NOLTA 2006
11-14 september
Bologna - ITALY

Entropy Enhancement in a Chaos-Based Random Bit Generator

T.Addabbo†, M.Alioto†, A.Fort†, S.Rocchi† and V.Vignoli†

† Dept. of Information Engineering, University of Siena
Via Roma, 56 - Siena 53100, Italy
Email: vignoli@dii.unisi.it

Abstract—In this paper the theoretical bases to design a true random bit generator circuit with a predefined minimum entropy are discussed. The approach is tailored to generators based on the Sawtooth chaotic map, and it is based on a feedback control system designed for the optimization of the system parameters. The design approach was validated by an hardware prototype implemented on a Field Programmable Analog Array.

1. Introduction

An ideal True Random Bit Generator (TRBG) is a discrete time memoryless binary source with unbiased symbols that presents a Shannon Entropy of 1 bit/time-step. In practical TRBG circuit implementations the entropy, which is a quantity strictly dependent on the unpredictability and the randomness degree of the generated sequences, is lower than 1 bit/time-step. Since the entropy level can not be increased by means of deterministic post-processing algorithms without decimation or compression (i.e. without lowering the TRBG throughput), design approaches that achieve an adequate entropy at full throughput are of practical interest [1]. In most of the chaos-based TRBGs (see e.g. [1]-[3]), the theoretical knowledge about the chaotic system information generation mechanism allows for setting the nominal entropy arbitrarily close to the ideal maximum Shannon limit [4]. Nevertheless, the behavior of a chaotic circuit is in general highly sensitive to the tolerances of the implementation process, and therefore the actual entropy of the TRBG circuit often results lower than its designed nominal value [4]. In this paper, a feedback strategy to enhance the entropy of TRBG circuits based on the Sawtooth map that preserves the circuit full throughput is proposed [1], [4].

2. Chaotic TRBG circuit model

Let us consider the dynamical system (Sawtooth map)

$$x_{n+1} = \begin{cases} Bx_n + A, & \text{if } x_n < 0 \\ Bx_n - A, & \text{if } x_n \geq 0, \end{cases} \tag{1}$$

where $x \in \mathbb{R}$, $A, B \in \mathbb{R}^+$. As discussed in [2], a possible implementation of (1) requires five different ideal analog blocks: one comparator which evaluates the sign of x_n, one gain block B, one adder, one $\pm A$ constant generator (driven by the comparator digital output S_n), and one delay block. In [1], [4] it is shown that a random bit sequence b_n can be obtained by just picking out the binary sequence defined by the comparator output variable S_n, that assumes the logic value '1' if $x_n > 0$, and '0' otherwise. Since parameter A just represents a scaling factor, the stochastic properties of the system generating b_n, hereafter indicated as TRBG, are dependent only on the parameter B, which must assume values greater than 1 for the system to be chaotic and not greater than 2 to avoid the state x being attracted to $\pm\infty$. For $B \in [\sqrt{2}, 2]$ the two unstable fixed points $\pi_1 = A/(1-B)$ and $\pi_2 = A/(B-1)$ define an interval $I \subset \mathbb{R}$ such that any initial state $x_0 \in I$ triggers a sequence x_n eventually attracted into the chaotic attractor $\Lambda = [-A, A) \subseteq I$. Under these hypotheses, different binary sequences are related to different initial conditions, and their generation probabilities are related to the state invariant probability density function (pdf) of the chaotic dynamical system [2],[1]. When $B = 2$ the system state x pdf in (1) approaches a stationary invariant uniform distribution over $[-A, A)$.
The amount of information supplied by a TRBG and its grade of redundancy are expressed by the Average Shannon Entropy (ASE) defined in what follows. Let be $O_n = \{b_0, \ldots, b_n\}$ an n-bit length generated sequence, where $b_n \in \{0, 1\}$. If $P(O_n)$ is the generation probability of the sequence O_n, the ASE in bit/time-steps is given by

$$\text{ASE} = \lim_{n\to\infty} \left(-\frac{1}{n} \sum_{O_n} P(O_n) \log_2 P(O_n) \right) \tag{2}$$

where the summation extends over all the generable sequences and the limit is approached from above [2]. Kocarev and Stojanovski showed in [4] that the ASE of the TRBG is an increasing function of the parameter B, and that it reaches the upper ASE limit of 1 bit/time-step for $B = 2$. This result focuses the trade-off imposed to the TRBG circuit designer: to avoid the B value being greater than 2 (also accounting for the actual circuit tolerances) while trying to achieve a B value as close as possible to this limit in order to maximize the ASE.

