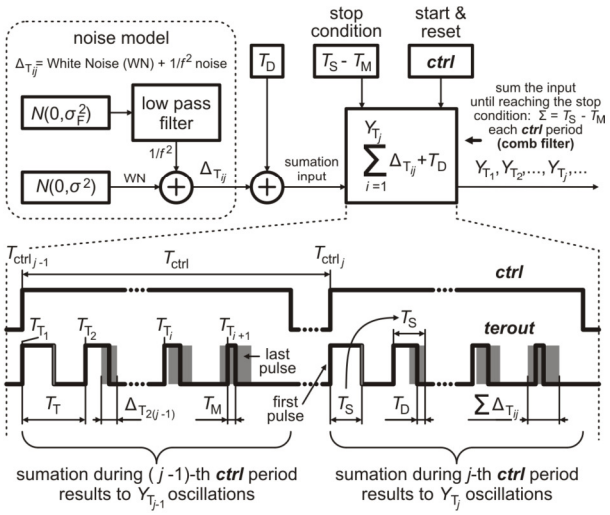

Fig. 4 Graphical interpretation of TERO mathematical model. The summation in it acts as an adaptive digital comb filter with variable (YTj) summing elements and thus TERO can extract noise of lower frequency better. The model is accompanied by an example of terout and ctr waveforms.

ISBN: 978-1-61804-271-2

5 Complete TERO-based TRNG with Internal Testing

This section introduce a complete TRNG based on a cluster of 16 TEROs. Entire TRNG is accompanied by the malfunction detector block used for internal testing that can uncover potential weaknesses of the entropy source.

5.1 Structure of the TERO-based TRNG

The block diagram of the complete TRNG is depicted in Fig. 5. The TRNG was implemented in the Actel Fusion FPGA. The TRNG consists of the Control FSM, 16 TERO channels, XOR-chain, and Malfunction Detector. Each TERO channel consists of: TERO loop (TERO 1-16), asynchronous counter (ACNT 1-16), sampler (SMPL 1-16), and comparator (CMPR 1-16). The purpose of TERO, explained in previous sections. The ACNT is asynchronous counter and counts number of oscillations. The SMPL samples value of ACNT at the end of generation period. The purpose of the CMPR is to decide whether number of oscillation TERO done fits to threshold region. The range of threshold region is implicit from both, the mathematical model (lower bound) and overflow of number of oscillations due to time-limited ctrl (upper bound). The practical range of the CMPR for the TERO structure from Section 9 can be approximately from TH= 70 (bottom threshold) to TH= 1000 (top threshold) oscillations. When number of oscillations fits to such range the random bit generated by particular TERO channel is considered as good. The Malfunction Detector block can generate alarm according to information from CMPR 1-16. The alarm is generated in the case when neither three TERO channels from sixteen does not fulfil described criteria. Experiments showed that the of two TERO outputs, which have sufficient oscillation variation can produce uncorrelated random bit stream of satisfactory statistical properties. However the minimal number of three is not mandatory. Malfunction detector is shown in Fig.6. It compares number of oscillations of each TERO with threshold values. If majority of TEROs does not produce enough oscillations, the alarm based on decision logic is signalized.

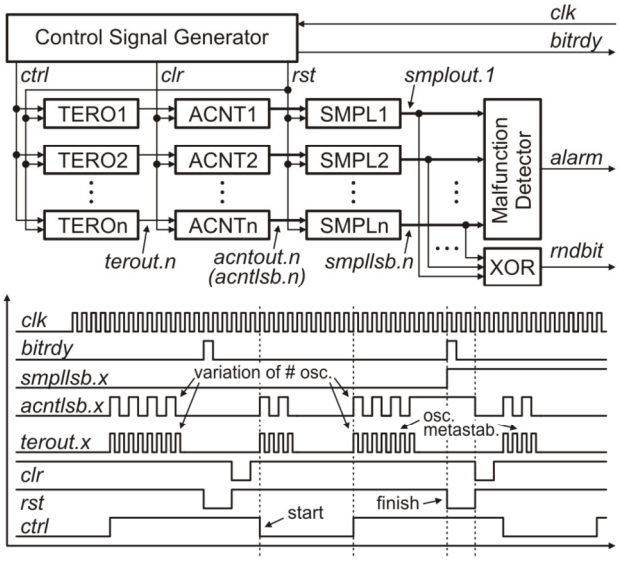

Fig. 5 The block diagram of the complete TRNG with built-in malfunction detection with waveforms of the signals

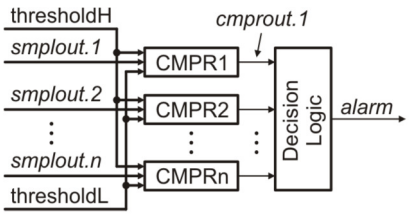

Fig.6 The malfunction detector schematic diagram – number of each TERO oscillations should fall into range between thresholdL and thresholdH.

