IEEE Open Journal of the
Industrial Electronics Society

Received 27 January 2023; revised 10 March 2023 and 23 March 2023; accepted 26 March 2023. Date of publication 30 March 2023; date of current version 10 April 2023. The review of this article was arranged by Associate Editor Gerhard Hancke.

Digital Object Identifier 10.1109/OJIES.2023.3262871

Suitability of Generalized GAROs on FPGAs as PUFs or TRNGs Considering Spatial Correlations

MIGUEL GARCIA-BOSQUE, ABEL NAYA, GUILLERMO DÍEZ-SEÑORANS, CARLOS SÁNCHEZ-AZQUETA, AND SANTIAGO CELMA

Group of Electronic Design, University of Zaragoza, 50009 Zaragoza, Spain

CORRESPONDING AUTHOR: MIGUEL GARCIA-BOSQUE (e-mail: mgbosque@unizar.es).

This work was supported in part by Ministerio de Ciencia e Innovación-Agencia Estatal de Investigación under Grant PID2020-114110RA-100, in part by Diputación General de Aragón under Grant LMP197_21, in part by Centro Universitario de la Defensa under Grant CUD-2021_02, and in part by Diputación General de Aragón fellowship to Guillermo Díez-Señorans.

ABSTRACT In the last years, guaranteeing the security in Internet of things communications has become an essential task. In this article, the bias of a wide set of oscillators has been studied to determine their suitability as both true random number generators (TRNGs) and physically unclonable functions (PUFs). For this purpose, a generic configurable structure has been proposed and implemented in an field programmable gate array (FPGA). With this implementation, by introducing some external signals it is possible to configure the system in different oscillator topologies. This way, we have managed to analyze 2730 oscillators composed by seven lookup tables (LUTs) without having to resynthesize the code each time. The performed analysis has included conventional ring oscillators, Galois ring oscillators, and newly proposed oscillator topologies. From this analysis, we have concluded that none of these oscillators behave as an ideal TRNG but ring oscillators present the closest to an ideal behavior. Regarding their suitability as PUFs, some of the newly proposed oscillators in this article present a high reproducibility, higher than that of conventional ring oscillator PUF (RO-PUF) and a high uniqueness. Furthermore, we have noticed that both their reproducibility and their uniqueness tend to improve when increasing the length of the oscillators, which opens the possibility of finding new oscillators with even better properties by studying oscillators of bigger lengths. Finally, by studying the spatial correlation of the bias of these oscillators, we have observed that they present a much lower spatial correlation compared to the ring oscillators, which opens the possibility of using these oscillators in PUF architectures that use more comparisons than typical RO-PUFs.

INDEX TERMS Authentication, FPGA, hardware security, Internet of things (IoT), physically unclonable function (PUF), secure key generation, true random number generator (TRNG).

I. INTRODUCTION

Industrial control systems (ICS) encompass various types of control systems that are used to optimize industrial processes and reduce human errors. They are often used in critical industrial facilities such as power plants, distribution systems, heavy industries, or water treatment facilities. Due to the importance of these facilities, a malicious attack or a human error can cause a great damage, so guaranteeing the security of these systems is an essential task [1].

In the past, these control systems used to be isolated in small networks, protected from the outside world. The workers needed to manually read each component and report the findings. Nowadays, with the great progress in the field of the Industrial Internet of Things and the machine-to-machine networks most of these processes are automated, being able to read and transmit much more useful data. Unfortunately, this advance has also made these control systems more vulnerable to targeted attacks.

While, traditionally, many proposed encryption systems assumed that an attacker could only have access to the encrypted information, in an ICS there can be insider threats with unlimited access to any device. This way, an attacker can extract information via side-channel or fault attacks [2], steal encryption keys that are not stored in a secure way, or impersonate other person to extract confidential data. Therefore, designing secure encryption algorithms is not enough and it is crucial to guarantee other security aspects such as: secure key generation, secure key storage, and authentication. In this context, two important cryptographic primitives are True random number generators (TRNGs) and physically unclonable functions (PUFs) [3], [4].

A TRNG can be defined as a device that generates random numbers from a physical process, rather than by means of an algorithm, whereas a PUF is defined as a physical object that, given an input and certain conditions (challenge), provides a physically defined output (response) that can be used as a unique identifier (often called a "digital fingerprint"). In other words, the same PUF instance always presents the same response for the same challenge, but different PUF instances present different responses to the same challenge.

Regarding TRNGs, besides being needed in many areas such as computer simulations, hazard games, or gambling, they are very important in the field of secure communications. Indeed, encryption algorithms as well as other cryptographic primitives such as message authentication code require the usage of secret keys. If those keys were generated by a user or a pseudo random number generator (PRNG) they could be somehow predictable and potentially vulnerable to cryptanalysis. Therefore, using a TRNG to generate keys is a good way to guarantee a maximum unpredictability [5].

As for PUFs, they can use the uncontrollable variations introduced during the semiconductor manufacturing process to provide low-cost authentication. Furthermore, in the ideal case, their responses are random so they can also be used for key generation and storage [6]. For these reasons, in the last years, PUFs have emerged as a potential solution to preserve a high level of security in IoT structures [7].

With regard to FPGA implementations, while many different structures have been proposed, most of the preferred solutions for both TRNGs and PUFs are based on ring oscillators (ROs). In the case of RO-TRNGs, they typically use the noise in frequency or phase (jitter) of ROs [8]. In the case of RO-PUFs, they are often based on the small differences in frequency between identical ring oscillators implemented in different locations [9]. In some cases, such as [10], [11], the same RO-based structure can be configured to work both as a PUF and as a TRNG.

In [12], a new set of oscillators called Fibonacci Ring Oscillators (FIRO) and Galois Ring Oscillators (GARO) were proposed as fast TRNGs. These systems have been widely studied and several TRNGs based on them have been implemented [13], [14]. However, this kind of structures is not completely understood, are not supported by a stochastic model and, therefore, there is not a way of guarantying a minimum entropy of these systems [15], [16]. Furthermore, some works have proven that the behavior of these systems can greatly depend on the location within the FPGA so that, in certain locations, these systems can present poor randomness results [17]. Based on this fact, Garcia-Bosque et al. [18] studied the possibility of using the variations with the location presented by the GAROs to construct a PUF. That work showed that the bias of these systems varied with the location in a similar manner as the frequencies of a ring oscillator and, therefore, it was possible to use GAROs to construct a PUF in an analogous manner as an RO-PUF but comparing biases instead of frequencies. However, the uniqueness of the tested systems seemed to be smaller than the ones presented by analogous RO-PUFs.

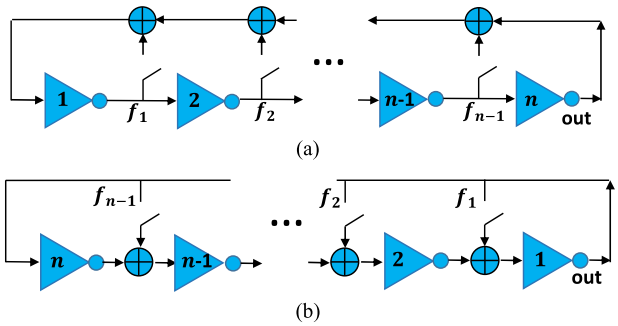

FIGURE 1. Scheme of (a) Fibonacci ring oscillator and (b) Galois ring oscillator.

