A Stochastic Algorithm to Design Min-Entropy Tuning Controllers for True Random Number Generators

Tommaso Addabbo, *Member, IEEE*, Ada Fort, *Member, IEEE*, Riccardo Moretti, *Member, IEEE*, Marco Mugnaini, *Member, IEEE*, Duccio Papini, and Valerio Vignoli, *Member, IEEE*

***Abstract*— We discuss a stochastic algorithm to design tuning controllers for cryptographic True Random Number Generators, compliant to NIST recommendations, as an effective low-complexity solution to counteract entropy variability in integrated architectures implementing tunable entropy sources. Taking as a reference the min-entropy concept, we discussed the proposal from both the theoretical and hardware design points of view, validating claims with proofs and experiments. Depending on the target accuracy, the proposed architecture is scalable, and its profitable use in TRNG design strongly depends on the kind of core entropy sources taken into account. Furthermore, we show that the low-complexity entropy measurement techniques exploited in this proposal can be used to design a legitimate alternative to the Adaptive Proportion Health Test recommended in the NIST 800.90B publication.**

***Index Terms*— True random number generators, cryptography, entropy sources, statistical testing.**

I. Introduction

INTEGRATED True Random Number Generators (TRNGs) are integrated circuits devised to generate sequences of truly random bits. TRNGs apply in different Information Technology (IT) fields, including Cryptography and Information Security, in which they are extensively used, e.g., in the initialization of cryptographic protocols [1]–[3]. The security of cryptographic algorithms using random numbers is critically related to the degree of unpredictability of TRNGs, that represent sensitive components subject to severe design constraints. In this regard, the U.S. National Institute of Standards and Technology (NIST) has produced a set of publications providing guidelines and recommendations for the design and verification of cryptographic TRNGs, widely adopted and referred to in literature [2]–[4].

Manuscript received October 9, 2021; revised December 24, 2021 and January 26, 2022; accepted February 1, 2022. This article was recommended by Associate Editor C. K. Ahn. *(Corresponding author: Tommaso Addabbo.)*

Tommaso Addabbo, Ada Fort, Riccardo Moretti, Marco Mugnaini, and Valerio Vignoli are with the Department of Information Engineering and Mathematics, University of Siena, 53100 Siena, Italy (e-mail: addabbo@dii.unisi.it).

Duccio Papini is with the Department of Mathematics, Computer Science and Physics, University of Udine, 33100 Udine, Italy.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCSI.2022.3151794.

Digital Object Identifier 10.1109/TCSI.2022.3151794

The randomness of a cryptographic TRNG is originated from the core entropy source S, that is, typically, a mixed-signal circuit around which different deterministic encryption/compression post-processing digital algorithms are carefully designed to make the final stream meet both adequate entropy levels and statistical compliance with NIST recommendations [1]–[3], [5]–[12]. In general, the lower is the entropy of S, the less efficient the TRNG is, in terms of throughput (information bit/s).

Circuit fabrication process variability, aging and minor hardware failures, circuit sensitivity to temperature and supply voltage variations are among the chief causes of entropy static or dynamic degradation in integrated entropy sources [13]–[20]. In most cases, to gain control of the entropy, several researchers proposed to include the TRNG core in a feedback loop, in which a monitoring task supervises the TRNG operation, analyzes its generated stream, and tunes the entropy core resorting to different technical solutions (e.g., by adjusting voltage/current offsets, propagation delays, sampling/clock frequencies [13]–[29]). In these TRNGs, adopting different strategies at the low-level design, the proposed solutions rely on *tunable core entropy sources*. In literature, depending on the solution, the tuning/controlling algorithm has been designed combining theoretical analysis, heuristic considerations, exhaustive numerical simulations and experiments [13]–[27].

Depending on the design of the TRNG core, the technical relation between entropy and tuning/controlling parameters can be strongly dependent on the implementation. This happens, e.g., in some fully digital TRNGs combining complex oscillators and metastable circuits, in which process-voltage-temperature variations can play relevant roles [13]–[17], [30]–[32]. In these solutions, finite parametric spaces are inspected searching for an optimum setup, according to different optimization criteria, that in the worst case agree with exhaustive investigations of the entire parametric space [13], [14], [30], [33]. From a theoretical point of view, this problem is equivalent to the selection of the best entropy source within a set of available ones. Within this technical framework, the study of low-complexity hardware techniques for entropy estimation is of interest [1]–[3], [22], [26], [30].

Referring to Fig. 1, in this work we discuss the design of a generic low-complexity hardware tuning controller based on the *min-entropy* concept given in the NIST publication 800-90B [2]. In detail, we discuss a design approach suitable for those integrated TRNGs in which the entropy source can be varied according to a finite set of parametric values. Adopting a well-defined theoretical framework, our proposal is based on estimation methods exploiting low-complexity entropy measurement techniques, and has a generic validity. In this context, the goal of the Tuning Controller in Fig. 1 is to select the best S_{best} among the set $\{S_1, \ldots, S_Z\}$ of entropy sources corresponding to, e.g., Z different parametric configurations of a same hardware core entropy source. Which one is the best depends on the adopted figures of merit, as made clearer in next sections.

This work is organized as in the following. In Sec. II and III we introduce the notation and fundamental theoretical results that justify the technical solutions proposed in this work. In detail, in Secs. III-A and III-B we present original theoretical results that are used to design a stochastic algorithm aiming to solve the Tuning Controller problem, presented in Sec. III-C. The algorithm hardware design, implementation and testing, with experiments, are discussed in Sec. IV. The experiments were specifically designed to test the capability of the Tuning Controller to select the best source S_{best}, considering different algorithmic/hardware complexities and introducing adequate performance evaluation metrics. To take adequate control of the test bench, in Sec. IV-A we first investigated the operation of the Tuning Controller when applied to an artificial Markov stochastic binary source, referring to precise theoretical links between the entropy and tuning parameters. In Sec. IV-B we also repeated the analysis considering parametric entropy sources implemented in FPGA, based on low-complexity Digital Nonlinear Oscillators (DNOs). Finally, in Sec. IV-C we show with theoretical arguments that the low-complexity entropy measurement techniques exploited by the Tuning Controller in Fig. 1 can be used to design a legitimate alternative to the Adaptive Proportion Health Test recommended by NIST [2]. Conclusion and Reference close the paper.

II. Entropy, Min-Entropy, Worst-Case and Best-Case Entropies

The information generation rate of a generic ergodic source S generating symbols taken from an alphabet of N symbols $\mathcal{A} = \{s_1, \ldots, s_N\}$ corresponds to the Average Shannon Entropy (ASE), expressed in bit per symbol (i.e., information bit per generated random symbol [bit/symb]), defined as

$$\mathcal{H}(S) = \lim_{k\to\infty} \mathcal{H}^k(S) = \lim_{k\to\infty} -\frac{1}{k} \sum_{i=0}^{N^k-1} P(w_i) \log_2 P(w_i). \quad (1)$$

In (1), the summation extends to the entire set of words w_i, made of k-tuplets of symbols having positive generation probability. In the special case of i.i.d. symbols, eq. (1) agrees with the well known Shannon entropy $\mathcal{H}(S) = -\sum_{i=1}^{N} p_i \log_2 p_i$, being $\mathbf{P} = (p_1, \ldots, p_N)$ the symbols generation probabilities.

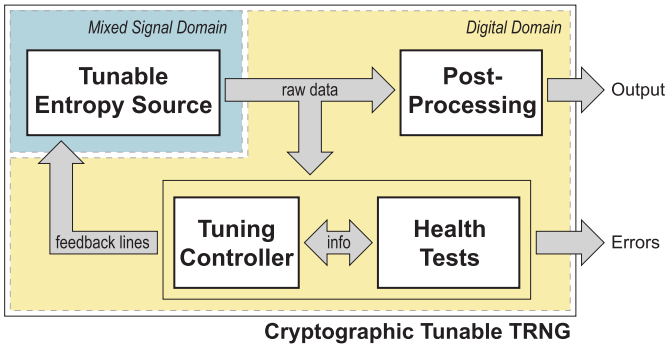

Fig. 1. The proposed architecture of a cryptography TRNG with tunable entropy source.

In any case, the result of (1) ranges between 0 and $\log_2 N$ bit/symb.

With few exceptions, for most entropy sources the estimation (or the measurement) of (1) is unfeasible, and in literature a number of methods have been proposed to calculate approximated estimations based on finite-time observations [2], [3].

Aiming to introduce an operational method to evaluate entropy sources, the NIST publication 800-90B introduces the *min-entropy* concept, defined as a conservative measurable lower-bound for the source entropy [2]. In other words, the min-entropy $\mathcal{H}_{\text{m}}(S)$ for a physical information source S can be operationally understood as a *measurable* information generation rate such that, *with reasonably high probability*, the average amount of information per symbol, issued by the source, is greater than $\mathcal{H}_{\text{m}}(S)$.

