# On-Line Method to Limit Unreliability and Bit-Aliasing in RO-PUF

Sergio Vinagrero Gutierrez, Giorgio Di Natale, Ioana Vatajelu

# On-Line Method to Limit Unreliability and Bit-Aliasing in RO-PUF

Sergio Vinagrero Gutierrez, Giorgio Di Natale, Elena-Ioana Vatajelu

Univ. Grenoble Alpes, CNRS, Grenoble INP*, TIMA, 38000 Grenoble, France
{sergio.vinagrero-gutierrez,giorgio.di-natale,ioana.vatajelu}@univ-grenoble-alpes.fr

***Abstract*—Physical Unclonable Functions (PUFs) allow generating an intrinsic signature in electronic device thanks to process variability. One of the most researched solutions for PUF implementation is the Ring Oscillator PUF (RO-PUF). This solution is based on the comparison of the frequency of 2 identically designed ROs in an IC. Ideally these 2 ROs would have the same frequency, however this in not the case in reality due to fabrication-induced process variability. By measuring and comparing their actual frequency, a 1-bit PUF response is generated. The RO-PUF has been demonstrated to satisfy the principal randomness requirements (uniformity and uniqueness) but it suffers from problems such as bitaliasing and unrepeatability (i.e. low reliability). In this paper we perform a thorough analysis of RO-PUF bitaliasing and reliability and propose a methodology for its analytical estimation based on the variability profile of the underling technology.**

***Index Terms*—device fingerprinting, ring oscillator, reliability, bitaliasing**

## I. Introduction

Physical Unclonable Functions (PUFs) are security primitives that serve as low cost, tamper-free mechanisms for unique signature and secret key generation, and device identification. To achieve this functionality without the need of resorting to non-volatile memories, PUFs exploit the intrinsic variability induced in the manufacturing process [3]. Indeed, during the manufacturing process, systematic and random variation are introduced and they are the source of the randomness that makes every device unique. This makes PUFs unclonable as it is impossible to reproduce the same physical behaviour even given complete mask information of the circuit. Mathematically, a PUF is a function [1] that maps an input (challenge) to an output (response). Applying the same challenge to different devices leads to different outputs. Applying the same challenge to the same device should lead always to the same output (i.e., reliability of the PUF response).

There are different PUF architectures depending on the physical characteristic they exploit [2], [7]. One of the most studied PUF is the Ring Oscillator (RO) PUF (see Fig. 1) due mainly to its simplicity. RO-PUFs leverage the oscillation frequency of CMOS inverters to generate responses. Different transistor strength in CMOS inverters generate signals with different delays, thus at different frequencies, which can be later compared to generate a response.

In this architecture, the outputs of the two selected ROs are passed through a counter to count the number of cycles in a predefined amount of time, and the output is generated by comparing the number of counted cycles. If the difference is larger than 0 the response is 1 and 0 otherwise. In order to generate multiple bits, various challenges are applied consecutively.

979-8-3503-4135-5/23$31.00 ©2023 IEEE

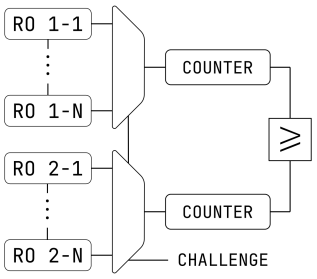


Fig. 1: Schematic of a RO-PUF.

There are a number of metrics in the literature to quantitatively measure the quality of a PUF [4]. The most common ones are uniformity, uniqueness, bit-aliasing and reliability. In this work, we focus on bitaliasing, reliability and their relationship. Bitaliasing occurs when multiple devices produce nearly identical responses. This is undesirable as challenges that present bit-aliasing need to be filtered out, reducing the entropy of the PUF and thus lowering the ability of the PUF to differentiate devices. Reliability is the ability of the PUF to produce the same response when the same challenge is applied multiple times and it is heavily dependent on environment conditions and aging. It is shown in [5] that the number of unreliable responses can be as high as 11% depending on the conditions. Reliability is also affected by aging [10], [11] but its effect is very difficult and expensive to measure. Certainly, it is important to study their relationship as challenges that are reliable across all devices are likely to be bitaliased due to an abnormally large frequency difference. The goal of this work is to analyse the relationship between reliability and bit-aliasing and define an on-line methodology to evaluate the reliability and provide the probability of bit-aliasing of the RO-PUF responses, based on the measured differences of the oscillation frequencies. We propose a methodology, based on electrical simulations, which allows the user to perform a pre-manufacturing PUF reliability study. This is based on finding the range of frequency difference between two ring oscillators

