Low Cost and Energy, Thermal Noise Driven, Probability Modulated Random Number Generator

Nicoleta Cucu Laurenciu, and Sorin D. Cotofana
Computer Engineering Laboratory,
Delft University of Technology, The Netherlands.
{N.CucuLaurenciu, S.D.Cotofana}@tudelft.nl

Abstract*—Random Number Generators (RNGs) constitute the essential foundation of many applications, e.g., cryptographic technology, and statistic based computing. The vast majority of RNG proposals, generate binary sequences with uncorrelated and equiprobable bits. However, in certain applications, such as stochastic computing, the binary sequence is required to be generated not with 0.5 logic "1" probability, but with a custom specified one. This paper presents a probability modulated True Random Number Generator (TRNG), that produces binary sequences with a desired probability of logic "1", according to a value resident in a register. The proposed circuit relies on the thermal noise as random signal source and it was implemented in $65nm$ CMOS technology. Simulation results reveal a quasi-linear dependence, between the output logic "1" probability sampled at $1GHz$, and the modulating voltage (obtained from the register value by mean of a D/A converter) over a range of $20mV$. For a desired probability of 0.5, the sequences randomness was validated by the NIST tests.

***Index Terms*—probability modulated random number generator, thermal noise, stochastic computing.**

I. Introduction

Random Number Generators (RNGs) are employed in many applications, e.g., communication systems, stochastic computing, cryptography. Several approaches have been proposed in the literature [1]: (i) pseudorandom techniques, which are provably non-random, and which suffer from strong long-range correlations, and (ii) truly random techniques, which extract randomness from physical phenomena that behave in a nondeterministic way, thus making them better candidates for true random number generation. However, for the latter case, challenges appear in the randomness extraction, i.e., the randomness has to be extracted in such a way that the ideal functioning of the physical randomness source is not disrupted by introducing deterministic biases which may favor resolution towards one logic state or another.

All these approaches generate sequences of equiprobable and uncorrelated bits. In certain applications, however, such is the case of stochastic computation [2] for instance, randomly generating bitstreams that are consistent with a particular probability is of great interest. Stochastic computing is presented as a promising avenue towards leveraging the hardware complexity and designing robust and energy-efficient circuits, in nanoscale process technologies. In stochastic computation, probabilities are represented by stochastic bit streams, which allows for the implementation of complex arithmetic operations via simple stochastic logic. In this way, the stochastic bit streams are processed in a serial manner (e.g., the multiplication of two probabilities, encoded as two stochastic bit streams is performed using only an AND gate serially processing the two stochastic streams). However, the computation precision is limited by the randomness properties of the stochastic bit sequences (i.e., of the input operands of a stochastic computation), with the correlation between stochastic bit streams possibly leading to an erroneous result. As different probabilities are required to be represented as binary sequences, low area and energy efficient probability modulated RNGs are desired as foundation blocks in stochastic computing.

To this end, we propose a true RNG-based approach. We consider a minimum size MOSFET as thermal noise source. The generated noise, sensed as a small AC voltage, is then amplified up to a certain level, such that the DC voltage level on top of which the AC noise component is superimposed, can be accurately thresholded, without bias, by an inverter. As a result of the noise, the inverter will undergo probabilistically a series of transitions from logic "1" to logic "0" and viceversa. The inverter logic "1" output probability is adjusted by varying the DC voltage level around the inverter switching point voltage. This is achieved by using a D/A converter and a register that drives it according to the stored desired probability value, i.e., the frequency of occurrence of logic "1" at the inverter output (e.g., a register value '00111' is equivalent to a probability of $7/31$). The proposed low cost, low energy (due to the exponential energy-probability dependence [3]) probability modulated RNG system was implemented in CMOS 65nm technology, its design parameters were characterized by means of Cadence simulation and its output bit sequences statistical properties were evaluated. Experimental results reveal an almost liner dependence between the probability of logic "1" measured at the inverter output and the D/A control voltage, over a range of $0 \div 20mV$ at a $1GHz$ output sampling frequency. Moreover, the randomness quality of the output bitstream for a modulating probability of 0.5, was successfully verified using the NIST test suite [4].

The remaining of the paper is organized as follows: We first present a general overview on RNG circuits, and their associated benefits and challenges in Section II. In Section III, we discuss the proposed circuit topology and its implementation details. Then, in Section IV, we present and discuss the experimental results. We conclude the paper with some final

978-1-4799-8391-9/15/$31.00 ©2015 IEEE

remarks in Section V.