In specific theoretical cases, when the stochastic model of a source is completely known, precise min-entropy expressions can be given. For example, in [2] the average min-entropy of an ergodic source generating i.i.d. symbols, taken from an alphabet of N elements with generation probabilities $\mathbf{P} = (p_1, \ldots, p_N)$, is defined as

$$\mathcal{H}_m(S) = -\log_2 p_H, \quad \text{[bit/symb]}, \quad (2)$$

where $p_H = \max_i p_i \in \mathbf{P}$, i.e., p_H is the maximum generation probability among the N symbols. By the way, let us notice that, in general, more than one symbol can have generation probability p_H.

As it can be appreciated from (2), $\mathcal{H}_m(S)$ is a monotonic decreasing function with p_H. Its maximum value, equal to $\log_2 N$ bit/symb, is obtained for $p_H = 1/N$, i.e., in case of uniform probability distributions. Furthermore, it is worth noting that there is an infinite set of different sources with N symbols sharing a same min-entropy level. For instance, all the generation probabilities $\mathbf{P}$ having p_H as maximum value provide the same result in (2).

In practical cases, from the operational point of view, in [2] the min-entropy is estimated with statistical estimators applied to the raw data sequences collected from the entropy source core, as in Fig. 1.

We conclude this section introducing the definition of worst-case and best-case entropies.

Definition 1: Given an arbitrary set Ω of entropy sources, we define the worst-case average entropy and the best-case average entropy in Ω as

$$\mathcal{H}_{WC}(\Omega) = \inf_{S\in\Omega} \mathcal{H}(S), \quad \mathcal{H}_{BC}(\Omega) = \sup_{S\in\Omega} \mathcal{H}(S), \tag{3}$$

respectively.

Given a source $S \in \Omega$, the relations between the min-entropy $\mathcal{H}_m(S)$, $\mathcal{H}_{\mathrm{WC}}(\Omega)$ and $\mathcal{H}_{\mathrm{BC}}(\Omega)$ depend on both the operational (or theoretical) definition of $\mathcal{H}_{\mathrm{m}}(S)$ and the stochastic sources in Ω. As shown in the next Section, in specific cases these relations can be expressed theoretically.

III. A Stochastic Algorithm to Design TRNG Tuning Controllers

We discuss the concept design of a low-complexity stochastic algorithm aiming to solve the Tuning Controller problem, that is to select the best entropy source S_{best} in a set of entropy sources $\Omega = \{S_1, \ldots, S_Z\}$. To investigate the proposal within a tractable theoretical framework, if not otherwise stated, in the following subsections we focus on ergodic stochastic sources S of i.i.d. symbols. We denote with $\mathbb{M}(N) \subset (0,1)^N \subset \mathbb{R}^N$ the set of the probability mass functions expressing different generation probabilities for the N symbols in $\mathcal{A}$. Accordingly, if $\mathbf{P} = (p_1, \ldots, p_N) \in \mathbb{M}(N)$, we have $\sum_{i=1}^{N} p_i = 1$ and $0 < p_i < 1$. Furthermore, we denote with $p_H = \max_i p_i \geq \frac{1}{N}$ the maximum generation probability of S.

A. Comparison of Entropy Sources: The Min-Entropy Approach

Given a finite set Ω of entropy sources, we propose to solve the Tuning Controller problem by comparing their min-entropies, as defined in (2). From a mathematical point of view, a partial order in Ω is given by the

Definition 2: Given two entropy sources $S_1, S_2 \in \Omega$,

$$S_1 \leq S_2 \Leftrightarrow \mathcal{H}_m(S_1) \leq \mathcal{H}_m(S_2) \Leftrightarrow p_{H_2} \leq p_{H_1}, \tag{4}$$

where p_{H_1}, p_{H_2} are the maximum generation probabilities of S_1, S_2, respectively.

As shown hereafter, the above defined ordering in Ω does not assure to select the source with highest ASE, but it provides a low-complexity sub-optimal effective solution to a difficult problem. Interestingly, given any $\tilde{S}$ in Ω, we can investigate the infinite set $\tilde{\mathcal{C}}$ of sources sharing the same maximum generation probability $\tilde{p}_H$. Accordingly, from the min-entropy point of view, the source $\tilde{S}$ can be considered a representative of an equivalence class $\tilde{\mathcal{C}}$, that is the subset of all possible sources of i.i.d. symbols having min-entropy $\mathcal{H}_m(\tilde{S}) = -\log_2 \tilde{p}_H$. As reported in the Appendix, we could prove the following

Theorem 1: Let $\tilde{S} \in \tilde{\mathcal{C}}$. By defining $F = \lfloor 1/\tilde{p}_H \rfloor \in \mathbb{N}$ and the function $h : (0,1] \to \mathbb{R}$ as $h(x) = -x \log_2 x$, it results

$$\mathcal{H}_m(\tilde{S}) \leq \mathcal{H}_{WC}(\tilde{\mathcal{C}}) \leq \mathcal{H}(\tilde{S}) \leq \mathcal{H}_{BC}(\tilde{\mathcal{C}}), \tag{5}$$

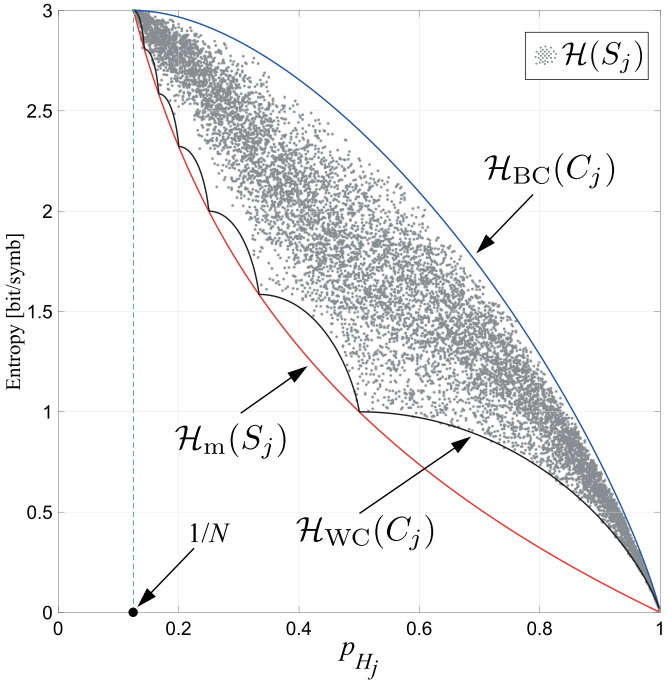

Fig. 2. Numerical verification of Theor. 1 by means of numerical Monte Carlo simulations, for different values of p_{H_j}, considering $N = 8$ (10.000 simulation points).

where

$$\mathcal{H}_{BC}(\tilde{\mathcal{C}}) = (N-1) \cdot h\left(\frac{1-\tilde{p}_H}{N-1}\right) + h(\tilde{p}_H). \tag{6}$$

$$\mathcal{H}_{WC}(\tilde{\mathcal{C}}) = F \cdot h(\tilde{p}_H) + h(1 - F \cdot \tilde{p}_H), \tag{7}$$

In other words, referring to the introduced theoretical framework, the knowledge of $\tilde{p}_H$ for the source $\tilde{S}$ provides precise lower and upper bounds for its Average Shannon Entropy. The above relations (5)-(7) can be inspected by means of numerical Monte Carlo simulations, as shown in Fig. 2. As it can be appreciated, the min-entropy provides a conservative under-estimation of the Shannon entropy, with a relative error that approaches zero for $p_H \to 1/N$. This represents a fundamental result, *from the considered application point of view*, since the better is the best source in Ω, the more reliable is the min-entropy selection criteria.

B. Low-Complexity Maximum Generation Probability Estimation

Directly from (2), the comparison among the different sources is based on the measurements of the generation probabilities p_H for the most-probable symbols. To this aim, hereafter we propose a specific estimator of p_H, involving low-complexity calculations.

Let us focus on the following problem: what is the number T of generation trials that we reach if we stop the experiment as soon as any element $s_j \in \mathcal{A}$ has been generated m times? In literature, this kind of problem is related to waiting-time problems for occupancy in urns [34]. Accordingly, each symbol $s_j \in \mathcal{A}$ can be associated to an urn that is randomly filled by indistinguishable balls with probabilities

$\mathbf{P} = (p_1, \ldots, p_N) \in \mathbb{M}(N)$. After K observations, the occurrences $\mathcal{T}_1, \ldots, \mathcal{T}_N$ of the symbols $s_1, \ldots, s_N$ (note that $\mathcal{T}_1 + \ldots + \mathcal{T}_N = K$) have a well known joint multinomial probability distribution

$$P\left(\cap_{j=1}^{N} \mathcal{T}_j = o_j\right) = K! \prod_{j=1}^{N} \frac{p_j^{o_j}}{o_j!}, \tag{8}$$

and $E\{\mathcal{T}_j\} = Kp_j$. However, if the experiment (generation trials) is stopped *as soon as one predefined symbol* reaches m occurrences, let us assume $\mathcal{T}_1$, the remaining $N-1$ variables $\mathcal{T}_2, \ldots, \mathcal{T}_N$ are distributed according to a *negative multinomial distribution* [34], [35], and their mean values are, for $j = 2, \ldots, N$, $E\{\mathcal{T}_j\} = \frac{p_j \cdot m}{p_1}$.