which guarantees a reliable responses and reduces the chance of bit-aliasing. In addition, we provide the means to analyse the trade-off between the reliability of the PUF responses, the probability of bit-aliasing and the PUF entropy i.e., the number of unreliable responses in a given design. In this study we take into account temperature and voltage variations. The work described here is a continuation of the work presented in [16], which is based on the results obtained from the simulation of 200 ROs, implemented with an industrial technological library.

The rest of this paper is organised as follows: the current state-of-the-art related to our work is summarised in section II; section III describes the preliminary mathematical formalism for reliability and bitaliasing estimation; the methodology is presented in sections V and VI; section IV presents the experimental setup; finally section VII concludes the paper.

## II. State of the art

All the PUF quality metrics described in the literature are equally important, but bitaliasing and reliability play a big role when it come to the adoption of PUFs in modern circuits. Due to bitaliasing and reliability issues of today's PUF's their implementation can allow only a small number of devices to be deployed on the field and with costs that render them unsuitable for industrial applications, as shown in [14]. To our knowledge there are no major works on reducing bitaliasing effects on PUFs, since challenges that present bitaliasing are directly filtered out. Nonetheless, important research efforts are dedicated today to develop (i) techniques for reliability analysis and evaluation, and (ii) techniques for reliability improvement.

### A. Techniques for reliability analysis and evaluation

Maes in [8] was among the first to demonstrate the trade off between the PUF reliability and its entropy. The paper proposes an ad-hoc framework for SRAM PUFs based on experimental data. The results show that some responses are more prone to unreliability than others. This is nowadays widely accepted, and it is also the hypothesis of this work.

Schaub et al. provide in [12] a generic probabilistic method for delay PUFs (RO-PUF, RO sum PUF and Loop PUF), where the trade off between reliability and entropy is modelled based on signal-to-noise ratio (SNR), and it is validated by real measurements. They show an analytical method to filter out the responses with high probability of unreliability. This method uses simplified models of delay distribution (due to fabrication-induced variability and thermal noise) to evaluate the SNR of a PUF response. Another work, [13] of Martin et al., provides a PO-PUF reliability evaluation metric based on FPGA-extracted data. Here the trade-off between reliability and entropy is estimated from experimental data. The method is based on extracting the actual distribution of frequencies under fabrication-induced variability and evaluating the frequency fluctuation associated to the operation environment variations (temperature and noise).

In contrast, we propose a simulation-based framework which can be applied before manufacturing the PUF, which allows determining the trade-off between overall reliability and entropy. The proposed framework enables higher accuracy in the results (since it is not based on predictive simplified models of the device variability and noise, but on actual technological electrical models) and higher versatility (since it is not based on measurements extracted from a single technology). The work proposed in this paper will improve the state-of-the-art as it provides a methodology to estimate reliable responses on-the-fly, based on an off-line study under different environmental conditions.