II. Preliminary Concepts and Related Studies

Generally speaking, there are two kinds of RNGs [1]: (i) Truly Random Number Generators (TRNGs), whose output sequences are produced by exploiting a random physical process, and (ii) Pseudo-Random Number Generators (PRNGs), which are algorithm based, and require an externally generated sequence as initial seed based on which, output sequences which appear to be random are deterministically generated.

In general, a TRNG architecture consists of three main components: (i) a physical noise source, (ii) an entropy detector circuit, which harvests as much entropy as possible via a preferably non-invasive circuit/mechanism (i.e., that doesn't disturb the physical randomness process, and thus doesn't introduce any deterministic bias), and (iii) a randomness extractor, which handles the sampling in order to obtain a discrete sequence of bits with the desired statistical properties. In some cases, additional binary output sequence post-processing circuitry might be required, to compensate/mitigate the entropy and randomness loss.

Numerous randomness entropy sources exploiting an underlying random physical process, e.g., electronic components noise, Brownian motion, quantum phenomena, nuclear decay, meta-stability of devices, chaos for deterministic analog signals, and chaotic optical signals from lasers and photonic ICs, have been proposed. In general, it is hard to find a hardware source with enough entropy and/or randomness. Furthermore, the statistical characteristics of the physical hardware source change over time as a result of environmental influences (e.g., temperature, supply voltage), aging induced wear-out, etc. Because of these effects, an initially relatively good random source may become less random in time.

As concerns the entropy harvesting circuit, relatively few approaches have been proposed, among which: direct amplification of a small AC signal produced by electronic noise [5], free-running oscillators with phase/frequency drift [6], and pipelined A/D converter based architecture for exploiting chaos on a deterministic random signal.

Typically a randomness extractor circuit, can be implemented by an A/D converter, a clocked comparator, an inverter, or by any other embodiment that allows the discrimination between two logic states.

The imperfections of both the entropy source, and of the randomness extractor circuit, caused by, e.g., Process, Voltage and Temperature (PVT) variations, components aging, and other perturbing deterministic phenomena, may lead to an output binary sequence which may not have the desired degree of randomness. To this end, various post-processing techniques have been proposed, in order to reduce/eliminate the output binary sequence randomness statistical characteristics drift. Von Neumman correction [7], XOR [7], and hash functions [8], are a few commonly employed post-processing techniques to enhance the randomness properties of the generated binary sequence.

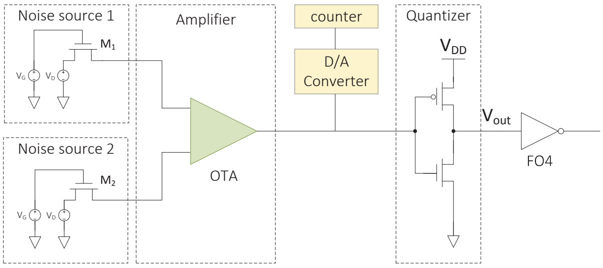


Fig. 1: Probability Modulated TRNG Architecture.

A TRNG aim is to produce independent and equiprobable bits, through a non-deterministic and unreproducible process. However, many applications, require generation of random bits whose occurrence frequency is given by a specified probability. Such a requirement adds further to a TRNG system design space complexity. In stochastic computing, a costly Feedback Shift Register based approach is usually used as probability biased RNG [9]. Furthermore, as the pseudorandom bitstreams are periodic, the entropy decreases asymptotically as more output sequences are produced, implying long-range correlations between the generated sequences. As far as the analog probability modulated RNG approaches are concerned, a related body of work is termed probabilistic switching [10]. The authors propose an analytical model which exploits the random behavior of a binary switch, specifically of a CMOS inverter, induced by noise accrued from multiple sources (i.e., power supply, and input/output coupled thermal noise sources). In [11], a $0.5\mu m$ Mosis AMI CMOS energy efficient implementation of a probabilistic switch based TRNG system is presented. A $10G\Omega$ resistor is employed as noise source, which provides thermal noise with a Root Mean Square (RMS), value of $4mV^2$. The small AC voltage is subsequently amplified by a two-stage common source amplifier, and then converted to a binary sequence of bits by using a CMOS inverter. In [12], following the same line of reasoning, the authors, in order to improve the randomness figures of merit, exploit the decreased susceptibility to process variability afferent to a Fully Depleted Silicon On-Insulator (FDSOI) technology for implementing the inverter.However, they employ synthetic AWGN noise sources, instead of the circuit with a noise amplifier circuit.