2.1. Modeling the non-idealities

In Fig. 1 a modified block diagram that accounts for the most important non-idealities of circuits implementing (1) is shown, and in (3) the corresponding modified map is

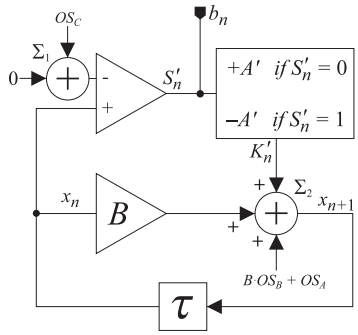


Figure 1: Block diagram implementing (1) with main non-idealities included.

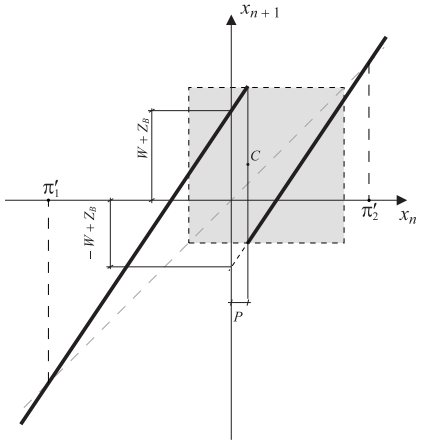


Figure 2: The map described by (4)

reported:

$$f'(x_n) = \begin{cases} B\,(x_n + OS_B) + A' + OS_A, & \text{if } x_n < OS_C \\ B\,(x_n + OS_B) - A' + OS_A, & \text{if } x_n \geq OS_C. \end{cases} \tag{3}$$

In Fig. 1 and in (3) the quantities A' , OS_A, OS_B, and OS_C have the following meaning:
A', OS_A: any actual implementation of the $\pm A$ costant in (1) is not perfectly symmetric, and issues two quantities $+A_1$ and $-A_2$; A' is defined as $(A_1 + A_2)/2$, whereas OS_A is defined as $(A_1 - A_2)/2$.
OS_B: any actual implementation of the gain block B suffers from non-linearities and the presence of offsets. Non-linearities can be reduced by proper analog design techniques to non-influent values [2]. Only the gain block offset, OS_B, was therefore taken into account in the model.
OS_C: this quantity accounts for the offset of the comparator, which modifies its threshold to OS_C. Therefore the comparator output S'_n assumes the logic value '1' if $x_n \geq OS_C$, and '0' otherwise.

Equation (3) can be rearranged as

$$x_{n+1} = \begin{cases} Bx_n + W + Z_B, & \text{if } x_n < P \\ Bx_n - W + Z_B, & \text{if } x_n \geq P, \end{cases} \tag{4}$$

where $W = A'$, $Z_B = B \cdot OS_B + OS_A$, and $P = OS_C$. The parameters (W, Z_B, P) determine the map scale factor and the map position in the plane (x_n, x_{n+1}), while the parameter B sets the slope of the map (Fig.2).