In this article, an analysis of the bias of a much wider set of oscillators referred to as generalized GAROs has been carried out to evaluate the suitability of these systems as both TRNGs and PUFs. Regarding the suitability of these systems as TRNGs, this analysis has focused on studying if their bias follows the binomial distribution that should be found in an ideal TRNG. Regarding the suitability of these systems as PUFs, this article presents an exhaustive analysis of the properties of these systems including their reproducibility, their uniqueness and their spatial correlation.

The rest of this article is organized as follows. Section II presents the generic structure of the oscillators that we have studied and a way to implement it in a configurable manner. Section III explains the experiment that we have carried out as well as the parameters calculated to evaluate the systems, Section IV presents the experimental results. Finally, Section V concludes this article.

II. STUDIED GENERIC STRUCTURE

A. BACKGROUND

In [12], with the aim of combining the true random properties of ROs and the pseudorandom properties of Linear Feedback Shift Registers (LFSRs), FIRO, and GARO were proposed. Their structure was analogous to an LFSR (with a Fibonacci or Galois structure) but used inverters instead of registers (see Fig. 1). As seen in the figure, the feedback connections can be defined with a set of coefficients f_r. The switches are

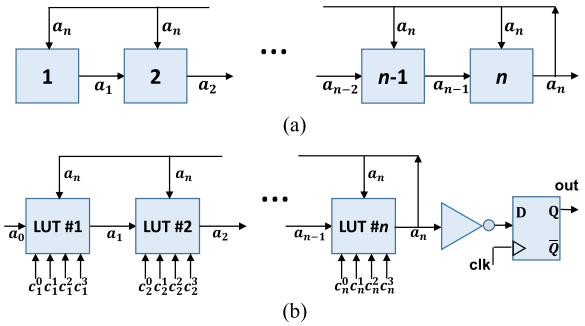

FIGURE 2. Scheme of the proposed (a) generic structure and (b) generic configurable structure.

shown for illustration purposes. In practice, if $f_r = 1$ there is a feedback connection in the ith position and if $f_r = 0$ there is not a feedback connection (and the XOR is not implemented).

In [18], it was shown experimentally that the bias of GAROs changed with their location in a reproducible way and, therefore, they could be used to construct a PUF. As a proof of concept, a seven-LUT PUF that compared the bias of neighboring GAROs was implemented achieving an average Intrachip Hamming Distance (Intra-*HD*) of ∼1% and an average Inter-*HD* of ∼39%.

In this article, a more generic structure has been studied to evaluate their suitability as both TRNGs and PUFs.

B. PROPOSED GENERIC CONFIGURABLE STRUCTURE

The proposed generic structure is shown in Fig. 2(a). It consists of an array of n logic blocks that perform a combinational operation where the output of each block a_r, $1 < r \leq n$, can be any function of the feedback signal a_n, and the output of the previous block a_{r-1}. In case of the first block, its output a_1, can only be the feedback signal a_n or its inverted signal $\overline{a_n}$. It can be trivially seen that this structure includes ROs (when all the blocks perform an inversion operation $a_r = \overline{a_{r-1}}$) and GAROs (when all the blocks perform an inversion or an XNOR operation) but also a large number of additional oscillators. This structure, however, does not include FIROs since these systems would require the implementation of additional LUTs in the feedback signal.

This article will analyze experimentally all the possible oscillator configurations emerging from the abovementioned general structure to see if any of them can be used to construct a good TRNG or a good PUF.

Typically, when an oscillator is implemented in an FPGA, it has a fixed connectivity and can only perform a fixed function (a ring oscillator and a GARO with a certain feedback polynomial). Creating a new implementation of each oscillator requires a large amount of time, which makes it unfeasible to perform a systematic analysis of the proposed generic structure by resynthesizing the FPGA each time a new oscillator architecture is analyzed. To solve this issue, this article presents a generic structure implemented in a configurable manner. Its scheme is shown in Fig. 2(b). Since each LUT can carry out any possible six-input function, it is possible to use two of the inputs as the inputs of the logic block (a_n and a_{r-1} in Fig. 2(a)) while using the other extra four inputs as configuration inputs $c_r = (c_r^0, c_r^1, c_r^2, c_r^3)$, which can be introduced externally, to determine the function $a_r = f_{c_r}(a_{r-1}, a_n)$ that the logic block is performing.

In case of LUT #1, it should be enough to use a single configuration input to determine if the LUT performs an inversion ($a_1 = \overline{a_n}$) or a delay operation ($a_1 = a_n$). This last operation just consists of a propagation of the unchanged input through the LUT (thus applying the inherent delay of the LUT). However, to have a more symmetric structure, the first LUT also includes an a_0 signal that is introduced externally and four configuration inputs so that the first LUT can perform any logic operation $f_{c_1}(a_0, a_n)$. Nevertheless, during all of our experiments, the external signal is always kept at $a_0 = 0$ and the function f_{c_0} is always a delay or an inversion. Finally, an inverter followed by a flip–flop is used to sample the system. This inverter is used to avoid any possible frequency couplings.

Regarding the implemented functions, there are 16 possible two-input functions that can be configured with the configuration signals c_r. However, in practice, some functions are not of interest since they create fixed points or their only effect is to reduce the effective size of the system. For this reason, the only functions that have been considered are the XOR, XNOR, OR, NOR, AND, NAND, DEL ($f_{c_r}(a_{r-1}, a_n) = a_{r-1}$) and INV ($f_{c_r}(a_{r-1}, a_n) = \overline{a_{r-1}}$). In any case, the implemented structure can perform any operation.

It must be noticed that there are a couple of extra functions $f_{cr}(a_{r-1}, a_n) = a_n$ and $f_{c_r}(a_{r-1}, a_n) = \overline{a_n}$ whose net effect is that the LUTs #1 to #r-1 do not have any influence on the output. This can be trivially seen since these functions make the output of LUT #r independent of the output of LUT #r-1 and, due to the characteristics of the proposed structure, it is already independent of the output of LUTs #1 to #r-2. Therefore, using these extra functions, it is possible to study any oscillator of size less than n.

III. EXPERIMENT DESIGN

A. EXPERIMENTAL SETUP

To study the bias of these oscillators, a seven-LUT configurable structure has been implemented in 101 different locations in 20 different FPGAs (using Pynq Z2 boards). More precisely, each oscillator is implemented in a different column and uses seven different rows (one row for each LUT). The reason for using 101 different locations is that it will allow us to generate 100-bit responses (explained below), which is a quite standard number. Furthermore, it is in line with the number of locations used in [18], which makes it easier to compare both works.

The structures have been physically placed so that the LUTs as well as the flip–flops within each structure are close to each other and implemented always in the same relative location, so that all oscillators are almost identical. We have not forced

the exact same relative routing (i.e., wires connecting LUTs), so some oscillators might present small differences, but we do not expect this fact to have a big impact on the results. Finally, the same bitstream file has been used to program each FPGA to make sure that the exact same structures are implemented in all FPGAs. To carry out the experiments, a Python script has been used to send instructions to the FPGAs (choose the configuration, start each measurement, reset the systems, …) and to collect the data from the FPGAs. The communications between the computer and the FPGAs have been carried out through serial RS-232 standard.