### B. Techniques for reliability improvement

Two main categories can be distinguished in literature: Error Correcting Codes (ECC) and filtering unreliable bits. The ECCs use circuit redundancy to detect and correct unreliable PUF responses by using helper data calculated during PUF enrolment (e.g. Maes et al. in [6]). They are very efficient in guaranteeing the PUF reliability but they are very expensive in terms of area and power consumption. Filtering unreliable bits requires knowledge of the PUF behavior under different environmental conditions and aging and it is based on removing from the PUF responses the bits with reliability lower than a certain threshold. The efficiency of filtering has been demonstrated by Bhargava et al in [9]. Moreover, Schaub et al. in [15], have compared the two techniques and showed that filtering is more efficient than ECC to improve PUF reliability. While efficient, the filtering technique has the disadvantage of requiring a heavy characterisation campaign to understand the reliability of each PUF bit under all possible operation conditions. All reliability estimation methods presented in the previous subsections have the ability to mitigate in part this shortcoming. Similarly, the reliability estimation technique proposed in this paper can be efficiently used for filtering unreliable bits. We show that using the proposed method we can estimate the number of PUF responses which guarantee a certain level of reliability. Moreover, the method can be implemented in hardware in order to provide, for each challenge, the information whether its response is reliable or not.

## III. Preliminaries

Reliability is defined as the ability of the PUF to produce the same response for a given challenge under different operation conditions and aging. In the case of a RO-PUF, the frequencies of two ROs are compared to generate a response. By convention, if the frequency of the first RO is larger than the frequency of the second, the PUF response is 1, otherwise is 0. If the two frequencies are very similar, the response is prone to be unreliable since a small shift in the frequency in one of the ROs due to noise or environmental conditions can alter the response. Therefore, analysing the frequency differences of all ROs in a PUF can give us a good measure of PUF reliability.

Based on the general agreement, the oscillation frequencies of all ROs in the PUF can be fitted to a normal distribution (left plot in figure 2). Frequency differences close to 0Hz are possibly unreliable. For this case, we define a threshold $T$ such that pairs for which $-T < f_{diff} < T$ are considered unreliable (area in yellow). Thus, reliability is calculated as $Reliability = 1 - [P(T) - P(-T)]$

Furthermore, we can use the distribution of frequency differences to estimate the time needed to obtain a response for a certain challenge. The measurements of each RO frequency is performed resorting to a counter, which count at each rising edge of the RO, and the PUF response is obtained by comparing two counters. It has been observed that RO pairs whose frequency difference is very large can assure a meaningful counter difference early on, while pairs whose frequency difference is very small take more time to provide a meaningful counter difference since the frequency difference might be masked by the sampling effect of the counters, therefore, the two counters can register the same value, until the frequency lag becomes significant enough to counteract this effect. Our methodology is based on the observations that two RO start oscillating at the same time and any two sine waves with different frequencies will experience simultaneous zero-crossing periodically, at intervals $T_{sync} = 1/f_{diff}$. As a result, any expected change in the counter difference must happen in a $T_{sync}$ interval. If we observe the counter difference at certain intervals $t_{sample}$, we can define the *expected number of samples until the counter difference changes* as $E = T_{sync}/t_{sample} = 1/(f_{diff} \cdot t_{sample})$. By introducing the notion of frequency difference threshold for reliability (i.e., $T$) we can correlate the lag of meaningful counter difference with the reliability of the corresponding response.

Based on our observations from simulations and the general agreement on the variability distribution, the oscillation frequencies of all ROs in the PUF can be fitted to a normal distribution $F \sim N(\mu_n, \sigma_n)$. As said before, similar frequencies (i.e., frequency difference close to 0Hz in this distribution) are possibly unreliable. For this case, we define a threshold $T$ such that pairs for which $-T < f_{diff} < T$ are considered unreliable (area in yellow), i.e. *area of unreliability*.

However, challenges with a large frequency difference should not be used directly, since the bigger the frequency difference, the smaller the chance that process variation makes a difference in the different ICs. Those challenges will thus be common among multiple instances (i.e. they will present bitaliasing). In the rest of the paper, we establish the relationship between reliability and bitaliasing.

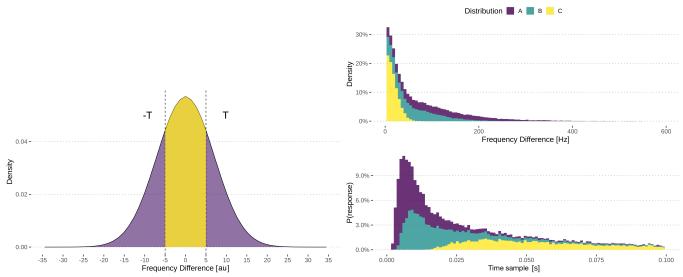


Fig. 2: On the left, distribution of frequency differences. Area in yellow marks unreliable responses. On the right, three different distributions of absolute frequency difference. On the bottom right, the probability of obtaining a response for each distribution.