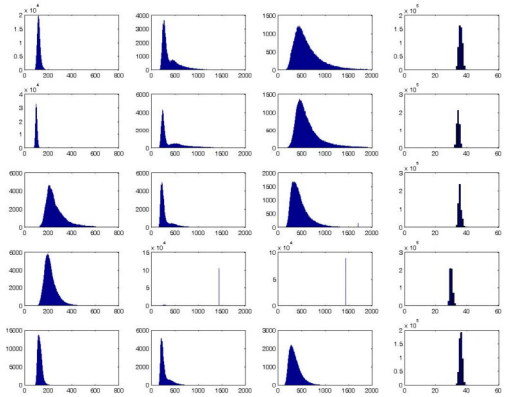

Fig. 7 Sensitivity to operating conditions variation. The TERO #1 #3 #11 and #15 channel (columns) under various temperatures and core power voltages using the Actel Fusion M1AFS600 PQG208 FPGA in rows: a) standard 1500mV, +20C, b) 1500mV, -10C, c) 1500mV +110C, d) 1210mV, +20C, e) 1640mV, +20C

ISBN: 978-1-61804-271-2

6 Experimental Results

The robustness of such TRNG has been evaluated using the Actel Fusion M1AFS600 PQG208 FPGA by a set of experiments under violated working condition such as wide core power voltage range (from VCORE = 1250mV to VCORE = 1640mV, while nominal value is VCORE = 1500mV) and wide temperature range (from T=−10C to T=110C). A long 1 Gbit random bit sequence has been acquired under each working condition setup (with the exception of temperatures above T=80C in order to prevent the destruction of the board). Such long sequences were evaluated by NIST 800-22 statistical tests suite with satisfactory results. The NIST 800-22 suite tested 1000 sequences of 220 kbits long each. The histograms of number of oscillations for various experiments are given in Fig. 7. The content of figures is explained in their captions. The NIST 800-22 tests pass even when working conditions are violated – the lowest temperature and lowest core power voltage among all experiments. TRNG synthesis in different boards and in different relative placements within single FPGA does not affect satisfactory results of the NIST 800-22 tests too.

7 Conclusion

A new TERO-based TRNG has been introduced. It possesses a great ability to extract internal noises of FPGA logic cells and this features contrasts with traditional RO-based TRNGs. The high statistical quality that is independent of the FPGA device, in particular P&R, in working conditions was underlined by statistical NIST 800-22 tests [25] that gave satisfactory results. The designed structure accompanied by a malfunction detector provides a practical solution for a TRNG ready for implementation in modern FPGA devices despite the fact that a single TERO needs more precise P&R than a traditional single RO.

Acknowledgment:
This work has been supported in part by the grant APVV-0586-11.

References:

[1] J. Menezes, P. Oorschot, and S. Vanstone, Handbook of Applied Cryptography. New York: CRC Press, 1997.

[2] K. Koç, Ç., Ed., Cryptographic Engineering. Springer, 2009.

[3] A. T. Markettos and S. W. Moore, "The Frequency Injection Attack on Ring-Oscillator-Based True Random Number Generators," in Proceedings of 11th International Workshop on Cryptographic Hardware and Embedded Systems (CHES), Lausanne, Switzerland, September 6-9, 2009, pp. 317 – 331.

[4] A. A. Abidi, "Phase Noise and Jitter in CMOS Ring Oscillators," IEEE Journal of Solid-State Circuits, vol. 41, no. 8, pp. 1803–1816, August 2006.

[5] B. H. Leung, "A Novel Model on Phase Noise of Ring Oscillator Based on Last Passage Time," IEEE Transactions on Circuits And Systems, vol. 51, no. 3, pp. 471–482, March 2004.

[6] M. Mandal and B. C. Sarkar, "Ring oscillators: Characteristics and applications," Indian Journal of Pure & Applied Physics, vol. 48, pp. 136–145, February 2010.

[7] B. Sunar, W. J. Martin, and D. R. Stinson, "A Provably Secure True Random Number Generator with Built-in Tolerance to Active Attacks," IEEE Transactions on Computers, vol. 56, no. 1, pp. 109–119, January 2007.