In the next Section we describe the proposed probability modulated TRNG architecture. We also exploit thermal noise, but in our case it is generated by a minimum size transistor. The noise is then amplified, while accounting for environmental influences, which might introduce deterministic biases, and finally a threshold decision is performed by using an inverter to discriminate whether the sampled RNG output is logic "1" or "0".

III. Circuit Architecture

The proposed probability modulated TRNG block scheme is graphically illustrated in Figure 1. It comprises the following:

A. Entropy Source

Among the natural randomness sources, thermal noise is the most widely employed one in TRNG circuits. It is a white

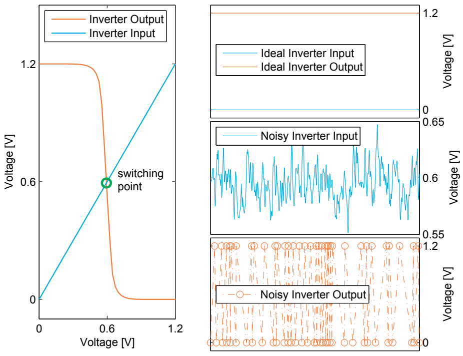


Fig. 2: Inverter Input and Output Voltage.

noise (i.e., frequency independent spectral density) with normal distribution, which is caused by the temperature induced Brownian motion of electrons in a conductor. Considering passive devices at room temperature, a $1k\Omega$ resistor can typically generate $4nV$ thermal noise Root Mean Square (RMS), while capacitors of $1fF$ and $10fF$ can yield $2mV$ and $640\mu V$ noise RMS, respectively. High impedance resistors, based on MOSFETs biased in linear region, are well known white noise generators. However, parasitic capacitances, which cause the thermal noise to be bandlimited, and thus reduces its amplitude, as well as non-zero DC current, which results in non-white (i.e., correlated noise samples) $1/f$ noise, have to be accounted for. We employ as thermal noise source, the nonlinear gate-to-channel parasitic capacitance of an NMOS transistor M_1 which operates in the triode regime, with its drain terminal connected to the gate of a transistor from the following TRNG stage [13] (terminal V_{in+} of the amplifier in Figure 6a). In this way, the noise magnitude is given only by the parasitic capacitance, and is independent of the M_1 drain to source resistance. Additionally, the M_1 drain to source voltage is equal to 0, alleviating thus the $1/f$ noise artifacts.

B. Quantizer

The switching point of an ideal deterministic inverter, i.e., the point where the input voltage is equal to the output voltage, as illustrated in Figure 2, allows the discrimination between a logic "1" and a logic "0" inverter output. Thus if instead of an either a logic "0" or a logic "1" input signal, the inverter is fed by an input voltage which varies randomly around its switching point, the corresponding inverter output undergoes spurious transitions to logic "1" and logic "0", and is thus probabilistic, as depicted in Figure 2. The inverter input signal consists of two components: (i) a DC voltage component, that is given by an A/D converter which level shifts a voltage equal to the inverter switching voltage with an amount modulated by the value stored in the register, and (ii) a random noise AC component, that is given by the entropy harvesting circuit described subsequently.

C. Entropy Harvesting Circuit

As the source generated thermal noise amplitude is much smaller than the semiconductor circuits logic levels (in the order $10^3\sigma$), amplification is required. To this end, we employ an Operational Transconductance Amplifier (OTA) [14] [15], as illustrated in Figure 6a. To reject the common-mode and power supply interferences, two noise sources are coupled at the positive and negative inputs of a differential amplifier. While the OTA output nominally is DC offset free, components mismatch, temperature, aging, etc. might cause a time-varying offset. Figure 3 illustrates the two voltages obtained at the OTA output, which are subtracted from one another, via a simple unity gain differential amplifier (see Figure 1), to achieve DC offset cancellation. An alternative architecture consists of an OTA with Common-Mode FeedBack (CMFB) and a unity gain amplifier in order to keep the DC output voltage at approximately the inverter switching voltage (i.e., $\approx 1/2V_{DD}$).