The frequency difference distribution is calculated as the difference of two identical normal distributions in the following manner:

$$\mu_{diff} = \mu_n - \mu_n \simeq 0 \tag{1}$$

$$\sigma_{diff} = \sqrt{\sigma_n^2 + \sigma_n^2} \simeq \sqrt{2} \cdot \sigma_n \tag{2}$$

The reliability can be calculated as the complement to the area of unreliability. Mathematically, it can be calculated as follows, where $P \sim CDF(\mu_{diff}, \sigma_{diff})$.

$$Reliability = f(\mu_n, \sigma_n, T) \tag{3}$$

$$= 1 - p_{unrel} \tag{4}$$

$$= 1 - \int_{-T}^{T} \frac{1}{\sqrt{2}\pi\sigma_{diff}} exp(-\frac{1}{2}(\frac{t}{\sigma_{diff}})^2)\, dt \tag{5}$$

$$= 1 - [P(T) - P(-T)] \tag{6}$$

Furthermore, we can use the distribution of frequency differences to estimate the time needed to obtain a response to a certain challenge. The measurements of each RO frequency is performed through a counter activated at each rising each of the RO. It has been observed that RO pairs whose frequency difference is very large (in the tails of the normal distribution) can assure a meaningful counter difference early on. In the opposite case where the RO frequencies are very close together, the frequency difference might be masked by the sampling effect of the counters, therefore, the two counters registering the same value, until the frequency lag becomes significant enough to counteract this effect.

$$f_1 + f_2 = 2 \cdot A \cdot cos\left[\frac{k_1 - k_2}{2}x - \frac{\omega_1 - \omega_2}{2}t\right] sin\left[\frac{k_1 + k_2}{2}x - \frac{\omega_1 + \omega_2}{2}t\right] \tag{7}$$

Here we propose a methodology to evaluate the time needed for a meaningful counter difference to be observed between two RO, as a function of their respective frequencies. This study is based on 2 observations: (i) the two RO start oscillating at the same time; (ii) any two sine waves with different frequencies will experience simultaneous zero-crossing periodically, at intervals $T_{sync} = 1/f_{diff}$. The resulting difference wave shown in equation 7 has two terms that oscillate at $f = 1/2(f_1 - f_2)$ and $f = 1/2(f_1 + f_2)$ respectively, and as a result of this periodical instantaneous coincidence, any expected change in the counter difference must happen in a $T_{sync}$ interval. If we observe the counter difference at certain intervals $t_{sample}$, we can define $E$ the *expected number of samples until the counter difference changes* as:

$$E = \frac{T_{sync}}{t_{sample}} = \frac{1}{f_{diff}\, t_{sample}} \tag{8}$$

The plot in the top-right corner in figure 2 illustrates the distribution of $1/f_{Diff}$ for 3 different PUFs. Distribution A assumes a fabrication process with a wide process variability,

i.e., large variance of RO frequency (large $\sigma_n$), while distribution C assumes a fabrication process with a narrow process variability, i.e., small variance of RO frequency (small $\sigma_n$). In the bottom-right corner, it is illustrated the probability of registering a counter difference after $t_{sample}$ (i.e., probability of $E \neq 0$). It should be noted that in case of distribution A the expected times to obtain a response are much shorter than in the case of distribution C. It is important to clarify that this estimation just provides the time for the highest probability of obtaining a response, but not if the response is reliable. The latter will be studied in section V.

By introducing the notion of frequency difference threshold for reliability (i.e., $T$) in eq. 6, for a pair of ROs, we can correlate the lag of meaningful counter difference with the reliability of the corresponding response and the likelihood of it being bitaliased.