[8] M. Dichtl, B. Meyer, and H. Seuschek, "SPICE Simulation of a "Provably Secure" True Random Number Generator," 2008. [Online]. Available: http://eprint.iacr.org/2008/403.pdf

[9] N. Bochard, F. Bernard, and V. Fischer, "Observing the randomness in RO-based TRNG," in International Conference on Reconfigurable Computing and FPGAs, Cancun, Quintana Roo, Mexico, December 9–11, 2009, 2009, pp. 237–242.

[10] V. Fischer, F. Bernard, N. Bochard, and M. Varchola, "Enhancing Security of Ring Oscillator-based RNG implemented in FPGA," in Proceedings of 18th International Conference on Field Programable Logic and Applications (FPL), Heidelberg, Germany, September 08–10, 2008, 2008, pp. 245–250.

[11] T. Pialis and K. Phang, “Analysis of Timing Jitter in Ring Oscillators Due to Power Supply Noise,” in Proceedings of the International Symposium on Circuits and Systems (ISCAS), Bangkok, Thailand, May, 25–28, 2003, 2003, pp. 685–688.

[12] V. Fischer and M. Drutarovsky, “True Random Number Generator Embedded in Reconfigurable Hardware,” in Proceedings of 4th International Workshop on Cryptographic Hardware and Embedded Systems (CHES), Redwood Shores, CA, USA, August 13-15, 2002, Springer, 2002, pp. 415–430.

[13] T. Kacprzak, “Analysis of Oscillatory Metastable Operation of an R-S Flip-Flop,” IEEE Journal of Solid-State Circuits, vol. 23, no. 1, pp. 260–266, February 1988.T. Kacprzak, “Analysis of Oscillatory Metastable Operation of an R-S Flip-Flop,” IEEE Journal of Solid-State Circuits, vol. 23, no. 1, pp. 260–266, February 1988.

[14] L. Reyneri, D. Corso, and B. Sacco, “Oscillatory Metastability in Homogenous and Inhomogeneous Flip-Flops,” IEEE Journal of Solid-State Circuits, vol. 25, no. 1, pp. 254–264, February 1990.

[15] J. U. Horstmann, H. W. Eichel, and R. L. Coates, “Metastability Behavior of CMOS ASIC Flip-Flops in Theory and Test,” IEEE Journal of Solid-State Circuits, vol. 24, no. 1, pp. 146–157, January 1989.

[16] Tokunaga, C.; Blaauw, D.; Mudge, T.; , "True Random Number Generator With a Metastability-Based Quality Control," Solid-State Circuits, IEEE Journal of , vol.43, no.1, pp.78-85, Jan. 2008

[17] Holleman, J.; Bridges, S.; Otis, B.P.; Diorio, C.; , "A 3 uW CMOS True Random Number Generator With Adaptive Floating-Gate Offset Cancellation," Solid-State Circuits, IEEE Journal of , vol.43, no.5, pp.1324-1336, May 2008

[18] Nakura, T.; Ikeda, M.; Asada, K., "Ring oscillator based random number generator utilizing wake-up time uncertainty," Solid-State Circuits Conference, 2009. A-SSCC 2009. IEEE Asian , vol., no., pp.121-124, 16-18 Nov. 2009

[19] I. Vasyltsov, E. Hambardzumyan, Y. S. Kim, and B. Karpinskyy, “Fast Digital TRNG Based on Metastable Ring Oscillator,” in Proceedings of 10th International Workshop on Cryptographic Hardware and Embedded Systems (CHES), Washington, DC, USA, August 10-13, 2008, Proceedings, ser. LNCS, vol. 5154. Springer, 2008, pp. 164–180.

[20] M. Varchola and M. Drutarovsky, “New High Entropy Element for FPGA based True Random Number Generators,” in Proceedings of 12th International Workshop on Cryptographic Hardware and Embedded Systems (CHES), Santa Barbara, CA, USA, August 17-20, 2010. Springer, 2010, pp. 351–365.

[21] M. Dichtl and B. Meyer, “Apparatus and Method for Generating a Random Bit Sequence,” German Patent WO/2010/031630, March 2010, assignee: Siemens Aktiengesellschaft, Germany. [Online]. Available: http://www.wipo.int/pctdb/en/wo.jsp?WO=2010031630

[22] L. Hars, “Random Number Generation Based on Oscillatory Metastability in Ring Circuits”, eprint.iacr.org/2011/637.pdf

[23] M. Epstein, L. Hars, R. Krasinski, M. Rosner, and H. Zheng, “Design and Implementation of a True Random Number Generator Based on Digital Circuit Artifacts,” in Proceedings of 5th International Workshop on Cryptographic Hardware and Embedded Systems (CHES) Cologne, Germany, September 8–10, 2003, ser. LNCS, vol. 2779. Springer, 2003, pp. 152–165.