To measure the bias of each system, the sampling frequency of the flip–flop shown in Fig. 2(b) is 100 kHz and, when the sampled value is 1, a counter is incremented. After 100 000 samples (1 s), the final value of the counter can be used as an estimation of the bias. These values for the sampling frequency as well as the total number of samples have been chosen for two reasons: first, they are the same as the ones used in [18] so, this way, it is easier to compare both works; second, according to [18], by choosing these values it is possible to estimate the bias with high precision without taking too much time to complete each measurement. To quantify this fact, if we assume that the sample bits follow a binomial distribution with $0.2 \lesssim p \lesssim 0.8$, after taking 100 000 samples, the bias can be estimated with an error of $\sim$0.3%.

Since one of the key properties that we want to measure is the reproducibility of the bias, each measurement is repeated 100 times. To sum up, for each configuration and each FPGA a matrix of integer numbers $A = \{A_i^j\}$ is generated where each element represents the final value of the counter at the ith measurement at the jth location.

Since the final value of the counter is trivially related to the bias of the oscillator, in order to simplify the language in this paper, from now on, we will refer to the final value of the counter as "bias".

B. MEASURED PARAMETERS

To evaluate whether each configuration can be used as a good PUF or as a good TRNG, four bias metrics have been calculated: randomness, reproducibility, uniqueness, and spatial correlation.

1) *Randomness of the bias:* If a certain oscillator was an ideal TRNG, the measured values of the bias should follow a binomial distribution of $p = 0.5$ and $N = 100\,000$. To measure how close the measured values are from a binomial distribution, we have calculated the root-mean-square error (RMSE) between the ideal binomial cumulative distribution function (cdf) and the obtained cdf

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{l=0}^{N} \left(\mathrm{cdf}_{\mathrm{measured}}(l) - \mathrm{cdf}_{\mathrm{bin}}(l)\right)^2}{N+1}} \quad (1)$$

where the cdf indicates the probability of measuring a bias with a value less or equal than l, i.e.,: $\mathrm{cdf} = P(\mathrm{bias} \leq l)$. It must be noticed that, even if some configurations behaved as ideal TRNGs, their values should not be exactly 0, due to the natural uncertainty that exists when we sample a TRNG. Therefore, to have a figure of merit to qualitatively see how an ideal TRNG looks like, we have also used a PRNG to simulate measuring an ideal TRNG, with the number of simulated measurements equal to the number of actual measurements in the tested configurations. We have then computed their RMSE values (referred to as "ideal RMSE") and plotted them in a histogram.

Among the possible methods that can be used to compare two distributions, we consider the RMSE method a good choice due to its simplicity and the ability of giving out a single parameter that can be used to easily compare, which distributions are closer to the ideal binomial distribution. Nevertheless, other metrics such as Kolmogorov–Smirnov, chi-square or Anderson–Darling tests could have been used for this purpose, possibly leading up to similar results.

It must be noticed that obtaining an ideal RMSE does not guarantee that the TRNG is good since it could have other issues such as having a high statistical dependency. Therefore, in case that a configuration presented a good RMSE value, more complex test such as the National Institute of Standard and Technology (NIST) tests [19] should be performed to determine if it can, indeed, work as an ideal TRNG. However, if a bad RMSE was obtained, it would already indicate that the TRNG is not ideal without having to apply any additional tests.

2) *Reproducibility of the bias:* To evaluate if a configuration can be used to construct a reproducible PUF, when measuring the bias in a certain location the result should always be approximately the same. More precisely, a PUF response is typically obtained by comparing the bias in two or more different locations so the differences between the column elements in A (which correspond to several measurements in the same jth location) should be much smaller than the differences between the row elements in A (that corresponds to the measured bias in different locations). A possible way to quantify this reproducibility is to divide the average standard deviations of the rows and columns as done in [18]. However, it can be difficult to interpret how this parameter would exactly affect the average Intra-*HD* of an actual PUF. For this reason, in this article, to measure the reproducibility, we have compared the bias of neighboring oscillators to obtain 100-bit responses and calculated their average Intra-*HD*s. In other words, for each measurement i we compare the values A_i^j and A_i^{j+1} (if $A_i^j > A_i^{j+1}$, the jth bit of the response is 1, otherwise is 0). By repeating this process for all values of j, with $0 \leq j \leq 99$, we obtain a 100-bit response for each measurement i (a total of 100 responses of 100 bits). Then, we calculate all the Intra-*HD*s between these responses and, finally, the average value. This is in line with the analysis made in [20].

It must be noticed that, by using this comparison strategy, the obtained 100 bits within each response will not be independent [21] and, therefore, the responses would not pass any comprehensive randomness evaluation such as the NIST tests. However, this strategy allows us to extract a higher total entropy compared to other approaches, such as pairwise comparison. For this reason, it is quite often used in the literature, although this evaluation pattern could be exploited by side channel attacks [22]. A detailed study of different approaches to generate the responses in this kind of PUFs and its impact on several parameters such as total entropy, entropy per oscillator, and entropy per bit can be found in [23].

3) *Uniqueness of the bias*: To check if a configuration can be used to construct a unique PUF, the average bias in a given location should be different in different FPGAs (more precisely, for a given location, the differences when changing the FPGA should be much bigger than the differences when repeating the measurement). In a similar way as explained before, this could be quantified comparing standard deviations but, again, we have chosen to generate 100-bit responses and calculate their average Inter-*HDs*. More precisely, for each FPGA and configuration, we have taken the most repeated 100-bit response. This way, we have obtained 20 responses (one for each FPGA). After that, we have calculated all the Inter-*HD*s between these responses and, finally, the average value.

In both cases (for the study of the reproducibility and the study of the uniqueness of the bias) we have obtained the fractional Hamming distance (FHD). Thus, given two m-bit responses $x = (x_1, \ldots x_m)$ and $x' = (x'_1, \ldots x'_m)$, their FHD have been calculated as

$$\mathrm{FHD}\left(x, x', m\right) = \sum_{k=1}^{m} \frac{x_k \oplus x'_k}{m} \ [\%] \tag{2}$$

4) *Spatial correlation of the bias:* Finally, it has been widely documented that the frequency of ring oscillators can present a strong spatial systematic component [24], [25]. This fact forces designers to use some comparison strategies (such as comparing only nearby oscillators) to reduce this effect at the cost of reducing the number of output bits. To measure the spatial correlation of these oscillators, we have used the Moran's I [26] as well as the Geary's C [27].

The Moran's I can take values between -1 and 1, where the 0 indicates the absence of correlation, 1 indicates perfect positive correlation, and -1 indicates perfect negative correlation. In case of Geary's C, it takes values between 0 and 2 where the value of 1 indicates the absence of correlation, 0 indicates perfect positive correlation and 2 indicates perfect negative correlation [28]. It must be noticed that, although both Moran's I and Geary's C are related, they are not identical. Moran's I is a measure of global spatial autocorrelation while Geary's C is more sensitive to local spatial autocorrelation.