## IV. Simulation setup and results

As this work is an extension of the work proposed in [16], the simulation setup is identical. The RO-PUF under study is composed of 200 ROs (two groups of 100 ROs, i.e., $RO_{1,1}$ to $RO_{1,100}$ and $RO_{2,1}$ to $RO_{2,100}$ in Fig. 1), each of them designed with three CMOS inverters in 65nm technology provided by ST Microelectronics. The output of a RO is connected to a counter (implemented in VerilogA) which increments its value at every rising edge of the oscillation. Hysteresis behaviour is implemented in the counter to avoid measurement errors. To reduce the simulation time, each RO has been simulated independently and its frequency has been indirectly measured by the counter. The state of the counter has been sampled at each 1ns.

In order to emulate the effect of manufacturing process variability, we applied random variation of the width, length and Vth of each transistor. Moreover, each simulation is carried out at different temperatures (24 to 30 degrees) and voltages (0.9 to 1.1 V). We have chosen a small range of variation for both voltage and temperature as we are not interested in the operating point of the device, rather the gradient of temperature and voltage of two ROs under the same device. These values are chosen under the assumption that there is almost no temperature gradient among the ROs as they are closely packed. The values of voltage have been chosen based on the assumption that it's possible to find ROs at different voltage due to, for example, the resistance of the power line.

## V. Proposed method for reliability estimation

In this section, we describe a methodology for PUF reliability estimation based on the mathematical preliminaries and the simulation results described in sections III and IV respectively. Without loss of generality, the methodology will be described and validated for a RO-PUF with 200 3-stages ROs designed in 65nm STMicroelectronics bulk-technology.

In general, a PUF has two operation phases: (i) the enrollment - when a set of challenges are applied to the PUF for the first time, under nominal environmental conditions and the golden responses are obtained; (ii) the mission - when the PUF is challenged by request. In mission mode, the PUF might be challenged under operation conditions which are different than the nominal (due to noise coming from the surrounding circuitry, or operation in extreme environments) which might affect the PUF response. A challenge applied to a PUF in mission mode which always generates the golden response, is defined as reliable. On the contrary, when it will not always generate the golden response is unreliable. The reliability of the PUF is determined by the reliability of all its challenges.

Our proposed method of reliability estimation is based on the observation (explained in section III) that RO challenges for which the frequency difference is very large can assure a meaningful counter difference early on, while RO challenges for which the frequency difference is very small (close to 0Hz) take more time to provide a meaningful counter difference, i.e. a response. For this reason, when simulating the RO, we do not only retain the counter value at the end of the simulation time (in our example 100ns), but we also record intermediary counter values at a fixed time-step (in our example 1ns)). In this way, when applying a challenge to the RO-PUF, we are able to calculate its response and also determine how fast this response can be obtained (i.e., after how many time steps we start observing a counter difference). *We postulate that challenges for which the counter difference is bigger than a value CD at a measurement time MT are reliable, while challenges for which the counter difference is smaller than CD are unreliable.* To demonstrate this assumption and to estimate the value of *CD* and *MT*, we simulate the full set of ROs under study for various operation temperatures (26°C, 27°C and 28°C) and supply voltages (0.95V, 1.00V and 1.05V).

The key point is counting the number of bit-flips to estimate the unreliability. A response is considered reliable if it gives the same response as the reference sample, which means, if the sign of difference in counters is the same as the sign for the nominal response.

The full algorithm is described in [16] but it will be briefly summarised here. For every pair of ROs, the nominal response is calculated for every time step of the counter, at the nominal voltage $V_n$ and nominal temperature $T_n$. For every other condition, we calculate the corresponding PUF response. If the response is equal to the one obtained at nominal conditions, the response is considered reliable. The overall PUF reliability is calculated by averaging the reliability of all ROs. Thus, the reliability will be 100% if none of the challenges differ from the nominal sample and 0% if all of them differ. As it has been previously stated, we gathered the value of the counter at different time steps in the simulation, so this analysis can be done at different times to study the evolution of reliability in time.