2.2. Equivalent dynamical systems

In this paper two different dynamical systems (4) are said *equivalent* if they define two TRBGs characterized by the same stochastic properties (i.e. for each binary sequence O_n the generation probability $P(O_n)$ is equal for both of the TRBGs). From this point of view, it has been proved by the authors in [2] that two systems (4) with a same B value are equivalent (and therefore the statistical properties of the output bit sequences $\{b_n\}$ are equal) if they have the same *central distance* D, where D is defined as

$$D = \frac{P(B-1) + Z_B}{W\sqrt{2}}. \tag{5}$$

Referring to Fig.2, the absolute value of D is equal to the Euclidean distance of the point $C = (P, BP + Z_B)$, in the plane (x_n, x_{n+1}), from the bisector of the first and third quadrants, normalized to W. Accordingly, for any given B value, systems described by (4) are equivalent to the following equivalence class representative dynamical system:

$$x_{n+1} = \begin{cases} Bx_n + 1 + D\sqrt{2}, & \text{if } x_n < 0 \\ Bx_n - 1 + D\sqrt{2}, & \text{if } x_n \geq 0. \end{cases} \tag{6}$$

Moreover, given B, systems (1) and (4) are equivalent (and therefore the corresponding TRBGs have the same ASE) if the central distance D of (4) is zero, that is if

$$P(B-1) = -Z_B. \tag{7}$$

3. Enhancing the ASE: map correction

When the central distance D is not zero, to avoid the state x of (4) to jump beyond the fixed points π'_1 and π'_2 (Fig.2) and to be attracted toward infinity (with a consequent electronic saturation in practical implementations), B must satisfy the relationship

$$1 < B \leq \frac{2}{1 + |D|\sqrt{2}} \leq 2. \tag{8}$$

According to (8), the quantity $|D|$ determines an upper bound for the parameter B which fixes the rate of mixing of the chaotic map [1], and and if D tends to zero, the maximum allowed B value tends to 2. Therefore, as discussed in [2], a non-zero central distance D determines also an upper bound for the ASE achievable by the TRBG. This point has

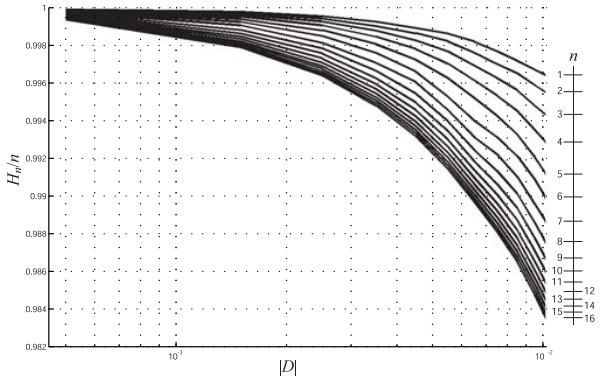


Figure 3: Upper bound for the partial ASE H_n/n (in bit/time-step) versus $|D|$, computed up to $n = 16$.

been investigated by the authors on the basis of the computation of the asymptotic invariant pdf of the dynamical system (6) (Fig. 3).

The ASE maximization procedure proposed in this paper is based on the fact that the two parameters D and B are sufficient for determining the stochastic properties of the TRBG. First of all, from (7), if B is kept constant, D can be zeroed by changing either the parameter P or the parameter Z_B, that is, referring to Fig.1, by injecting just one correction offset either in the adder Σ_1 or in the adder Σ_2. Once minimized D with a further correction B can be increased according to (8).

In practical systems the required exact values of the correction signal for D and B are not available, since the quantities OS_A, OS_B, and OS_C are not known. Nevertheless, the values of the correction signals for D and B can be obtained from the statistical analysis of the generated binary sequences, as it was discussed by the authors in [2]. In particular, if $P(b_n = 1)$ is the probability for the n-th binary output to take the '1' logic value, and if $\mu = (P(b_n = 1) - 0.5)$ for $n \to +\infty$, the authors showed that for each B value the parameter μ is a monotonic function of D, and that $\mu = 0$ when $D = 0$. As a consequence, the sign estimation of μ can be used to mimize $|D|$ [2]. Anyway, since the upper ASE limit of 1 bit/time-step is reached when $B = 2$ (Fig. 3), and due to the upper limit for B given by (8), the ASE is maximized if both $D \to 0$ and $B \to 2$. Summarizing, the simplest way to drive system (4) to a desired operating point $(B_{\max}, |D|_{\min})$, where $B_{\max}$ is the maximum value of B according to (8) given $|D|_{\min}$, is to iteratively minimize the central distance magnitude $|D|$, exploiting the estimation of the μ sign, and then increase B until the divergence of the state orbit is detected [2].