In the next paragraphs, we adopt the following notation. We denote with T_i the random variable describing the minimum number of generation steps necessary to obtain m occurrences of the symbol s_i. Thus, we have $T_i = \sum_{i=1}^{N} \mathcal{T}_i$ and $\mathcal{T}_i = m$. It is worth noting that if $m \to \infty$ also $T_i \to \infty$, since T_i can not be smaller than m (at least m steps are necessary to generate m occurrences of any symbol, i.e., $P(T_i < m) = 0$). Accordingly, $0 \leq \liminf_{m\to\infty} \frac{m}{T_i} \leq \limsup_{m\to\infty} \frac{m}{T_i} \leq 1$. Regarding the ratio $\frac{m}{T_i}$, the following proposition holds.

Proposition 1: Let $s_i \in \mathcal{A}$ be a symbol with generation probability p_i. If T_i is the minimum number of generation trials such to have s_i generated m times, then the statistics $\{\hat{p}_i(m) = \frac{m}{T_i}, m = 1, 2, \ldots\}$ is a consistent estimator of p_i.

Proof: Since $m > 0$, let us focus on the sequence of random variables $Y_m = \frac{1}{\hat{p}_i(m)} - \frac{1}{p_i} = \frac{T_i}{m} - \frac{1}{p_i}$, for $m = 1, 2, \ldots$. It suffices to show that $\mu_{Y_m} = E\{Y_m\} = 0$ and that the variance of Y_m vanishes with increasing m, i.e., $\lim_{m\to\infty} \sigma^2_{Y_m} = 0$.

The number of trials T_i can be written as $T_i = m + \eta$, where the random variable $\eta = \sum_{i=1, i\neq j}^{N} \mathcal{T}_i$ has a negative binomial distribution with mean value and variance [35]

$$\mu_\eta = \frac{m(1-p_i)}{p_i}, \quad \sigma^2_\eta = \frac{m(1-p_i)}{p_i^2}. \tag{9}$$

As a result, since $\mu_{T_i} = m + \mu_\eta = \frac{m}{p_i}$, we have $\mu_{Y_m} = \frac{E\{T_i\}}{m} - \frac{1}{p_H} = 0$. On the other hand, recalling that $\sigma^2_\eta = E\{\eta^2\} - \mu^2_\eta$ and using (9)

$$\begin{aligned} \sigma^2_{Y_m} &= E\{Y_m^2\} = E\left\{\frac{T_i^2}{m^2} + \frac{1}{p_i^2} - \frac{2T_i}{m\ p_i}\right\} \\ &= \frac{m^2 + 2m\mu_\eta + E\{\eta^2\}}{m^2} - \frac{1}{p_i^2} = \frac{1-p_i}{m\ p_i^2}, \end{aligned} \tag{10}$$

i.e., $\lim_{m\to\infty} \sigma^2_{Y_m} = 0$. ∎

As far as the distributions of the random variables $T_i = m + \eta_i$ is considered, we have the following proposition.

Proposition 2: Let us consider the symbols $s_i, s_j \in \mathcal{A}$, having generation probabilities p_i and $p_j < p_i$, respectively. Let the random variables T_i and T_j represents the minimum number of generation steps to have the symbols s_i, s_j, respectively, counted m times. It results

$$\lim_{m\to\infty} P(T_i \geq T_j) = 0. \tag{11}$$

Proof: By writing $T_i = m + \eta_i$ and $T_j = m + \eta_j$, the limit (11) is proved if we show that the random variable $S = (T_i - T_j)/m = (\eta_i - \eta_j)/m$ satisfies the limit $\lim_{m\to\infty} P(S \geq 0) = 0$. Recalling (9), it results $\mu_S = E\{S\} = (E\{\eta_i\} - E\{\eta_j\})/m = (1-p_i)/p_i - (1-p_j)/p_j = (p_j - p_i)/p_j p_i < 0$. For the variance of S we have that $m^2\sigma_S^2 = \sigma^2_{\eta_i} + \sigma^2_{\eta_j} - 2\mathrm{Cov}(\eta_i, \eta_j) = \sigma^2_{\eta_i} + \sigma^2_{\eta_j} - 2\rho_{\eta_i,\eta_j}\sigma_{\eta_i}\sigma_{\eta_j} \leq \sigma^2_{\eta_i} + \sigma^2_{\eta_j} + 2\sigma_{\eta_i}\sigma_{\eta_j} = (\sigma_{\eta_i} + \sigma_{\eta_j})^2$, where ρ_{η_i,η_j} is the correlation coefficient for η_i and η_j. As a result,

$$\sigma_S \leq \sigma_0 = \frac{(\sigma_{\eta_i} + \sigma_{\eta_j})}{m} = \frac{\sqrt{1-p_i}}{p_i\sqrt{m}} + \frac{\sqrt{1-p_j}}{p_j\sqrt{m}}. \tag{12}$$

By noting that $P(S \geq 0) = P(S - \mu_S \geq -\mu_S) < P(|S - \mu_S| \geq -\mu_S)$ we can exploit the Chebyshev's inequality stating that, for any $k > 1$, $P(|S-\mu_S| \geq k\sigma_0) \leq P(|S-\mu_S| \geq k\sigma_S) \leq 1/k^2$. Indeed, by setting

$$k = \frac{-\mu_S}{\sigma_0} = \frac{\sqrt{m}(p_i - p_j)}{\sqrt{1-p_i} + \sqrt{1-p_j}} \tag{13}$$

in the previous inequalities, we obtain

$$P(S \geq 0) < \frac{\left(\sqrt{1-p_i} + \sqrt{1-p_j}\right)^2}{m(p_i - p_j)^2}, \tag{14}$$

that imples $\lim_{m\to\infty} P(S \geq 0) = 0$. ∎

The above proposition states that, for increasing values of m, the random variables η_i associated to those symbols having large generation probabilities, have increasing probabilities to be among the first to reach m occurrences. This is important when considering a counting experiment that is halted as soon as *any* symbol in $\mathcal{A}$ reaches m occurrences. In this case, the number of generation steps is equal to $T = \min_i T_i = m + \min_i \eta_i$, where the random variables η_i have mean values and variances given in (9).

As a direct result of the previous propositions, the ratio $\frac{m}{T} = \frac{m}{\min_i T_i}$ converges in probability to p_H, i.e., for any $\varepsilon > 0$

$$\lim_{m\to\infty} P\left(\left|\frac{m}{T} - p_H\right| > \varepsilon\right) = 0. \tag{15}$$

As an example, the convergence in probability is represented in Fig. 3, which reports the mean value and standard deviation of $\hat{p}_H(m) = \frac{m}{T}$ estimated on the basis of the numerical investigation of an arbitrary source generating $N = 4$ symbols with probabilities $\mathbf{P} = (0.185, 0.279, 0.291, 0.245)$.

As it can be appreciated from the figure, the estimator $\hat{p}_H(m)$ is affected from a positive bias that vanishes as m increases. As discussed hereafter, both the convergence rate and the bias depend on the generation probability $\mathbf{P}$.

The Fig. 4 reports the statistical distributions of the random variables $\eta_1, \eta_2, \eta_3, \eta_4$ and $\min_i \eta_i$, estimated on the basis of 10.000 randomized experiments ($m = 64$), for a source of $N = 4$ symbols with generation probabilities $\mathbf{P} = (0.185, 0.279, 0.291, 0.245)$. The red curves in the upper plots reports the theoretical frequency distribution of negative binomial random variables, with mean value and variances given in (9). The larger is m, the smaller is the standard deviation of η_i/m, and the more accurate is the estimator $\hat{p}_H(m) = m/T$.

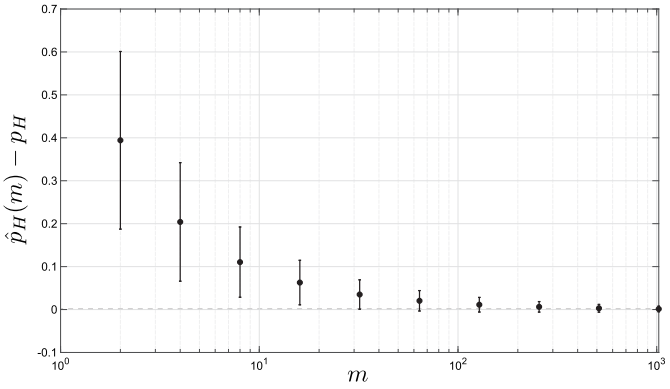

Fig. 3. Convergence in probability of the estimator $\hat{p}_H(m) = \frac{m}{T}$, for a source of $N = 4$ symbols with generation probabilities $\mathbf{P} = (0.185, 0.279, 0.291, 0.245)$, for different values of m. Mean value and standard deviation ($\mu \pm \sigma$) reported, 10.000 randomized experiments.