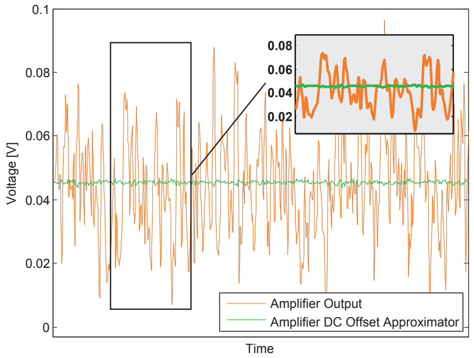


Fig. 3: Amplifier DC Offset Cancellation.

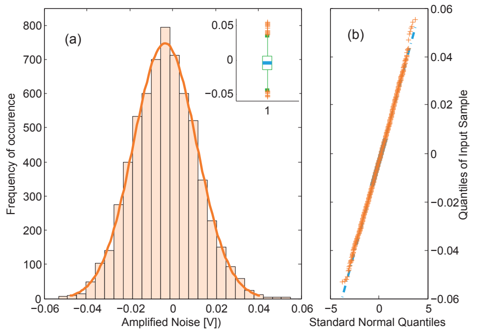


Fig. 4: Amplified Noise Distribution.

IV. Simulation Results

In this section, we first present the noise related figures, followed by the proposed system probabilistic output validation. The probability modulated TRNG system was implemented in a commercial CMOS 65nm Low Power technology. The figures of merit were obtained at nominal operating conditions ($V_{DD} = 1.2V$ and $t = 27°C$) using Cadence Spectre.

As noise sources, we used two minimum size NMOS transistors, which generate additive white Gaussian noise with a RMS value of $461.9\mu V^2$. The noise is then amplified, achieving an output RMS value of $22.14mV^2$. It is important to consider that the main task of the proposed design is to grant the output generated sequences the desired statistical properties, and not to achieve a certain characteristic for the amplifier. In consequence, we omit subsequently the amplifier characteristics and focus on the statistical properties. Figure 4 depicts the noise distribution at the amplifier output, and its corresponding linear quantile-quantile plot, in order to enable a quantitative and visual assessment of whether or not the

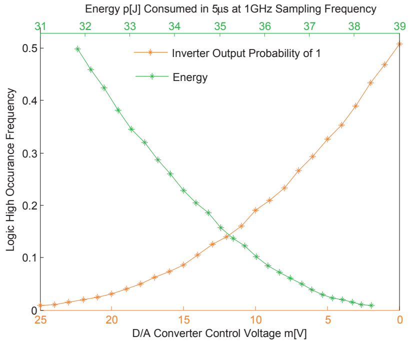


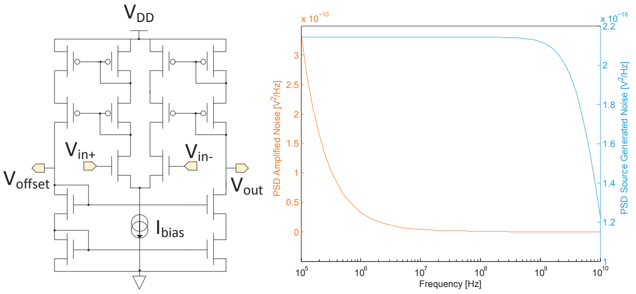
Fig. 5: Voltage Controlled Probability.

noise distribution is sampled from a Gaussian distribution. As far as the inverter is concerned, the measured cut-off frequency, f_{-3dB}, above which the inverter behaves as a low pass filter and attenuate the noise, is $1.87GHz$ for a Fan Out of 1 ($FO1$) load. Thus, to ensure the noise is not filtered, but propagated to the inverter output, for experimental purposes we employed a $1GHz$ sampling frequency of the inverter output. The power spectral density of both the source generated noise, and the amplified noise are depicted in Figure 6b. As one can observe, at a sampling frequency of $1GHz$, the noise spectrum is flat, which implies that the noise samples are uncorrelated, as desired. The probability of a bit “1” at the inverter output is plotted in Figure 5 against the D/A control voltage. The probability of a logic “1” at the inverter output is determined as the ratio between the number of logic “1” voltage samples and the total number of samples. For each probability we measured a total number of 2000 samples. The inverter probability granularity is set by the resolution of the D/A converter and the register that drives it. In Figure 5, one can also observe the total energy consumed by the inverter during a time interval of $5\mu s$ (for 1GHz sampling frequency), which increases with the probability of “1”.