4. Correction signals resolution and worst case ASE

As discussed in the previous section, two correction signals S_B and S_D are sufficient for correcting the map: the former signal is used for changing the amplification B of the gain block in Fig.1, whereas the signal S_D is used to minimize D. Even if the correction technique described above may be used to drive system (4) to any desired operating point $(B_{\max}, |D|_{\min})$ arbitrarily close to the optimum point (2, 0), in practical realization the precision of the correction system is limited by the physical resolution of signals S_D and S_B. Accordingly, by denoting with δ_B and δ_D the minimum variations effectively achievable on gain B and on signal S_D respectively, a worst case operating point $(B_{\max}, |D|_{\min})_{wc}$ can be estimated as a function of δ_B and δ_D. In particular, when injecting the correction signal S_D in node Σ_1 (a similar analysis can be carried out considering the injection of S_D in node Σ_2), the map (3) must be modified as

$$f'(x_n) = \begin{cases} B\,(x_n + OS_B) + A' + OS_A, & \text{if } x_n < OS_C + S_D \\ B\,(x_n + OS_B) - A' + OS_A, & \text{if } x_n \geq OS_C + S_D. \end{cases} \quad (9)$$

According to (5), since D for the system (9) is equal to

$$D = \frac{OS_C(B-1) + OS_A + B \cdot OS_B + S_D(B-1)}{A\sqrt{2}}, \quad (10)$$

D is zeroed if

$$S_D = \frac{OS_C(B-1) + OS_A + B \cdot OS_B}{1-B}. \quad (11)$$

In the worst case it can be assumed that, after the $|D|$ minimization, the S_D actual value differs from (11) by the quantity $\delta_D/2$, and we can write

$$|D|_{\min} \leq \frac{(B-1)\delta_D}{2\,A'\sqrt{2}} = D_0. \quad (12)$$

If, once minimized $|D|$, also the B value is changed by S_B, the value of D changes due to (5). In detail, if B is increased by δ_B, (12) is modified according to

$$|D|_{\min} \leq D_0 + \frac{\delta_B}{A'\sqrt{2}}\left(\frac{|OS_B|_{\max} + |OS_A|_{\max}}{B-1} + \frac{\delta_D}{2}\right). \quad (13)$$

The upperbounds expressed in (8) and (13) represent the worst case limit for the control system, and in practical realizations, even if performing the correction, a worste case operating point $(B_{\max} < 2, |D|_{\min} > 0)_{wc}$ must be taken into account. This limit, that depends on the resolution of the correction signals, identifies the minimal guaranted ASE for the corrected TRBG.

5. Experimental results and conclusions

The proposed correction procedure, described in detail in [2], was tested on a prototype TRBG circuit implemented by a Field Programmable Analog Array (FPAA Anadigm AN212E04). This SRAM based device can be dynamically reconfigured by an host processor while the old configuration is still active and running. The activation

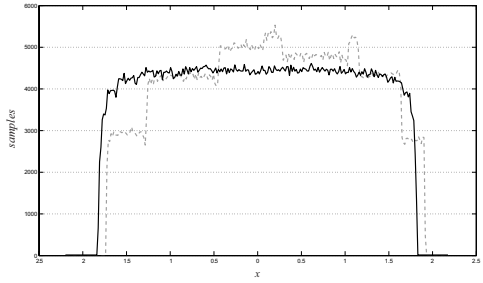


Figure 4: Histograms of two typical 10^6 sample analog output sequences collected from the prototype circuit implementing the Sawtooth map (1): before (dashed-gray) and after (solid-black) correction.