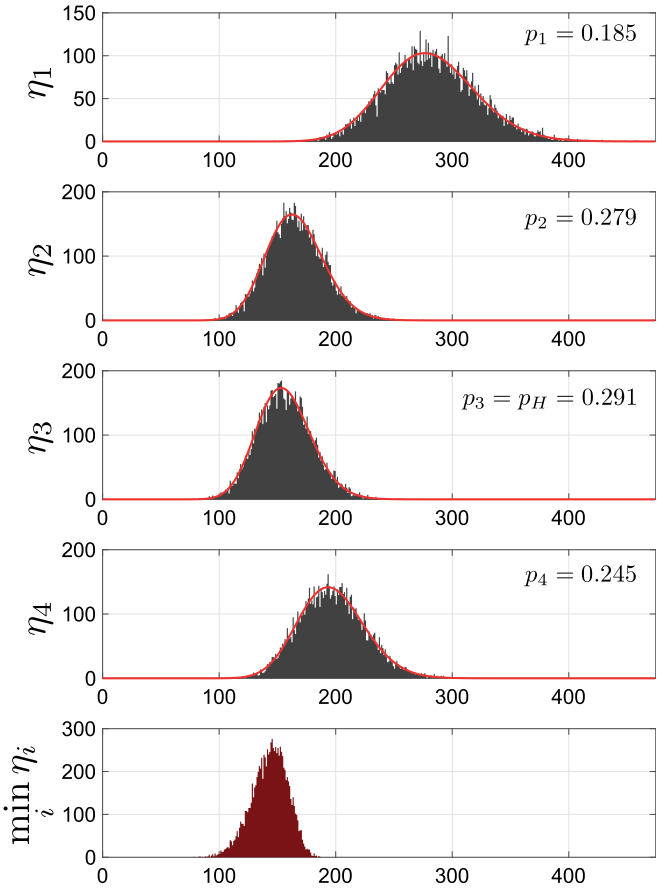

Fig. 4. Statistical distributions of the random variables $\eta_1, \eta_2, \eta_3, \eta_4$ and $\min_i \eta_i$, estimated on the basis of 10.000 randomized experiments ($m = 64$), for a source of $N = 4$ symbols with generation probabilities $\mathbf{P} = (0.185, 0.279, 0.291, 0.245)$. The red curves in the upper plots report the theoretical frequency distribution of negative binomial random variables, with mean value and variances given in (9).

On the other hand, if the source has two or more symbols with generation probability close to p_H or, alternatively, if m is not large enough to "separate" the distributions of the η_i variables, the random variable $T = \min_i T_i = m + \min_i \eta_i$ has mean value lower than $m + \mu_{\eta_H}$. In detail, the statistical distribution of the random variable $\min_i \eta_i$ is related to a cumulative distribution function $\Phi(z) = P(\min_i \eta \leq z) = 1 - P(\min_i \eta_i > z) = 1 - P(\eta_1 > z, \ldots, \eta_N > z)$. Unfortunately, the statistical dependency of the random variables η_i depends on m, N and the probability mass function $\mathbf{P}$. Accordingly, deriving a generic exact expression for F is not trivial. For increasing values of N and m, an heuristic approximated result can be achieved assuming the random variables η_i statistically independent, obtaining a cumulative probability distribution

$$\Phi(z) \approx 1 - \prod_{i=1}^{N} P(\eta_i > z) = 1 - \prod_{i=1}^{N} (1 - \Phi_{\eta_i}(z)), \quad (16)$$

where the product is considering the cumulative probability distributions Φ_{η_i} of the random variables η_i, that are negative binomial with mean value and variance (9).

Furthermore, given m, the higher is p_H the smaller will be the average estimation time T. This latter aspect is relevant to assessing the efficiency of the process, that is increasing with sources having decreasing entropy (increasing values of p_H). This point will be discussed more in detail the next paragraphs.

C. A Stochastic Algorithm to Select Best Entropy Sources

Recalling the partial order in Ω, given in Def. 2, according to the previous results, a low-complexity stochastic algorithm to determine the best source in Ω is presented hereafter.

Since, for a given m, the estimation $\hat{p}_H(m) = \frac{m}{T}$, the higher is T, the smaller is $\hat{p}_H(m)$, the higher is the estimated min-entropy. As a result, we can select the best source in Ω in two steps:

1) For each $S_j \in \Omega$ measure the minimum number $T(S_j)$ of generation trials such to have any symbol generated m times.
2) Select S_{best} as the source with highest T.

It is clear from the previous discussion that the above selection algorithm has the following strengths.

- For weak entropy sources, the expected estimation time μ_T is reduced, and is lower than mN in the worst case ($p_H = 1/N$).
- As shown in Fig. 5, given m, when $p_H \to 1/N$ the value of $\mathcal{H}_m(S) \to \log_2 N$ is systematically underestimated (since $\min_i \eta_i$ has mean value lower than μ_{η_H}), but its variance reaches the minimum. In general, this is favorable when selecting S_{best}. The variance of $\hat{p}_H(m)$ (as well as the variance of the relative estimation errors in Fig. 5) can be effectively reduced increasing m, as previously discussed.

IV. Design, Hardware Implementation and Testing

The design of the best source selection algorithm depends on both the statistical characteristics of the entropy sources in Ω and the targeted reliability of the results (application dependent).

For this reason, in this Section we discuss both the system architecture and its hardware implementation taking into

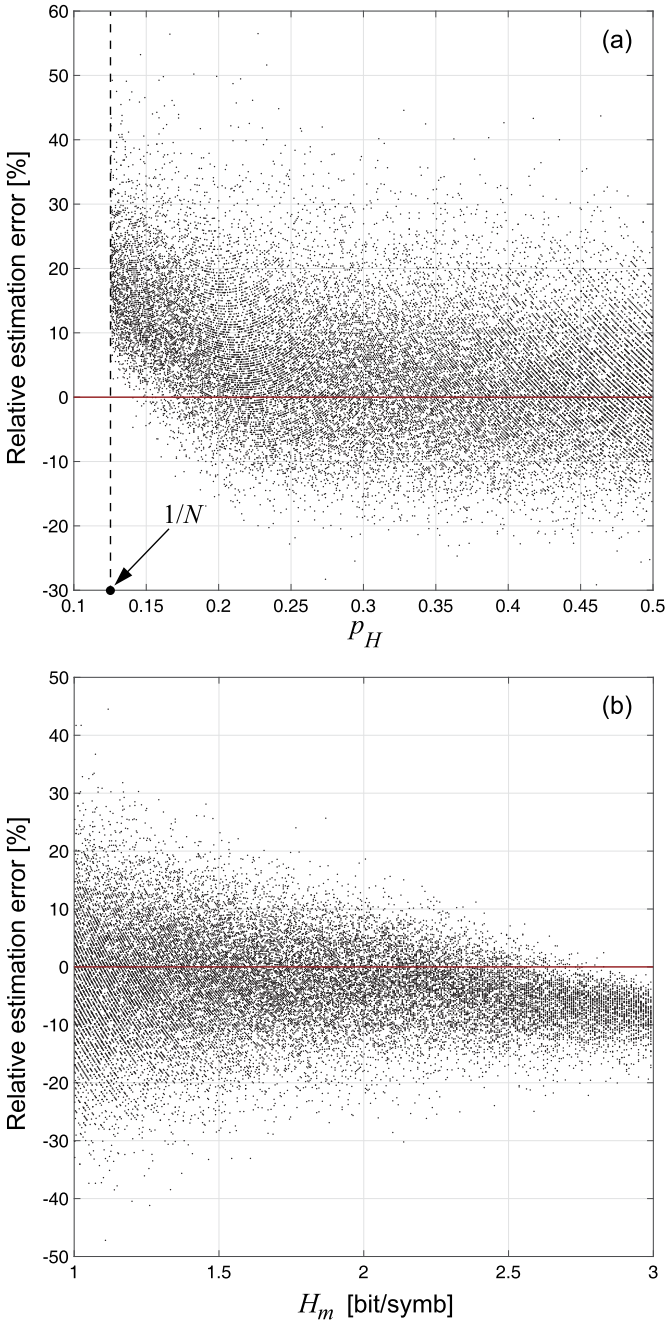

Fig. 5. Monte Carlo simulations (2500 simulation points) evaluating the relative estimation error for $\hat{p}_H(m)$ (subplot a.) and the corresponding min-entropy $\mathcal{H}_m$ (subplot b.) for different p_H values ranging between 1/8 and 1/2, considering $m = 64$ and $N = 8$.

account generalized parametric setting. To assess the validity of the proposal and to inspect the operation of the controller, in SubSec. IV-A and IV-B we present experimental results based on two selected case studies. Finally, in SubSec. IV-C we discuss the possible integration of the proposed algorithm with the Adaptive Proportion Health Test included in the NIST 800.90B publication [2].