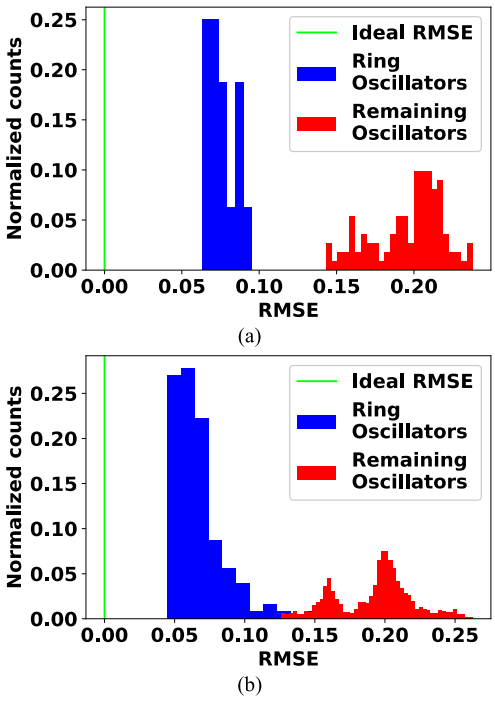

FIGURE 3. Histogram of the obtained RMSE values for (a) five-length oscillators and (b) n-length oscillators ($n \leq 7$). For each configuration, a single RMSE value has been obtained considering all repetitions, locations, and FPGAs.

IV. EXPERIMENTAL RESULTS

A. PRELIMINARY TEST

With the chosen values of sampling frequency, number of samples and number of repetitions it takes 100 s to measure each configuration. With the initially chosen functions there is a total of $2 \times 8^6 = 524\,288$ configurations of length 7 so it is unfeasible to measure all of them (this expression is trivially obtained since the first LUT can perform two different operations while the other six LUTs can perform eight different operations). Therefore, to see, which configurations are more interesting to be studied, a preliminary experiment has been carried out using only five-length configurations in five different FPGAs. Furthermore, of all possible five-length configurations ($2 \times 8^4 = 8192$) we have only measured those ones that do not have a logical fixed point (2048 in total). From this initial test, some preliminary results have been obtained.

First, by looking at the obtained RMSE values, we have noticed that none of the configurations behave as an ideal TRNG [see Fig. 3(a)]. Note that the ideal RMSE green line is actually a histogram of the RMSE values obtained by a PRNG but, since all values are very close to 0, they are contained in a single box. Furthermore, it can be seen that the ring oscillators (all configurations that only have an odd number of inverters and an even number of delays) have lower RMSE values than the rest of the configurations, indicating that their cdfs are closer to the ideal binomial cdfs expected in case that the sampled bits were perfectly random.

TABLE 1 Five-Length Configurations With the Highest Average Inter-HDs

Configuration	Inter-HD (%)	Intra-HD (%)	RMSE
INVb XNOR INVa XOR DELa	30.267	1.357	0.575
INVb XNOR XNOR DELa DELa	30.000	1.374	0.559
DELb XOR XNOR DELa DELa	29.467	17.593	3.455
DELb XNOR INVa XNOR DELa	29.333	1.208	0.623
DELb XNOR DELa XOR DELa	29.333	1.051	0.596
DELb DELa INVa INVa NOR	29.067	14.612	2.701
INVb INVa INVa XOR XOR	28.800	1.817	0.574
INVb XOR XOR XOR XOR	28.533	1.471	0.596
INVb AND DELa NAND NAND	28.533	10.391	3.217
INVb XNOR DELa XNOR DELa	28.533	0.603	0.597
DELb INVa XNOR XNOR DELa	28.267	0.693	0.590
DELb NOR DELa OR XOR	28.267	19.999	3.934
DELb XOR DELa XNOR DELa	28.000	5.342	0.988
INVb INVa INVa XNOR XNOR	27.867	1.265	0.579
DELb DELa DELa DELa NOR	27.867	0.740	0.583
INVb INVa DELa XNOR XOR	27.600	1.998	0.573
INVb INVa XOR XNOR DELa	27.467	0.864	0.597
DELb DELa XOR XNOR DELa	27.467	1.052	0.615
INVb INVa DELa XOR XNOR	27.200	0.643	0.579
INVb XOR NAND INVa NAND	27.200	8.871	3.484

*A color scale has been used to visualize the magnitude of the parameters

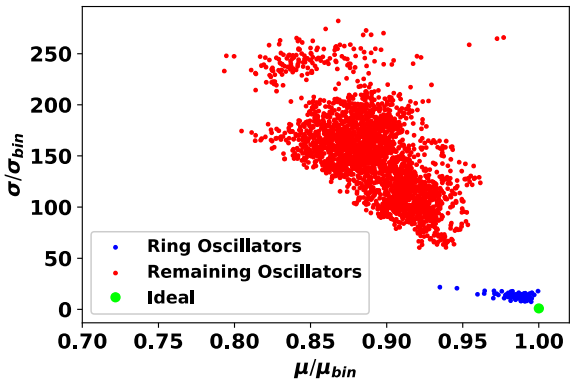

FIGURE 4. Means (μ) and standard deviations (σ) of Ring Oscillators (blue) and the remaining oscillators (red) divided by the ideal values of a binomial distribution (μ_{bin}, σ_{bin}). Each value has been obtained considering all repetitions, locations and FPGAs of a given configuration.

It must be noticed that it is common to find TRNGs that present a bias and it can be easily removed using some postprocessing techniques. Therefore, although the bias is an important parameter used to evaluate the quality of a TRNG, other aspects apart from the bias are usually considered to determine the suitability of a system as a TRNG. A very important parameter is the statistical dependency between the bits, i.e., how likely it is to predict the value of a bit by knowing some previous or following bits. This analysis, however, would require to generate long binary sequences in all configurations, which would be unfeasible for this article.

Therefore, this analysis does not allow us to accurately determine how well these systems would perform as TRNGs. However, it allows us to conclude that none of these systems would behave as an ideal TRNG, unless some kind of postprocessing was used. In a similar way, this analysis does not necessarily mean that ring oscillators are always a better choice as TRNGs than the other tested configurations since their sampled bits could present higher statistical dependency. However, for slow sampling frequencies where the statistical dependency tends to decrease, this result indicates that ring oscillators would usually be better TRNGs.

The second thing that we have noticed is that the measured average Inter-*HD*s are all lower than the ideal value of 50%.

In Table 1, we can see the top configurations ordered by their average Inter-*HD*. From these values it can be seen that most of the top elements have in common that they do not have an AND, NAND, OR, or NOR gate. Furthermore, the few of them that have any of those functions and high average Inter-*HD*, also present a quite high average Intra-*HD*.

This result could be expected since, when GAROs were initially proposed, AND, NAND, OR, and NOR gates were explicitly discarded because their asymmetry was feared to lead to suboptimal properties. This preliminary test gives supporting evidence of this reasoning. Nevertheless, future research could look into this in more detail.

B. FINAL FULL EXPERIMENT

Based on these preliminary results, we have carried out the full experiment with n-length configurations ($n \leq 7$) in 20 FPGAs but using only the XOR, XNOR, DEL, and INV operations. Of all possible configurations, we have only measured those that do not present a fixed point (a total of 2730). It must be noticed that, even after discarding these operations, the number of possible configurations is much larger compared to that of conventional GARO, which only present $2^6 = 64$ possible configurations of size 7, many of them with logical fixed points. This experiment has been carried out at a temperature of 20 °C. From this experiment, several conclusions have been made.