The left plot in figure 4 shows the minimum counter difference CD to achieve 100% reliability is shown for every measurement time MT of the counter. The values from 0ns to 30ns have been discarded since the values of the counter are too low to provide enough information. The minimum value to achieve 100% reliability for each RO at every MT is represented in the right plot of the same figure. Challenges

with counter differences CD that fall under the line at a given time are marked as unreliable, while challenges with counter differences above the line are reliable. By using this approach, the PUF reliability can be improved using the filtering method, as described in section II. However, as it will be presented in the following section, challenges that are very far from the green line are likely to present bitaliasing.

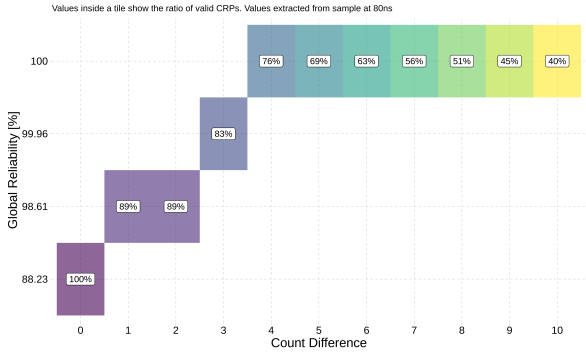


Fig. 3: Relationship between PUF reliability, counter difference and PUF entropy.

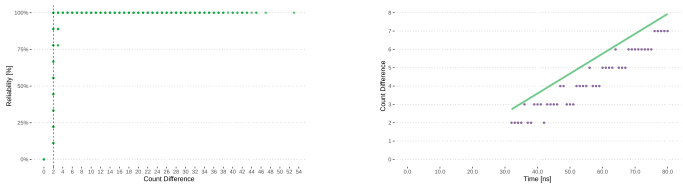

Fig. 4: On the left, the reliability vs counter threshold of a single RO. On the right, the model representing the minimum counter threshold for every MT to consider reliable challenges.

Figure 3 correlates the PUF reliability with the counter difference CD (measured at MT=50ns) and the PUF entropy (number of used CRPs). In this study, all CRPs for which the counter difference is lower than CD are filtered out. The number in each coloured block represents the percentage of remaining (reliable) CRPs after filtering out the unreliable ones. We can observe that increasing the CD increases the reliability of the PUF but reduces its entropy. For instance, when CD=4 is used, the reliability increases from 88.23% (without filtering) to 100% (with 24% of CRPs filtered out). If a more conservative approach is desired, with larger margins for reliability (to account for instance for unforeseen disturbances), the CD can be set to a larger value, hence filtering out more CRPs.

## VI. Proposed method for bitaliasing estimation

The method previously described allows filtering responses that can be deemed unreliable due to a low frequency difference. This section focuses on providing information on whether a challenge is likely to present bitaliasing given the frequency difference.

The oscillation frequency of a single RO can be modelized with eq. 9. Indeed, the oscillation frequency of a RO is shifted from it's design value due to manufacture variability and environmental conditions. Yet, there are design choices and problems during manufacture that can further alter the oscillation frequency (denoted in the equation by $\delta_{design}$). This additional shift in frequency would translate in a bias of the responses of all devices produced by the set of all affected ROs, commonly known as *bitaliasing*, proving a security risk where an external attacker can gather enough information to guess the responses of other devices, as the PUF loses the ability to correctly identify all devices in the system.

$$f_{RO} = f_{nom} \pm \delta_{var} \pm \delta_{env} \pm \delta_{design} \tag{9}$$

It has already been proven that pairs with a low frequency difference are unreliable. Yet, we cannot chose pairs that lay in the other side of the spectrum, as the effect of process variability will be diminished and making the responses likely to be the same for all devices.