As concerns the randomness validation, the US National Institute of Standards and Technology (NIST) $800 - 22$ test suite [4], is the typical bona fide approach. It comprises 15 tests which check different non-randomness kinds that could exist in a sequence. However, these tests reflect the performance of random sequences with equiprobable bits, and not of probability biased sequences, which renders them relevant in our case only when the bias probability is 0.5. The tests pass criteria are determined by both the sequences pass proportion as a function of their length (10 sequences of 100Kbit), and by the significance level ($\alpha = 0.01$) (i.e., about 1% of the binary sequences are expected to be nonrandom). The sequences were tested and passed the Frequency, Block-Frequency, CumulativeSums, Runs, and FFT tests, which quantitatively confirms that further post-processing of the bit sequences is not required any longer as in the general case in order to eliminate deterministic artifacts.

V. Conclusions

We have proposed a low cost, low energy, probability modulated RNG circuit, implemented in CMOS 65nm process, which is able to provide bit samples at 1GHz, of good statistical quality at low cost, and small area. The results reveal that, for a desired probability of 0.5 there is no crosscorrelation between different generated bit sequences, and that the considered NIST tests are passed.



(a) (b)

Fig. 6: a) Amplifier; b) Source Generated and Amplified Noise Power Spectral Density.

Acknowledgment

This work was supported by the Seventh Framework Programme of the European Union, under Grant Agreement number 309129 (i-RISC project).

References

[1] A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone, *Handbook of Applied Cryptography.* CRC Press, 1996.

[2] R.J. Baker, “Stochastic Computing Systems.” in *Advances in Information Systems Science, Springer*, 1969, pp. 37–172.

[3] K.V. Palem, L.N. Chakrapani, B.E.S. Akgul, and P. Korkmaz, “Realizing Ultra Low-Energy Application Specic SoC Architectures through Novel Probabilistic CMOS (PCMOS) technology.” in *International Conference on Solid State Devices and Materials*, 2005.

[4] http://csrc.nist.gov/groups/ST/toolkit/rng/index.html.

[5] W.T. Holman, J.A. Connelly, and A. Dowlatabadi, “An Integrated analog/digital random noise source.” in *IEEE Transactions on Circuits and Systems, vol. 44*, 1997.

[6] L. Letham, D. Hoff, and A. Folmsbee, “A 128K EPROM using Encryption of Pseudorandom Numbers to Enable Read Access.” in *IEEE Journal on Solid State Circuits, vol. 21, no. 4*, 1986, pp. 881–889.

[7] B. Sunar, W. Martin, and D. Stinsong, “A Provably Secure True Random Number Generator with Built-In Tolerance to Active Attackss.” in *IEEE Transactions on Computers, vol. 56*, 2007, pp. 109–119.

[8] D. Holcomb, W. Burleson, and K. Fu, “Power-Up SRAM State as an Identifying Fingerprint and Source of True Random Numbers.” in *IEEE Transactions on Computers, vol. 58*, 2007, pp. 1198–1210.

[9] P.K. Gupta, and R. Kumaresan, “Binary Multiplication with PN Sequences.” in *IEEE Trans. Acoustics Speech Signal Process. (36)*, 1988, pp. 603–606.

[10] K. Palem, “Energy Aware Algorithm Design via Probabilistic Computing: From Algorithms and Models to Moore’s Law and Novel (Semiconductor) Devices.” in *International Symposium on Verifcation (Theory and Practice)*, 2003, p. 524.

[11] S. Cheemalavagu, P. Korkmaz, K.V. Palem, B.E.S. Akgul, and L.N. Chakrapani, “A Probabilistic CMOS Switch and its Realization by Exploiting Noise.” in *IFIP International Conference on VLSI*, 2005.

[12] J. Kim, and S. Tiwari, “Inexact Computing using Probabilistic Circuits: Ultra Low-Power Digital Processing.” in *ACM Journal on Emerging Tehcnologies in Computing Systems, vol. 10, no.2*, 2014, pp. 16–23.

[13] P. O’Connor, G. De Geronimo, “Prospects for Charge Sensitive Amplifiers in Scaled CMOS.” in *Nuclear Instruments and Methods in Physics Research, 480*, 2001, pp. 713–725.

[14] R.J. Baker, *CMOS Circuit design, Layout, and Simulation.* John Wiley and Sons, Inc., 2010.

[15] N. Shrivastava, G. Bhargava, D.S. Ajnar, and P.K. Jain, “Design of Operational Transconductance Amplifier for Biquad Filter Applications in 0.18m Technology.” in *International Journal of Engineering Research and Applications*, 2012, pp. 562–565.