For sources of i.i.d. symbols, the design problem reduces to the design of m, that is the unique parameter defining the estimator $\hat{p}_H(m)$. When the sources in Ω have entropies

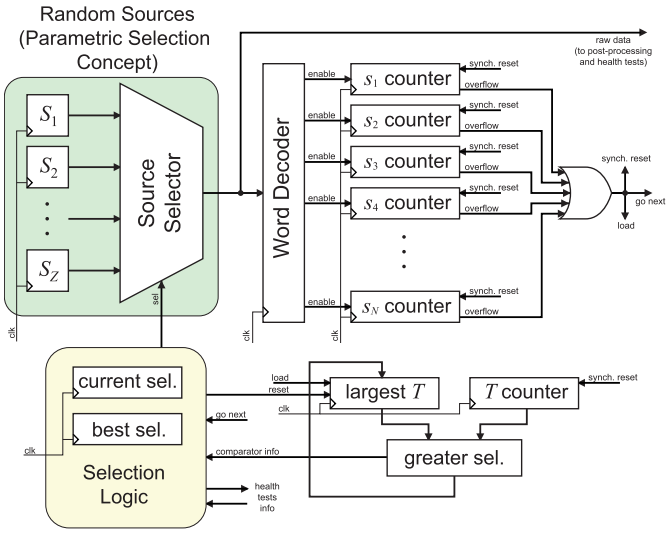

Fig. 6. Architecture of the proposed Tuning Controller (Concept Design).

with significant variance, the estimator can provide satisfactory results even for relatively small values of m. From an intuitive point of view, it is easier to guess a decent source among weak sources, rather than guessing the best source among very similar ones. On the other hand, from the application point of view, when the sources are similar, failing in the selection of the best source is an issue of minor entity.

When considering physical binary sources ($N = 2$), in most cases they are modeled as ergodic processes with vanishing statistically dependency among bits. In these circumstances, the measurement of the ASE can be approximated truncating the series (1) to a convenient number k of terms, approximating the original source assuming to deal with a process of 2^k i.i.d. k-tuplets (words). As a result, in actual applications, the general design of the Tuning Controller algorithm depends on two parameters: m and k.

The block diagram of a digital architecture implementing the proposed algorithm is shown in Fig. 6. The scope of the Selection Logic block is to drive the parametric setup of the entropy source, performing the best source selection. Proceeding in sequential order, the algorithm follows the steps described in Sec. III. As soon as a new word (k-tuple) is collected, it is counted. Since the counting process must stop at m occurrences of any word, the system uses 2^k binary counters, one per k-tuple, each one made of $q = \lceil \log_2 m \rceil$ bits. To simplify the architecture, m can be chosen as a power of 2, i.e., 2^q. In this way, at the end of the counting phase, one of the counter overflow lines gets a high value.

When this event occurs the T counter is compared with a register storing the temporary maximum. In case this latter register is updated, the Selection Logic stores the parametric source address as a temporary best source. The counter T must count up to $N(m-1)+1$ steps, and it has $\lceil \log_2(N(m-1)+1) \rceil = \lceil \log_2(2^k(2^q-1)+1) \rceil$ bits. As a result, the overall complexity of the sequential logic devoted to the counting phase is proportional to a resource consumption of $\approx (k+q) + q2^k \approx q2^k = 2^k \log_2 m$ flip-flops. It is clear

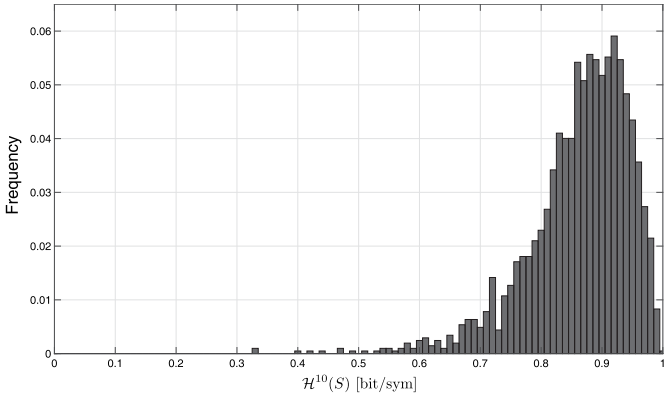

Fig. 7. Distribution of the ASE $\mathcal{H}^{10}(S_n)$ for the 128 binary Markov chains used in the experiment described in Sec. IV-A. The average entropy is ≈ 0.868 bit/sym.

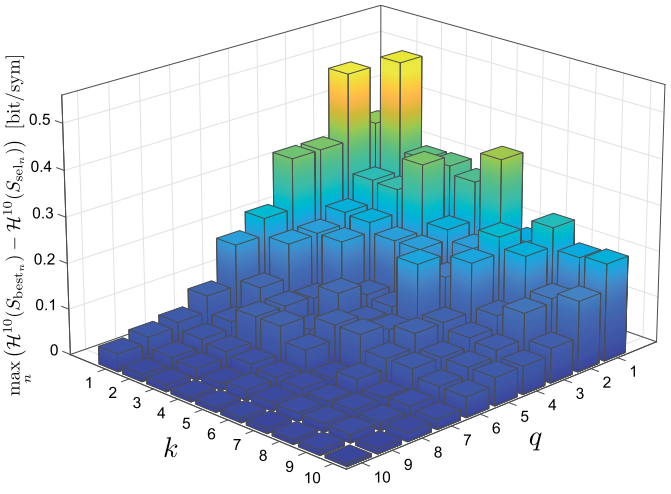

Fig. 8. The worst case maximum selection error given in (18), for different values of k and q.

from the analysis that the design parameter k is critical. As shown in the following Subsections, in most practical cases values of k between 3 and 5 allow to design reliable solutions.

A. Experiments: Application to a Markov Stochastic Source

We have designed a test bench to simulate randomized numerical experiments. For $n = 1, \ldots 128$, we built 128 sets Ω_n of 16 binary Markov chains, characterized by the parametric state transition matrix

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ 1/2 & 1/2 \end{pmatrix}, \tag{17}$$

where $0 < \alpha < 1$ and the element p_{ij} of P is the probability to have the state transition $i \to j$, for $i, j \in \{0, 1\}$. The Markov chain was treated as a binary TRNG generating one bit at each step, according to the current state. The goal of the test bench was to assess the performance of the algorithm described in Sec. III-C when used to guess the binary source S_{best_n} with highest entropy in each set Ω_n.

The designed Markov chains have vanishing statistical dependency among generated symbols. Nevertheless, to establish a reference ordering criterion among sources, we could effectively estimate the sources ASE levels truncating the limit (1) at $k = 10$, i.e., referring to the term $\mathcal{H}^{10}(S)$. Accordingly, for each Markov chain we estimated its ASE on the basis of 1 million bits.

We inspected the operation of the stochastic algorithm for different parametric setting, investigating the range $[1, 2, \ldots, 10]$ for both design parameters k and q (recall, $2^q = m \in \{2, 4, 8, \ldots, 1024\}$).

For each source we randomly set α in (17) according to a Gaussian distribution with mean value and standard deviation equal to 0.25 and 0.0625, respectively, obtaining different levels of entropies, as shown in Fig. 7.

In Fig. 8 we reported the worst case selection error among the 128 sets, defined as

$$\Delta\mathcal{H}^{10}(k, q) = \max_n \left(\mathcal{H}^{10}(S_{\text{best}_n}) - \mathcal{H}^{10}(S_{\text{sel}_n}(k, q))\right), \tag{18}$$

where $S_{\text{sel}_n}(k, q)$ is the source selected by the algorithm. According to this definition, $\Delta\mathcal{H}^{10}(k, q) = 0$ bit/sym only if the algorithm properly selects the source with highest entropy in each of the 128 sets. As it can be appreciated, the higher is the algorithm complexity (larger values of k, q), the better is the result. Interestingly, worst case error drops below 0.1 bit/sym for $q \geq 7$ and $k \geq 3$, whereas it drops below 0.05 bit/sym for $q \geq 9$ and $k \geq 4$. The *average error*, not reported in the figure, drops below 7e-3 bit/sym for $q \geq 7$ and $k \geq 3$, and below 1e-3 bit/sym for $q \geq 9$ and $k \geq 4$. Recalling the hardware architecture shown in Fig. 6, for the considered case study an average selection error below 7e-3 bit/sym can be obtained with only $k = 2^3 = 8$ registers (to count 3-bit word occurrences), being the size of each register equal to $q = 7$ bits. These results reveal the high sensitivity of the selection process, making the proposal particularly efficient in terms of resource consumption (i.e., computational costs and chip area consumption).