C. ANALYSIS OF THE RMSE VALUES

First, by analyzing the RMSE values [see Fig. 3(b)], the results are consistent with the results in the preliminary five-length test (i.e., no configurations have a bias that follows a binomial distribution with $p = 0.5$ but the ring oscillators are the closest ones to this ideal binomial distribution). Therefore, none of these systems could work as an ideal TRNG and would always need some postprocessing.

In order to further study the differences between the distributions of the bias and the ideal binomial distribution, the mean and standard deviation over repetitions FPGAs and locations of all distributions have been calculated and compared to the ideal values of a binomial distribution. A scatter-plot is shown in Fig. 4, where each point represents a different configuration and its x and y coordinates correspond to its mean and standard deviation, respectively. To better visualize these data, these values have been normalized by dividing them by the ideal values.

From this graph, it can be seen that all configurations have lower means and higher standard deviations than the ideal values. These deviations from the ideal values (for both means

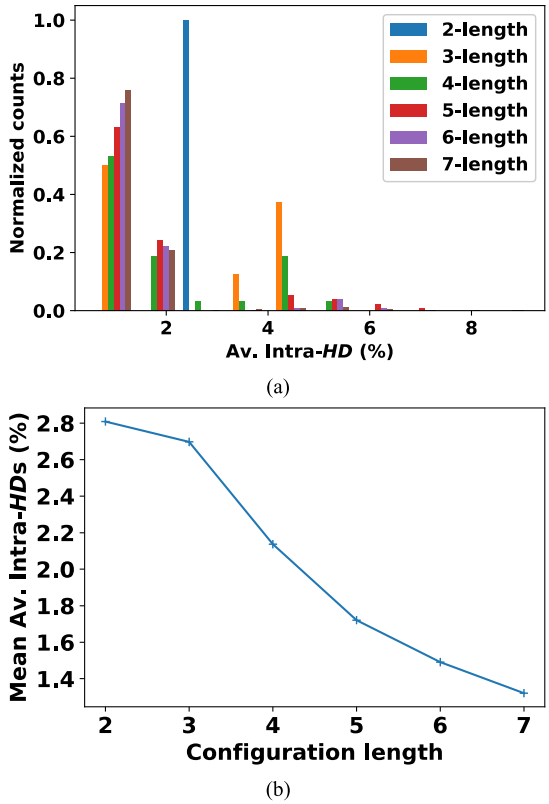

FIGURE 5. (a) Histograms of the average Intra-*HD*s for configurations of different lengths. (b) Mean of the average Intra-*HD*s for different lengths.

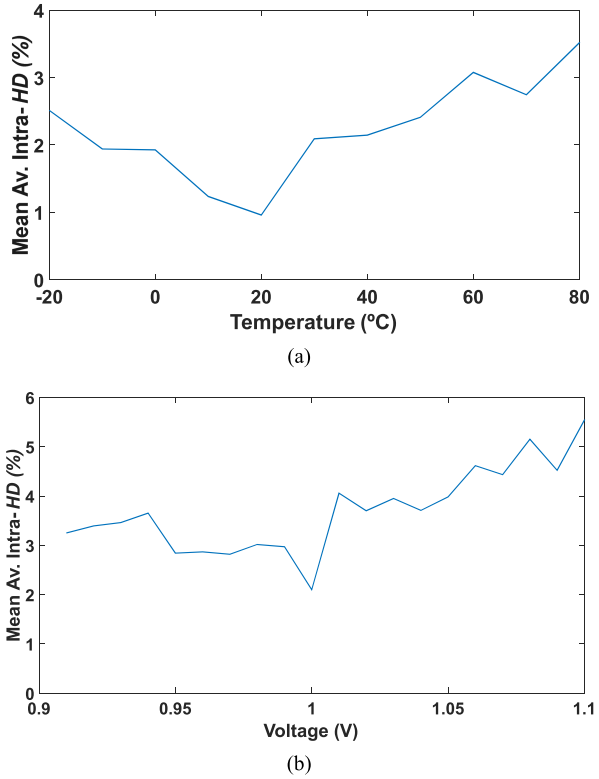

FIGURE 6. Mean of the average Inter-*HD*s of five different configurations for (a) different temperatures and (b) different supply voltages.

and standard deviations) explain why all configurations failed the RMSE test. It can also be seen that, in the case of ring oscillators, these deviations are smaller compared to the rest of oscillators, which explains why they performed better in the RMSE analysis. Finally, it can be seen that, while in the case of the ring oscillators both deviations (lower means and higher standard deviations) are somewhat comparable, in the case of the remaining oscillators, the effect of having higher standard deviations is much more noticeable. This implies that the distributions of the bias are quite wide, i.e., there is a wide range of possible bias that are likely to be measured. While this is not a good property for a TRNG, it could beneficial for a PUF based on comparing biases.

D. ANALYSIS OF THE REPRODUCIBILITY

Second, to study the reproducibility of possible PUFs, we have plotted the histograms of the average Intra-*HD*s of all configurations of each length in Fig. 5(a). From this figure, we can see that most of these oscillators tend to have a high reproducibility. In addition, to check how significant our results are, we have calculated the error (standard error of the mean) of each value of average Intra-*HD*. Although these errors vary depending on the chosen configuration, on average the error was 0.12%, which indicate that the measured values are quite accurate. Moreover, we can see that there is a big influence of the length of the configuration in the measured Intra-*HD*s since configurations of bigger lengths tend to have smaller average Intra-*HD*s. This can be seen more clearly in Fig. 5(b) where we have plotted the mean value of the average Intra-*HD*s of all configurations of each length.

Furthermore, to analyze the influence of the temperature in these systems, we have chosen five of these oscillators with low average Intra-*HD*s ($< 2\%$) and measured their responses at 11 different temperatures from $-20\,^{\circ}\text{C}$ to $80\,^{\circ}\text{C}$. Each measurement has been repeated 100 times to obtain 100 responses at each temperature per oscillator. Then, for each temperature and oscillator, we have calculated the Intra-*HD*s by comparing the measured responses with the most common response obtained at standard conditions ($20\,^{\circ}\text{C}$, 1 V) and calculated the average value (obtaining an average Intra-*HD* per oscillator). The mean values are shown in Fig. 6(a). As the figure shows, while temperature changes can affect the average Intra-*HD*s, the impact is not critical.

In addition, the average Intra-*HD*s of these five configurations have been measured using 20 different supply voltages from 0.91 V to 1.10 V. In a similar way as in the previous case, the Intra-*HD*s at each voltage have been obtained by comparing the responses with the most common response obtained at standard conditions ($20\,^{\circ}\text{C}$, 1 V). As it can be seen by analyzing their mean values [see Fig. 6(b)], small changes in those voltages do not have a great impact on their behavior.