In order to calculate if a challenge is likely to present bitaliasing, the frequency distribution of the corresponding ROs of all devices is used to calculate a frequency difference distribution with eq. 7. The mean of this distribution provides information about bitaliasing. If the distribution is centred around 0, it is likely to have not present bitaliasing, as the ratio between the right and left halves of the curves is close to 1. However, if the ratio of areas is skewed towards positive or negative values, the challenge is likely to present bitaliasing, as in most cases it will provide a 1 or 0 respectively. An example of this is represented by figure 5. Each row on the heatmaps represents the distribution of frequencies of each device and each column represents the same RO across different devices. The top heatmap shows the unbiased scenario, where any possible pair of columns is going to provide a frequency difference distribution centred at around 0. However, in the bottom heatmap, we see that the frequencies provided by the ROs near the 10th and 40th RO exhibit abnormally large frequency, either due to a design error or a manufacturing problem. In this case, any frequency difference from all pairs of these ROs are going to be skewed towards the positive side and will then present bitaliasing. It is important to mention that this skewness of the frequency difference distribution will also present in the case where the RO frequency is abnormally low.

We have built upon the results shown in figure 4 and calculated the likelihood of bitaliasing based on the frequency difference distribution of each RO pair. The graph represents the entropy of the bitaliasing so a value of 1 represents that there is no bitaliasing while a value of 0 means that the challenge is clearly aliased. Frequencies that are very close together and provide no difference in the counter provide high entropy due to the fact that their frequency difference distribution will always be a symmetric Gaussian centred at 0, so these responses should never be considered. On the other extreme, the responses with a very high counter threshold provide little to no entropy, since most likely the high frequency difference is due to aliasing. The shaded rectangle in the graph shows the range of counter threshold that provides responses with high enough reliability and bitaliasing.

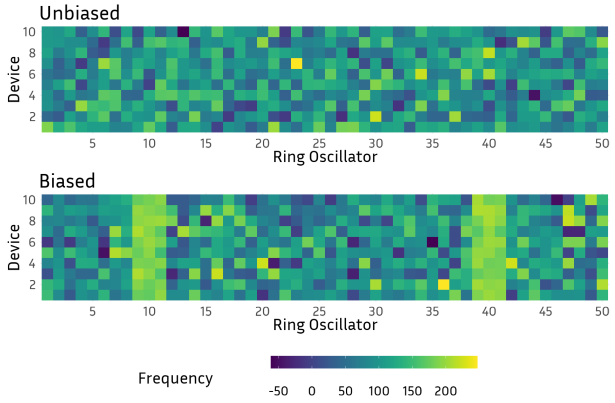


Fig. 5: Heatmaps containing the frequency distribution of 10 devices with 50 ROs. The top heatmap represents an unbiased scenario. The bottom heatmap represents a biased scenario where the ROs 10th and 40th are abnormally high.

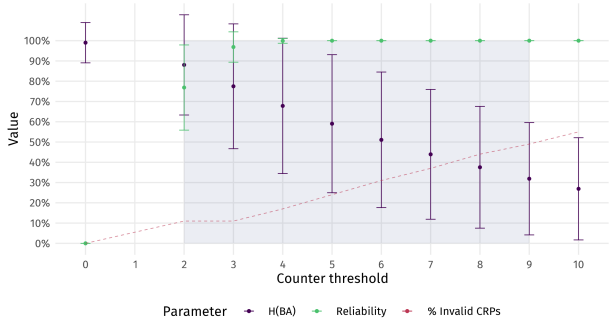


Fig. 6: Relationship between bitaliasing and reliability taking into account the counter threshold.

This study can be used to implement an online test methodology for RO-PUF reliability estimation. We propose a design-for-test solution which modifies the classical RO-PUF scheme in Fig. 1 by adding two blocks: an absolute-value subtraction block which computes the difference between the two counters in absolute value; and a comparator, which compares this difference against two constant values, which are the counter thresholds described previously. This new design will provide, besides the response to a challenge, the information whether the response is reliable or not and its likeliness to be aliased.

## VII. Conclusion

Guaranteeing sufficient entropy and improving the reliability of PUFs is a big progress in order to achieve massive adoption for their everyday use. Several methodologies have been proposed to calculate the reliability. In this paper we have presented a simulation based method for the estimation of reliability and bitaliasing of an RO-PUF based on a counter difference threshold. The method allows the early detection of unreliable challenges and provides the likelihood of challenges being aliased. This information along with filtering techniques can be used to develop an online PUF test strategy in order to achieve 100% reliable PUFs. Moreover, this methodology can be tuned to the characteristic variations of other technologies, different environmental conditions and different number of RO stages.