B. Experiments: Application to FPGA Digital Nonlinear Oscillators

We have evaluated the capability of the proposed algorithm to select the best source among a set of sixteen based on 7-nodes Galois Ring Oscillators [30], [33], [36], [37], implemented in a Xilinx Artix 7 xc7a35 FPGA. The entropy of such Digital Nonlinear Oscillators (DNOs) is sensitive to the variability of the delays introduced by the FPGA routing circuitry in the feedback loops [30], [33], [37], and the 16 entropy sources were deliberately obtained varying the routing paths. The nonlinear dynamical systems were digitized (1bit digitization) according to a sampling frequency of 100MHz, as discussed in [30], [33], and [37].

For a detailed investigation of the selection algorithm, each source was characterized collecting sequences of 1 million bits, that were used to estimate the Average Shannon Entropy $\mathcal{H}^k(S)$ for different values of k. After preliminary design investigations, we opted to implement the architecture shown in Fig. 6 for $k = 4$ and $q = 7$. The hardware consumption

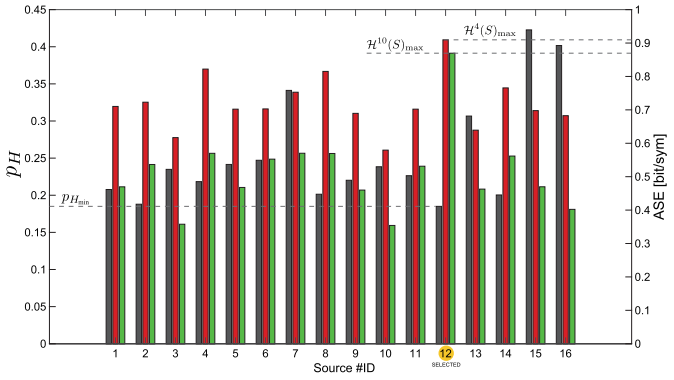

Fig. 9. Statistical characterization of the 16 sources belonging to the set described in Sec. IV-B, processed by the proposed selection algorithm method. To assess the validity of the result, reference values of p_H, $\mathcal{H}^{10}(S)$ and $\mathcal{H}^{10}(S)$ were accurately estimated from sequences of 1 million bits.

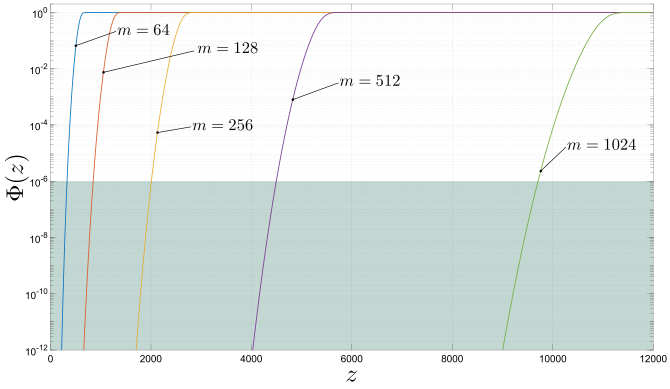

Fig. 10. The cumulative distribution function (19) assuming $p_H = \frac{1}{16} \cdot 130\%$, for different values of m. The grayed area represents the levels for which $P(\min_i \eta_i \leq z_0) < \alpha = 2^{-20}$.

TABLE I
FPGA XILINX ARTIX 7 XC735A HW RESOURCES UTILIZATION FOR THE IMPLEMENTATION OF THE ARCHITECTURE OF FIG. 6 ($k = 4, q = 7$)

Resource	Utilization	Available	Utilization %
LUT	97	20800	0.47
FF	158	41600	0.38

of the design has been reported in Tab. I. Thanks to the simplicity of the design, the authors successfully implemented the architecture in the Artix 7 FPGA considering different clock frequencies up to 400MHz (the maximum for the specific FPGA speed-grade), obtaining an overall power consumption lower than 150mW (entire chip).

We have repeated the selection process 400 times, and in 100% of cases (no exceptions) the algorithm selected the source with highest entropy (source ID = 12), as shown in Fig. 9. In this figure, we reported the values $\mathcal{H}^{10}(S)$, $\mathcal{H}^{10}(S)$ and the estimation of p_H obtained from the above mentioned source characterization. As shown in the figure, in the considered set of sources the one with maximum entropy is the one with minimum p_H. Recalling the results (5)-(7) of Theorem 1, the capability of the algorithm to select the best source in a set (or close to the best, in terms of entropy), strongly depends on the distribution of the sources in the plane $p_H, \mathcal{H}^k(S)$. For this reason, depending on the variability of the entropy sources, larger design parameters k, q can be taken into account for improved performance. As discussed at the beginning of this Section, the algorithm design problem is source-set dependent. In general, an adequate statistical characterization inspecting the variability of the entropy of the considered class of generators is necessary for a proper design of the algorithm.

C. Integration With TRNG Health Test Design

The NIST 800.90B publication provides recommendations about the design of cryptographic TRNG continuous health tests based on the min-entropy concept [2]. The publication defines two approved tests (the Adaptive Proportion and the Repetition Count tests), allowing for the use of developer-defined alternatives to them [2, Sec. 4.5]. In this sub-section we show with theoretical arguments that the low-complexity entropy measurement techniques exploited by the Tuning Controller in Fig. 1 can be used to design a legitimate alternative to the Adaptive Proportion Health Test recommended by NIST. More in detail, the Adaptive Proportion Health Test included in the NIST 800.90B publication cyclically checks if some symbols are generated too frequently than expected, given a reference min-entropy level. At the beginning of each testing cycle, the test takes a sample from the noise source and counts the occurrences B of that sample within an observation window of W samples. If B is greater than or equal to a cutoff value C, the test declares an error. For sources with $N > 2$ symbols the value of W is set to 512, whereas the threshold C is set such to have, for a given expected min-entropy level, $P(B \geq C) < \alpha$, being $2^{-20} \leq \alpha \leq 2^{-40}$ (Type I Error) [2].

Alternatively, the same result can be obtained exploiting the counting phase of the Tuning Controller proposed in this work. In detail, given a reference min-entropy level ($\mathcal{H}_{\mathrm{m}}(S) = -\log_2 p_H$), the number of generation steps $T = m + \min_i \eta_i$ necessary to collect m occurrences of any symbol should be greater than a given threshold T_0, such that $P(T \leq T_0) < \alpha$.

Given p_H we know from Theorem 1 that the worst case entropy is obtained when $F = \lfloor 1/p_H \rfloor$ symbols have generation probability p_H. In such case, the cumulative distribution (16) for the random variable $\min_i \eta_i$ can be approximated as

$$\Phi(z) \approx 1 - (1 - \Phi_{\eta_H})^F, \tag{19}$$

where Φ_{η_H} is the cumulative probability distribution of a negative binomial random variable with mean value $m(1 - p_H)/p_H$.

Using (19), proper values of T_0 can be chosen such to have $P(m + \min_i \eta_i \leq T_0) < \alpha$. For example, in Fig. 10 we reported the numerical computation of (19) assuming $p_H = \frac{1}{16} \cdot 130\%$. The grayed area represents the levels for which $P(\min_i \eta_i \leq z_0) < \alpha = 2^{-20}$. In this example, if $m = 256$, z_0 should be

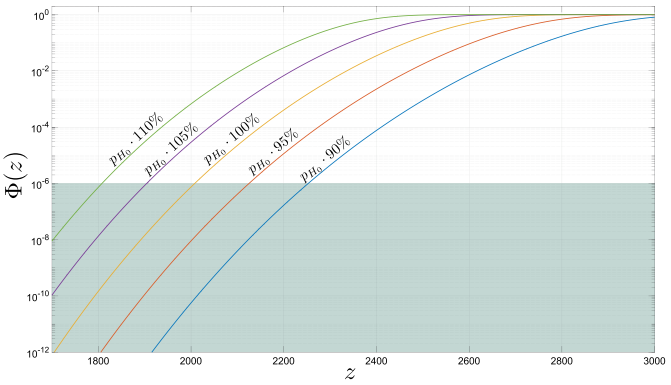

Fig. 11. Sensitivity of the cumulative distribution function (19) with respect to p_H variations, for $m = 256$ ($p_{H_0} = \frac{1}{16} \cdot 130\%$).

set equal to 2006. As a result, the threshold for the hypothesis testing results $T_0 = m + z_0 = 2262$ (Type I Error).