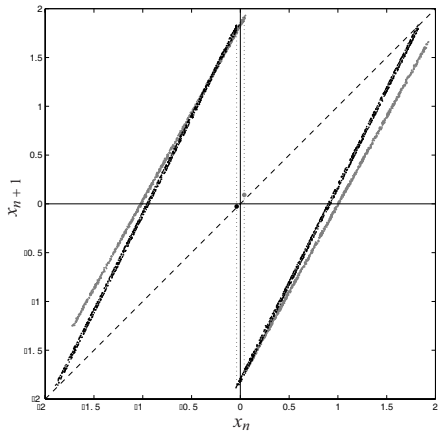


Figure 5: Plot of the implemented map before the correction (gray) and after the correction (black) obtained on the basis of 2,000 samples taken from the same sequences used for Fig. 4.

of the new configuration happens in real time, allowing for an easy implementation of the proposed ASE maximization strategy. The gain block B in Fig.1 was implemented using a FPAA library adjustable gain block, whose profile of gain versus control voltage was specified to linearly vary within the nominal range $1.4 \div 2.08$ with an experimentally verified resolution $\delta_B \approx 2 \cdot 10^{-2}$ for the working frequency of 250 kHz. The control and correction procedure for the ASE maximization was software implemented on a PC equipped with an acquisition board (National PCI MIO64E) for acquiring the TRBG output, and whose 12 bit D/A converters, with a proper scaling, provided the control signals S_B and S_D. The correction system was tested using an experimentally verified resolution $\delta_D \approx 5 \cdot 10^{-3}$, while setting for the map parameter A in (1) an equivalent voltage nominal value of $1.8V$. With the aim of emulating different combined effect of OS_A, OS_B and OS_C (that depend on the specific circuiit implementing the map), exploiting the finest resolution of the PCI MIO64E device at disposal, several *initial operative conditions* were forced to the prototype by adding random offsets to the correction signals [2]. Partial ASE values as high as 0.99 bit/time-step (for word lengths up to 16 bit) were experimentally achieved at the maximum allowed speed of 250 kHz, without performing any kind of digital post-processing on the output bit sequences. This result agrees with the theoretical worst case ASE obtainable considering Fig. 3 and the upper bounds (8) and (13), evaluated referring to the typical value of OS_A, OS_B and OS_C obtainable from the FPAA data sheet [2]. In Fig. 4 and 5 the effect of the correction procedure on the actually implemented map is shown. In each tested case, it has been estimated that after the correction the prototype was working with B values not lower than 1.99, obtaining an invariant distribution of the state over the chaotic attractor very close to the ideal uniform distribution theoretically achievable for $B = 2$. These results were obtained with resolutions of the correction signals S_B and S_D lower than the actual resolution normally achievable in accurate-designed integrated analog circuits.

Without post-processing, the proposed generator was able to issue sequences passing the FIPS 140.2 tests with a success rate higher than 98%, while typically failing only 4 of 18 Diehard tests [2]. On the other hand, with the direct bit-by-bit XOR with the output of a 8 bit Linear Feedback Shift Register, the sequences of the proposed TRBG passes all of the considered tests at full throughput.

The effectiveness of the proposed technique, proved from a theoretical point of view, is not tied to the specific implementation presented in this paper, and can be successfully applied also to integrated designs with an increase in the output bit rate of up to two order of magnitude, with state of the art CMOS technologies, with respect to the FPAA implementation.

References

[1] T. Stojanovski and L. Kocarev, "Chaos based random number generators Part I: Analysis", *IEEE Tran. CAS I*, vol. 48, no. 3, pp. 281–288, Mar. 2001.

[2] T. Addabbo, M. Alioto, A. Fort, S. Rocchi, V. Vignoli, "A Feedback Strategy to Improve the Entropy of a Chaos-Based Random Bit Generator", *IEEE Tran. CAS I*, Vol. 53, No.2, Feb. 2006.

[3] M.E. Yalçin, J.A.K. Suykens and J. Vandewalle, "True Random Bit Generation From a Double-Scroll Attractor", *IEEE Tran. CAS I*, vol. 51, no. 7, pp. 1395–1404, July 2004.

[4] T. Stojanovski, J. Pihl and L. Kocarev, "Chaos based random number generators Part II: Practical Realization", *IEEE Tran. CAS I*, vol. 48, no. 3, pp. 382-385, Mar. 2001. random noise generator", *Analog Integr. Circuits Signal Process.*, vol. 34, pp. 97105, 2003.