## References


[1] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld, "Physical one-way functions," *Science*, vol. 297, no. 5589, pp. 2026–2030, 2002.

[2] B. Gassend, D. Clarke, M. Van Dijk, and S. Devadas, "Delay-based circuit authentication and applications," in *Proceedings of the 2003 ACM symposium on Applied computing*, 2003, pp. 294–301.

[3] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Proceedings of the 44th Annual Design Automation Conference*, ser. DAC '07, San Diego, California: Association for Computing Machinery, 2007, pp. 9–14, ISBN: 9781595936271. DOI: 10.1145/1278480.1278484. [Online]. Available: https://doi.org/10.1145/1278480.1278484.

[4] A. Maiti, J. Casarona, L. McHale, and P. Schaumont, "A large scale characterization of ro-puf," in *2010 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*, IEEE, 2010, pp. 94–99.

[5] S. Katzenbeisser, Ü. Kocabaş, V. Rožić, A.-R. Sadeghi, I. Verbauwhede, and C. Wachsmann, "Pufs: Myth, fact or busted? a security evaluation of physically unclonable functions (pufs) cast in silicon," in *International Workshop on Cryptographic Hardware and Embedded Systems*, Springer, 2012, pp. 283–301.

[6] R. Maes, A. Van Herrewege, and I. Verbauwhede, "Pufky: A fully functional puf-based cryptographic key generator," in *International Workshop on Cryptographic Hardware and Embedded Systems*, Springer, 2012, pp. 302–319.

[7] Z. Cherif, J.-L. Danger, F. Lozac'h, Y. Mathieu, and L. Bossuet, "Evaluation of delay pufs on cmos 65 nm technology: Asic vs fpga," in *Proceedings of the 2nd International Workshop on Hardware and Architectural Support for Security and Privacy*, 2013, pp. 1–8.

[8] R. Maes, "An accurate probabilistic reliability model for silicon pufs," in *International Conference on Cryptographic Hardware and Embedded Systems*, Springer, 2013, pp. 73–89.

[9] M. Bhargava and K. Mai, "An efficient reliable puf-based cryptographic key generator in 65nm cmos," in *2014 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, IEEE, 2014, pp. 1–6.

[10] M. Barbareschi et al, "A ring oscillator-based identification mechanism immune to aging and external working conditions," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. PP, pp. 1–23, Aug. 2017. DOI: 10.1109/TCSI.2017.2727546.

[11] N. Karimi, J.-L. Danger, and S. Guilley, "Impact of aging on the reliability of delay pufs," *Journal of Electronic Testing*, vol. 34, no. 5, pp. 571–586, 2018.

[12] A. Schaub, J.-L. Danger, S. Guilley, and O. Rioul, "An improved analysis of reliability and entropy for delay pufs," in *2018 21st Euromicro Conference on Digital System Design (DSD)*, 2018, pp. 553–560. DOI: 10.1109/DSD.2018.00096.

[13] H. Martin et al, "On the reliability of the ring oscillator physically unclonable functions," in *2019 IEEE 4th International Verification and Security Workshop (IVSW)*, IEEE, 2019, pp. 25–30.

[14] A. Ali Pour et al, "Puf enrollment and life cycle management: Solutions and perspectives for the test community," in *2020 IEEE European Test Symposium (ETS)*, 2020, pp. 1–10. DOI: 10.1109/ETS48528.2020.9131578.

[15] A. Schaub, J.-L. Danger, O. Rioul, and S. Guilley, "The big picture of delay-puf dependability," in *2020 European Conference on Circuit Theory and Design (ECCTD)*, IEEE, 2020, pp. 1–4.

[16] S. V. Gutierrez, G. Di Natale, and E.-I. Vatajelu, "On-line reliability estimation of ring oscillator puf," in *2022 IEEE European Test Symposium (ETS)*, 2022, pp. 1–2. DOI: 10.1109/ETS54262.2022.9810418.