As shown in Fig. 11, even for relatively small values of m, the threshold z_0 exhibits a good sensitivity with respect to p_H variations. In this example a variation of -10% in p_H corresponds to a variation of -10% of z_0 ($z_0 = 1806$ for $\alpha = 2^{-20}$).

According to the above discussion, minor modifications of the architecture shown in Fig. 6 allow to use the tuning controller for both adjusting and monitoring the entropy sources, easily integrating in the same block one of the health tests recommended by NIST [2].

V. Conclusion

We have discussed a stochastic algorithm to design tuning controllers for cryptographic True Random Number Generators, compliant to NIST recommendations, as an effective low-complexity solution to counteract entropy variability in integrated architectures implementing tunable entropy sources. Taking as a reference the min-entropy concept introduced by NIST, we discussed the proposal from both the theoretical and hardware design points of view, validating claims with proofs and experiments. Depending on the target accuracy, the proposed solution is scalable, and its profitable use in TRNG design strongly depends on the kind of core entropy sources taken into account. Nevertheless, the results show the general validity of the presented solution, that in the investigated cases is capable to accomplish remarkable performance with a reasonable amount of hardware resource consumption. The scalability of the solution involves both hardware consumption and monitoring latency that have a cost in the overall TRNG design that must be properly evaluated when targeting lightweight designs. The authors are currently investigating the applicability of the proposed Tuning Controller to the design of efficient TRNGs, focusing on a family of all-digital tunable entropy sources based on DNOs, exhibiting chaotic dynamics.

Appendix

Proof of Theor. 1: Let us note that the function $h : (0, 1] \to \mathbb{R}$ is infinitely differentiable and strictly concave in $(0, 1]$. Furthermore, we note that the sources in $\tilde{\mathcal{C}}$ are related by a bijection to the subset of probability mass functions $\mathcal{K} = \{\mathbf{P} \in \mathbb{M}(N) : \max_i p_i = \tilde{p}_H\} \subset \mathbb{M}(N)$. For simplicity, for any $S \in \tilde{\mathcal{C}}$ we write the Shannon entropy $\mathcal{H}(S)$ as $\mathcal{H}(\mathbf{P}) : \mathcal{K} \to \mathbb{R}$ with $\mathcal{H}(\mathbf{P}) = \sum_{i=1}^{N} h(p_i)$, noting that any permutation of the components of $\mathbf{P}$ provides the same entropy.

For any $1 \le j \le N$ we define $\mathcal{K}_j$ as the subset of $\mathcal{K}$ such that $p_j = \tilde{p}_H$. We note that: 1) $\mathcal{K}$ is the set of permutations of $\mathcal{K}_j$; 2) $\mathcal{K}_j$ is a convex compact set; 3) and $\mathcal{H}(\mathbf{P})$ is strictly concave on $\mathcal{K}_j$. As a result, the strict concavity of $\mathcal{H}$ implies that it has an unique maximum in $\mathcal{K}_j$ that occurs for the probability mass function having $p_j = \tilde{p}_H$ and the other components have same value. Indeed, ab absurdum, if for some $i_1, i_2 \neq j$, $p_{i_1} \neq p_{i_2}$, we would have different permutations of $\mathbf{P}$ (that are also different points in $\mathcal{K}_j$) providing the same maximum value of entropy, contradicting the uniqueness of the maximum. Since $p_j = \tilde{p}_H$ and since if $i_1, i_2 \neq j$ it must be $p_{i_1} = p_{i_2}$, for the normalization property of probability mass functions we have $i \neq j \Leftrightarrow p_i = \frac{1-\tilde{p}_H}{N-1}$. We can use these values to calculate the maximum entropy in $\mathcal{K}_j$, that is $\mathcal{H}_{BC}(\mathcal{K}_j) = \sum_{i \neq j} h(p_i) + p_j = (N-1)h(\frac{1-\tilde{p}_H}{N-1}) + h(\tilde{p}_H)$. Recalling that the elements of $\mathcal{K}$ are permutations of $\mathcal{K}_j$ (thus sharing same entropy levels), we have $\mathcal{H}_{BC}(\mathcal{K}_j) = \mathcal{H}_{BC}(\tilde{\mathcal{C}})$ and (6) has been proved.

To prove (7), we note that $1 \le F \le N$, recalling that $\frac{1}{N} \le \tilde{p}_H < 1$. Accordingly, $\mathcal{H}_{WC}$ is the entropy of a probability mass function in $\mathcal{K}_j$ such that $F = \lfloor 1/\tilde{p}_H \rfloor$ components are equal to $\tilde{p}_H$, at most one is equal to $1 - F\tilde{p}_H$ and the remaining ones (if any) are equal to zero. Since $\mathcal{K}_j$ is compact, $\mathcal{H}$ has a minimum in it. First, let us show that if $\mathbf{P} \in \mathcal{K}_j$ has two components $\tilde{p}_H > p_{i_1} = x \ge p_{i_2} = y > 0$, for two arbitrary indices $i_1 \neq i_2$, then $\mathcal{H}(\mathbf{P})$ can not be a minimum. Indeed, we can build a of probability mass function $\mathbf{P}'(\delta) \in \mathcal{K}_j$ that, depending on $0 \le \delta < \min\{y, \tilde{p}_H - x\}$, can differ from $\mathbf{P}$ by the two components $\tilde{p}_H > p'_{i_1} = x + \delta > p'_{i_2} = y - \delta < 0$. Accordingly $\mathbf{P}'(0) = \mathbf{P}$.

The entropy of $\mathbf{P}'(\delta)$ can be written as $\mathcal{H}(\mathbf{P}') = H_0 + h(x+\delta) + h(y-\delta) = f(\delta)$. Is is immediate to verify that for $\delta \ge 0$ $\frac{df}{d\delta}(\delta) = -\log_2 \frac{x+\delta}{x-\delta} \le 0$, i.e., $\mathcal{H}(\mathbf{P}')$ decreases for δ increasing. As a result, the minimum entropy in $\mathcal{K}_j$ is obtained for probability mass functions having $N-1$ components equal to 0 or $\tilde{p}_H$ and no more than one component between 0 and $\tilde{p}_H$. On the other hand, the normalization property of probability mass functions is satisfied if the components of $\mathbf{P}_{\min}$ for some natural F satisfy the equation $F\tilde{p}_H + x = 1$ with $0 \le x < \tilde{p}_H$. Accordingly, it exists $0 \le \alpha < 1$ such that $\tilde{p}_H(F + \alpha) = 1 \Rightarrow F + \alpha = \frac{1}{\tilde{p}_H} \Rightarrow \lfloor F + \alpha \rfloor = F = \lfloor 1/\tilde{p}_H \rfloor$. □

References

[1] A. J. Acosta, T. Addabbo, and E. Tena-Sánchez, "Embedded electronic circuits for cryptography, hardware security and true random number generation: An overview," *Int. J. Circuit Theory Appl.*, vol. 45, no. 2, pp. 145–169, Feb. 2017.

[2] *Recommendation for the Entropy Sources Used for Random Bit Generation*, document NIST Special Publication 800-90b, Jan. 2018, doi: 10.6028/NIST.SP.800-90B.

[3] *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*, document NIST Special Publication 800-22 Rev.1a, Apr. 2010. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-22r1a.pdf

[4] *Recommendation for Random Number Generation Using Deterministic Random Bit Generators*, document NIST Special Publication 800-90a Rev.1, Jun. 2015, doi: 10.6028/NIST.SP.800-90Ar1.

[5] S. Taneja and M. Alioto, "Fully synthesizable unified true random number generator and cryptographic core," *IEEE J. Solid-State Circuits*, vol. 56, no. 10, pp. 3049–3061, Oct. 2021.

[6] Y. Cao, X. Zhao, W. Zheng, Y. Zheng, and C.-H. Chang, "A new energy-efficient and high throughput two-phase multi-bit per cycle ring oscillator-based true random number generator," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 1, pp. 272–283, Jan. 2022.

[7] X. Wang *et al.*, "High-throughput portable true random number generator based on jitter-latch structure," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 2, pp. 741–750, Feb. 2021.

[8] Y. Luo, W. Wang, S. Best, Y. Wang, and X. Xu, "A high-performance and secure TRNG based on chaotic cellular automata topology," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 12, pp. 4970–4983, Dec. 2020.

[9] Q. Zhao, W. Zheng, X. Zhao, Y. Cao, F. Zhang, and M.-K. Law, "A 108 F2/bit fully reconfigurable RRAM PUF based on truly random dynamic entropy of jitter noise," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 11, pp. 3866–3879, Nov. 2020.

[10] P. Z. Wieczorek and K. Gołofit, "True random number generator based on flip-flop resolve time instability boosted by random chaotic source," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 4, pp. 1279–1292, Apr. 2018.

[11] Y. Liu, R. C. C. Cheung, and H. Wong, "A bias-bounded digital true random number generator architecture," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 1, pp. 133–144, Jan. 2017.