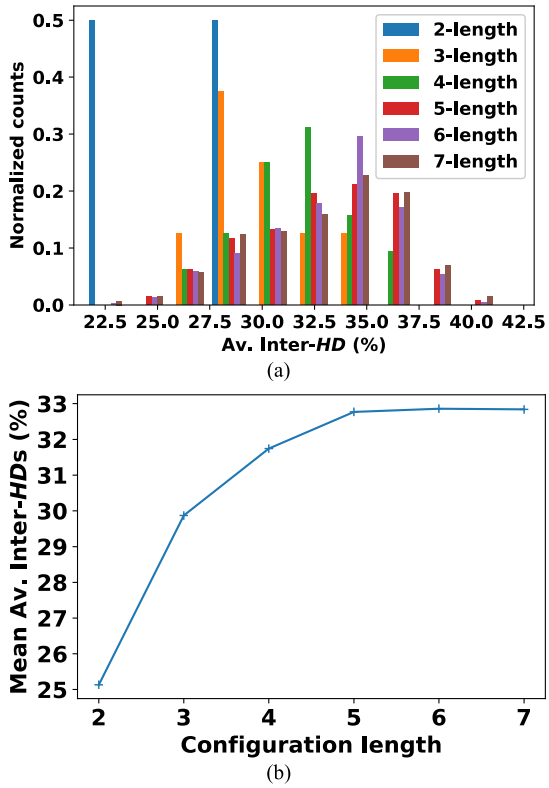

FIGURE 7. (a) Histograms of the average Inter-*HD*s for configurations of different lengths. (b) Mean of the average Inter-*HD*s for different lengths.

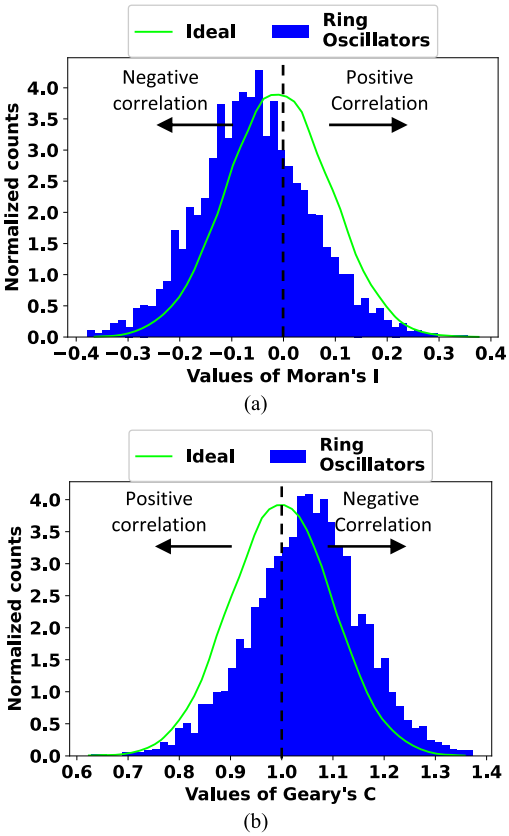

FIGURE 8. Histograms of the (a) Moran's *I* and (b) Geary's *C* for the measured rings oscillators.

E. ANALYSIS OF THE UNIQUENESS

In a similar manner, by analyzing the obtained average Inter-*HD*s we have not found any configuration that achieves the ideal value of 50%. The highest obtained value has been 41% for the configuration "DEL-DEL-DEL-DEL-XNOR-DEL-XOR". In this case, we have also noticed that bigger-length configurations tend to have higher average Inter-*HD*s. This can be seen in Fig. 7(a) where we have plotted the histograms of the average Inter-*HD*s of all configurations of each length and, more clearly, in Fig. 7(b), where the mean value of the average Inter-*HD*s of the configurations of each length has been plotted. It must be noticed that this tendency seems to slow down for high lengths and there is not a big difference between the six-length and seven-length configurations. However, even if the mean value did not change for further bigger lengths, since the number of possible configurations increases exponentially with their length, there could be bigger-length configurations with higher Inter-*HD*s (close to 50%).

It must be noticed that the oscillator with the highest average Inter-*HD* (41.2%) also presents a very low average Intra-*HD*, 1.38%. For comparison, by implementing a regular seven-LUT RO-PUF in the same FPGAs, we have obtained a better uniqueness (an average Inter-*HD* of 47.1%) but a worse reproducibility (an average Intra-*HD* of 1.69%).

Finally, in a similar way as done in the previous subsection, we have calculated the error of each value of the average Inter-*HD*. On average, those errors were 0.91%. These errors are larger compared to the errors when measuring the average Intra-*HD*s due to the fact that we are only using 20 FPGAs to estimate each value. Although this error does not have a big impact in the presented results [29], if new configurations were found with average Inter-*HD*s close to the ideal 50%, it would be advisable to use a bigger number of FPGAs to estimate the average Inter-*HD*s with higher precision.

F. ANALYSIS OF THE SPATIAL CORRELATION

To study the spatial correlation of the bias, we have calculated the Moran's *I* and Geary's *C* of each configuration and each FPGA. In Fig. 8, we have plotted the histograms of the obtained values of Moran's *I* and Geary's *C* for the measured bias of the ring oscillators only. Furthermore, we have plotted the ideal curves that would be obtained if the biases were completely not correlated. From this figure, we can see that the histograms of the ring oscillators seem to deviate from the ideal curves, indicating that the bias of the ring oscillators present some spatial correlation. Indeed, the average value of the Moran's *I* is clearly negative while the average value of the Geary's *C* is clearly bigger than 1. In other words, both metrics tend to present a negative spatial correlation.

For contrast, in Fig. 9, we have plotted the histograms of the obtained Moran's *I* and Geary's *C* for all the tested

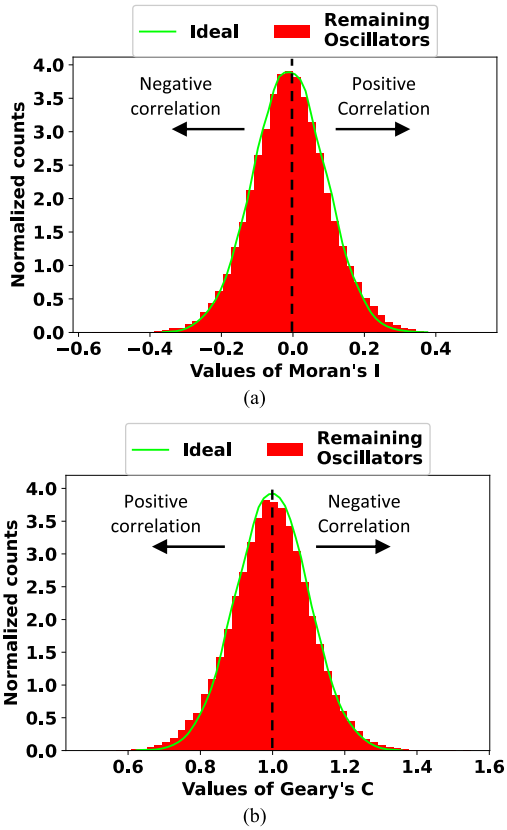

FIGURE 9. Histograms of the (a) Moran's I and (b) Geary's C for the measured remaining oscillators.

configurations, excluding the ring oscillators. Both histograms fit very well to the ideal curves, indicating that the studied structures do not present a significant spatial correlation.