[12] P. Z. Wieczorek, "Lightweight TRNG based on multiphase timing of bistables," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 7, pp. 1043–1054, Jul. 2016.

[13] A. P. Johnson, R. S. Chakraborty, and D. Mukhopadyay, "An improved DCM-based tunable true random number generator for Xilinx FPGA," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 4, pp. 452–456, Apr. 2017.

[14] K. Yang, D. Blaauw, and D. Sylvester, "An all-digital edge racing true random number generator robust against PVT variations," *IEEE J. Solid-State Circuits*, vol. 51, no. 4, pp. 1022–1031, Apr. 2016.

[15] 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," *IEEE J. Solid-State Circuits*, vol. 47, no. 11, pp. 2807–2821, Nov. 2012.

[16] X. Xu *et al.*, "An all-digital and jitter-quantizing true random number generator in SRAM-based FPGAs," in *Proc. IEEE 27th Asian Test Symp. (ATS)*, Oct. 2018, pp. 59–62.

[17] M. T. Rahman, K. Xiao, D. Forte, X. Zhang, J. Shi, and M. Tehranipoor, "TI-TRNG: Technology independent true random number generator," in *Proc. 51st Annu. Design Autom. Conf. Design Autom. Conf. (DAC)*, Jun. 2014, pp. 1–6.

[18] V. B. Suresh and W. P. Burleson, "Entropy and energy bounds for metastability based TRNG with lightweight post-processing," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 7, pp. 1785–1793, Jul. 2015.

[19] T. Addabbo, A. Fort, M. Mugnaini, N. Petra, H. Takaloo, and V. Vignoli, "Self-tunable chaotic true random bit generator in current-mode CMOS circuit with nonlinear distortion analysis," *Int. J. Circuit Theory Appl.*, vol. 47, no. 12, pp. 1877–1892, Dec. 2019.

[20] T. Addabbo, A. Fort, D. Papini, S. Rocchi, and V. Vignoli, "Invariant measures of tunable chaotic sources: Robustness analysis and efficient estimation," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 4, pp. 806–819, Apr. 2009.

[21] S. Tupparwar and N. Mohankumar, "A hybrid true random number generator using ring oscillator and digital clock manager," in *Proc. 6th Int. Conf. Inventive Comput. Technol. (ICICT)*, Jan. 2021, pp. 290–294.

[22] L. B. Carreira, P. Danielson, A. A. Rahimi, M. Luppe, and S. Gupta, "Low-latency reconfigurable entropy digital true random number generator with bias detection and correction," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 5, pp. 1562–1575, May 2020.

[23] X. Li, P. Stanwicks, G. Provelengios, R. Tessier, and D. Holcomb, "Jitter-based adaptive true random number generation for FPGAs in the cloud," in *Proc. Int. Conf. Field-Programmable Technol. (ICFPT)*, Dec. 2020, pp. 112–119.

[24] F. Tehranipoor, P. Wortman, N. Karimian, W. Yan, and J. A. Chandy, "DVFT: A lightweight solution for power-supply noise-based TRNG using dynamic voltage feedback tuning system," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 6, pp. 1084–1097, Jun. 2018.

[25] H. I. Kaysici and S. Ergun, "Random number generator based on metastabilities of ring oscillators and irregular sampling," in *Proc. 27th IEEE Int. Conf. Electron., Circuits Syst. (ICECS)*, Nov. 2020, pp. 1–4.

[26] T. Addabbo, M. Alioto, A. Fort, S. Rocchi, and V. Vignoli, "A variability-tolerant feedback technique for throughput maximization of TRBGs with predefined entropy," *J. Circuits, Syst. Comput.*, vol. 19, no. 4, pp. 879–895, Jun. 2010.

[27] T. Addabbo, M. Alioto, A. Fort, S. Rocchi, and V. Vignoli, "A feedback strategy to improve the entropy of a chaos-based random bit generator," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 2, pp. 326–337, Feb. 2006.

[28] D. Liu, Z. Liu, L. Li, and X. Zou, "A low-cost low-power ring oscillator-based truly random number generator for encryption on smart cards," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 6, pp. 608–612, Jun. 2016.

[29] A. Muthukumar, N. Sivasankari, and K. Rampriya, "Anti-aging true random number generator for secured database storage," in *Proc. 4th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Jan. 2017, pp. 1–7.

[30] T. Addabbo, A. Fort, R. Moretti, M. Mugnaini, V. Vignoli, and M. G. Bosque, "Lightweight true random bit generators in PLDs: Figures of merit and performance comparison," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2019, pp. 1–5.

[31] T. Addabbo, A. Fort, R. Moretti, M. Mugnaini, H. Takaloo, and V. Vignoli, "A new class of chaotic sources in programmable logic devices," in *Proc. IEEE Int. Workshop Metrol. Ind. 4.0 IoT*, Jun. 2020, pp. 6–10.

[32] T. Addabbo, A. Fort, R. Moretti, M. Mugnaini, H. Takaloo, and V. Vignoli, "Chaos in fully digital circuits: A novel approach to the design of entropy sources," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Oct. 2020, pp. 1–5.

[33] T. Addabbo, A. Fort, M. Mugnaini, V. Vignoli, and M. Garcia-Bosque, "Digital nonlinear oscillators in PLDs: Pitfalls and open perspectives for a novel class of true random number generators," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–5.

[34] N. L. Johnson and S. Kotz, *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory*. Hoboken, NJ, USA: Wiley, 1977.

[35] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Discrete Multivariate Distributions*. Hoboken, NJ, USA: Wiley, 1997.

[36] J. D. J. Golic, "New methods for digital generation and postprocessing of random data," *IEEE Trans. Comput.*, vol. 55, no. 10, pp. 1217–1229, Oct. 2006.

[37] T. Addabbo, A. Fort, R. Moretti, M. Mugnaini, H. Takaloo, and V. Vignoli, "A new class of digital circuits for the design of entropy sources in programmable logic," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 7, pp. 2419–2430, Jul. 2020.

Tommaso Addabbo (Member, IEEE) received the Ph.D. degree in information engineering from the Department of Information Engineering and Mathematics, University of Siena, Italy, in 2006. He is an Associate Professor with the Department of Information Engineering and Mathematics, University of Siena. He has authored or coauthored more than 130 international journals or conference papers. His research interests include the design and study of nonlinear circuits and systems, embedded systems, analog and mixed-signal circuits, front-end electronics for sensing systems, and automated measurements. He is a member of the IEEE NCAS TC.

Ada Fort (Member, IEEE) received the laurea degree in electronic engineering and the Ph.D. degree in nondestructive testing from the University of Florence, Italy, in 1989 and 1992, respectively. She is currently an Associate Professor with the Department of Information Engineering and Mathematical Sciences, University of Siena, Italy. Her research interests include the development of measurement systems and automatic fault diagnosis systems.

Riccardo Moretti (Member, IEEE) received the Ph.D. degree in information engineering and science from the University of Siena, Italy, in 2020. He is currently a Research Associate in electrical engineering with the University of Siena. His main research interests include analysis of nonlinear circuits and systems, stochastic aspects of chaotic dynamics and analog circuits design, and design of electronic embedded systems.

Marco Mugnaini (Member, IEEE) received the laurea degree *(cum laude)* in electronics engineering with a major in non-linear automatic controls and the Ph.D. degree in reliability availability and logistics from the University of Florence, Italy, in 1999 and 2003, respectively. Currently, he is the Manager of the Electronics Training Laboratory and an Associate Professor with the University of Siena, Italy. He was a Faculty Member and a Professor with the Electrical and Electronics Technology Department, Higher Colleges of Technology, Abu Dhabi, UAE, from 2012 to 2013. He was awarded as IEEE I&M Distinguished Lecturer from 2017 to 2020.

Duccio Papini received the Ph.D. degree in functional analysis from the International School for Advanced Studies, Trieste, Italy, in 2000. From 2001 to 2002, he was a Research Associate at the University of Turin and the University of Siena. From 2002 to 2014, he was an Assistant Professor at the Department of Information Engineering, University of Siena. Since December 2014, he has been an Associate Professor with the Department of Mathematics, Computer Science and Physics, University of Udine. His research interests include differential equations, with an emphasis on topological methods for nonlinear boundary value problems, properties of complex dynamics, and stability of non-linear systems.

Valerio Vignoli (Member, IEEE) received the graduate degree in electronic engineering and the Ph.D. degree in non-destructive controls from the University of Florence, Italy, in 1989 and 1994, respectively. Since 2020, he has been a Full Professor of electronics with the Department of Information Engineering and Mathematics, University of Siena. He holds seven patents and has authored more than 240 journals or international conference papers. His research interests include design, characterization, and modeling of advanced sensors and development of data acquisition and processing systems.