A possible explanation of this low spatial correlation could be that the behavior of these systems presents a very high sensitivity on the inherent delay mismatch due to manufacturing of the used components such as LUTs or flip–flops. Therefore, even if components placed in nearby locations tend to have more similar parameters, the oscillators implemented in those locations present a much different behavior, which translates into the measured uncorrelated bias.

Due to this small spatial correlation, a PUF based on comparing bias of these oscillators would allow a much bigger challenge set (i.e., number of possible comparisons) compared to a regular RO-PUF since it has been widely documented in the literature that the frequencies of ring oscillators present a high spatial correlation. Therefore, this novel family of PUFs presents a clear advantage with respect to standard RO-PUFs.

G. COMPARATIVE ANALYSIS

Finally, to prove the potential of this family of oscillators, we have carried out a comparative analysis between a PUF based on comparing the bias of one of the proposed oscillators and a standard RO-PUF. The chosen oscillator has been the one with a configuration "DEL-DEL-DEL-DEL-XNOR-DEL-XOR" since it was one of the measured oscillators that

TABLE 2 GARO-PUF Implementation Resources

	RO-PUF [20]	Tested RO-PUF	Proposed Oscillator-PUF
Platform	Spartan 3E	Pynq Z2	Pynq Z2
Number of LUTs	5	7	7
Av. Inter-*HD* (%)	47.3	47.1	41.2
Av. Intra-*HD* (%)	0.86	1.69	1.38
Spatial autocorrelation	High	High	Low

presented good properties. Regarding the RO-PUF, we have used the average values obtained in the large scale characterization presented in [20]. However, that work was carried out in a different FPGA, the Spartan 3E and it used five LUTs instead of seven. For this reason, to have a better comparison, we have also tested a single seven-LUT ring oscillator PUF in the 20 Pynq Z2 boards. The results are summarized in Table 2.

From this comparison, we can see that the proposed PUF presents values similar to the standard RO-PUF. When implemented in the same platform (Pynq Z2), it seems to present a somewhat worse uniqueness but a better reproducibility. However, it has the great advantage of presenting a low spatial correlation, which allows the possibility of designing PUF architectures with a much bigger challenge-response set by allowing comparisons of oscillators located far away from each other.

It must be noticed that several implementations of this RO-PUF architectures can be found in the literature, with different average Intra-*HD*s and average Inter-*HD*s. This is due to the fact that the quality of the PUF can depend on several parameters such as the number of stages, the location of the oscillators, the routing or the used platform. The same could apply to the tested oscillators. Therefore, although, when implemented in the same platform (Pynq Z2) with the same locations, the chosen oscillator presented higher reproducibility and lower uniqueness than the RO-PUF, more experiments could be carried out in other platforms with different locations to have a better comparison between both architectures.

V. CONCLUSION

In this article, we have proposed a generic structure named generalized GAROs that includes previously proposed oscillators (such as ROs and GAROs) as well as a new set of oscillators. Furthermore, we have proposed a way to implement this structure in a configurable manner so that, with the same implementation (i.e., the same bitstream file), it is possible to make the system work as any of the possible oscillators. Thanks to this configurable implementation, we

have analyzed all configurations of length five or less, excluding the ones with logic fixed points, to determine their suitability as PUFs or TRNGs. This analysis has shown that configurations with AND, NAND, OR, or NOR gates tend to present a worse behavior. Finally, all configurations of length seven or less, excluding the ones with logic fixed points or with AND, NAND, OR, or NOR gates have been analyzed, to check their suitability as TRNGs or PUFs. From this analysis, several important conclusions have been extracted.

The first conclusion is that it is impossible to create an ideal TRNG based on sampling an oscillator of this kind (with seven or less LUTs). Therefore, to generate perfect random sequences some kind of postprocessing will always be needed. We believe that this result is very important since many previous works have proposed using ROs, GAROs or other similar oscillators as TRNGs. While we cannot rule out the fact that it might be possible to build an ideal TRNG using one of these configurations in a particular FPGA or chip in a specific location with a certain routing, this could not be easily replicated in other implementations (such as ours).

Second, in order to look for an oscillator to construct a good PUF, it seems advisable to try only XOR, XNOR, DEL, and INV functions. It must be noticed, however, that this is an assumption based on a preliminary experiment, so we cannot neglect the possibility of finding some configurations with other functions such as AND, NAND, OR, or NOR that presented good PUF properties.

Third, with some seven-length configurations, it is possible to construct some PUFs with a quite high uniqueness ($>40\%$) and very high reproducibility (some of them are better than a standard RO-PUF).

In fourth place, the reproducibility and uniqueness of these oscillators tend to improve when increasing the configuration length. Combining this result with the fact that there is a huge number of oscillators with bigger lengths, it is likely that there are some configurations of bigger lengths that are suitable to construct much better PUFs.

Finally, the analysis of the Moran's I and Geary's C values shows that the bias of these oscillators, excluding the ring oscillators, present a very low spatial autocorrelation. This could help relax the constraints on where on the chip to place these oscillators in a weak PUF application.

To sum up, this article proofs that a PUF based on comparing the biases of the studied family of oscillators is a viable option and should be considered as an alternative to the standard RO-PUF. Some of the studied oscillators present a high reproducibility and a high uniqueness. Furthermore, they present a small spatial correlation. This presents a great advantage with respect to the standard RO-PUFs that are usually limited to a small challenge set due to the high spatial correlation. A possible drawback is that they seem to present a lower uniqueness than the RO-PUF. However, we believe that it is likely that other unstudied configurations of this family (i.e., with a length bigger than seven) do not have this problem. Future works could study bigger configurations to check this assumption.

REFERENCES

[1] A. Pérez-Resa, M. Garcia-Bosque, C. Sánchez-Azqueta, and S. Celma, "Physical layer encryption for industrial ethernet in gigabit optical links," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3287–3295, Apr. 2019.

[2] Y. Li, M. Chen, and J. Wang, "Introduction to side-channel attacks and fault attacks," in *Proc. Asia-Pacific Int. Symp. Electromagn. Compat.*, 2016, pp. 573–575.

[3] J. Zhang and G. Qu, "Physical unclonable function-based key sharing via machine learning for IoT security," *IEEE Trans. Ind. Electron.*, vol. 67, no. 8, pp. 7025–7033, Aug. 2020.

[4] Y. Zheng, Y. Cao, and C.-H. Chang, "UDhashing: Physical unclonable function-based user-device hash for endpoint authentication," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9559–9570, Dec. 2019.

[5] B. Sunar, "True random number generators for cryptography," in *Cryptographic Engineering*. New York, NY, USA: Springer, 2009, pp. 55–73.

[6] M. Garcia-Bosque, G. Díez-Señorans, C. Sánchez-Azqueta, and S. Celma, "Introduction to physically unclonable fuctions: Properties and applications," in *Proc. Eur. Conf. Circuit Theory Des.*, 2020, pp. 1–4.

[7] M. Kheir, M. M. Tentzeris, A. Abdelgawad, and I. You, "Guest Editorial special issue on intrinsic hardware security for Internet of Things infrastructure," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 321–324, Feb. 2019.

[8] B. Sunar, W. J. Martin, and D. R. Stinson, "A probably secure true random number generator with built-in tolerance to active attacks," *IEEE Trans. Comput.*, vol. 56, no. 1, pp. 109–119, Jan. 2007.

[9] A. Maiti and P. Schaumont, "Improved ring oscillator PUF: An FPGA-friendly secure primitive," *J. Cryptology*, vol. 24, pp. 375–397, 2011.

[10] M. A. Prada-Delgado, C. Martínez-Gómez, and I. Baturone, "Auto-calibrated ring oscillator TRNG based on jitter accumulation," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2020, pp. 1–4.

[11] C. Martínez-Gómez and I. Baturone, "Calibration of ring oscillator PUF and TRNG," in *Proc. Eur. Conf. Circuit Theory Des.*, 2020, pp. 1–4.

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

[13] M. Dichtl and J. D. J. Golić, "High-speed true random number generation with logic gates only," in *Proc. Cryptographic Hardware Embedded Syst.*, 2007, pp. 45–62.

[14] K. Demir and S. Ergun, "Random number generators based on irregular sampling and Fibonacci-Galois ring oscillators," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 10, pp. 1718–1722, Oct. 2019.

[15] L. Matuszewski and M. Jessa, "An auxiliary source of randomness for combined TRNG based on ring oscillators," in *Proc. Poznaskie Warsztaty Telekomunikacyjne*, 2011, pp. 1–4.

[16] M. Dichtl, "Fibonacci ring oscillators as true random number generators–A security risk," in *Proc. IACR Rep.*, 2015, pp. 1–13.

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

[18] M. Garcia-Bosque, G. Díez-Señorans, C. Sánchez-Azqueta, and S. Celma, "Proposal and analysis of a novel class of PUFs based on Galois ring oscillators," *IEEE Access*, vol. 8, pp. 157830–157839, 2020.

[19] A. L. Rukhin et al., "A statistical test suite for random and pseudo-random number generators for cryptographic applications," Nat. Inst. Standard Technol., NIST Special Pub. 800-22, Rev. 1a, Gaithersburg, MD, USA, 2010.

[20] A. Maiti, J. Casarona, L. McHale, and P. Schaumont, "A large scale characterization of RO-PUF," in *Proc. IEEE Int. Symp. Hardware Oriented Secur. Trust*, 2010, pp. 94–99.

[21] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Proc. ACM/IEEE 44th Des. Automat. Conf.*, 2007, pp. 9–14.

[22] D. Merli, D. Schuster, F. Stumpf, and G. Sigl, "Side-channel analysis of PUFs and fuzzy extractors," in *Proc. 4th Int. Conf. Trust Trustworthy Comput.*, 2011, pp. 33–47.

[23] G. Diez-Señorans, M. Garcia-Bosque, C. Sánchez-Azqueta, and S. Celma, "Digitization algorithms in ring oscillator physically unclonable functions as a main factor achieving hardware security," *IEEE Access*, vol. 9, pp. 147343–147356, 2021.

[24] A. Maiti and P. Schaumont, "Improving the quality of a physical unclonable function using configurable ring oscillators," in *Proc. Int. Conf. Field Programmable Log. Appl.*, 2009, pp. 703–707.
[25] M. Pehl, T. Tretschok, D. Becker, and V. Immler, "Spatial context tree weighting for physical unclonable functions," in *Proc. Eur. Conf. Circuit Theory Des.*, 2020, pp. 1–4.
[26] P. A. P. Moran, "Notes on continuous stochastic phenomena," *Biometrika*, vol. 37, no. 1/2, pp. 17–23, 1950.
[27] R. C. Geary, "The contiguity ratio and statistical mapping," *Incorporated Statistician*, vol. 5, no. 3, pp. 115–146, 1954.
[28] T. Arul, N. A. Anagnostopoulos, S. Reißig, and S. Katzenbeisser, "A study of the spatial auto-correlation of memory-based physical unclonable functions," in *Proc. Eur. Conf. Circuit Theory Des.*, 2020, pp. 1–4.
[29] F. Wilde and M. Pehl, "On the confidence in bit-alias measurement of physical unclonable functions," in *Proc. IEEE 17th Int. New Circuits Syst. Conf.*, 2019, pp. 1–4.

MIGUEL GARCIA-BOSQUE was born in Zaragoza, Spain. He received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Zaragoza, Zaragoza, Spain, in 2014, 2015, and 2019 respectively.

He is currently a member of the Group of Electronic Design, Aragon Institute of Engineering Research, University of Zaragoza and an Assistant Professor with the Centro Universitario de la Defensa, Zaragoza. He has coauthored 12 technical papers and more than 22 international conference contributions. He has participated in 8 national and international research projects, two of them as principal investigator. His research interests include chaos theory, true random number generation, cryptography algorithms, and physically unclonable functions.

ABEL NAYA was born in Zaragoza, Spain. He received the B.Sc. degree in mathematics, the M.Sc. degree in computer science, and the B.Sc. degree in computer science from the University of Zaragoza, Zaragoza, Spain, in 2016, 2018, and 2021, respectively.

He is currently working in a leading technology company and collaborating with the Group of Electronic Design, Aragon Institute of Research, University of Zaragoza. He has coauthored one technical article and has developed several apps and tools. His research interests include the field of compilers, programming languages, and programming in general.

GUILLERMO DÍEZ-SEÑORANS was born in Huesca, Spain. He received the B.Sc. and M.Sc. degrees in physics in 2016 and 2017, respectively, from the University of Zaragoza, Zaragoza, Spain, where he is currently working toward the Ph.D. degree in physics with the Group of Electronic Design, Aragón Institute of Engineering Research.

He has participated in five national research projects and coauthored two technical papers. His research interests include physically unclonable functions, cryptography, and physics of complex systems.

CARLOS SÁNCHEZ-AZQUETA was born in Zaragoza, Spain. He received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Zaragoza, Zaragoza, in 2006, 2010, and 2012, the Dipl.Ing. degrees in electronic engineering from the Complutense University of Madrid, Madrid, Spain and from the Helsinki University of Technology, Helsinki, Finland, in 2009, and the Ph.D. degree in education from the University of Zaragoza, in 2022.

He is currently an Assistant Lecturer with the Department of Applied Physics, University of Zaragoza, and member of the Group of Electronic Design, Aragon Institute of Engineering Research, University of Zaragoza. He has coauthored more than 45 technical papers and 150 conference contributions. He has participated in 25 national and international research projects, 12 of which as principal investigator. His research interests include mixed signal integrated circuits, high-frequency analog communications, cryptography applications, and quantum computing.

SANTIAGO CELMA was born in Zaragoza, Spain. He received the B.Sc., M.Sc., and Ph.D. degrees in physics from the University of Zaragoza, Zaragoza, in 1987, 1989, and 1993, respectively.

He is currently a Full Professor with the Group of Electronic Design, Aragon Institute of Engineering Research, University of Zaragoza. He has coauthored more than 130 technical papers and 320 international conference contributions. He has coauthored four technical books and the holder of four patents. He has participated in 70 national and international research projects, 40 of which as principal investigator. His research interests include cyber-physical systems, hardware security and cryptosystems, analog and mixed signal processing, front-ends for wireline and wireless communications, RFIC and MMIC integrated circuits, and cryo-CMOS design for quantum computing.