Order Statistics based Collision Analysis for PUFs

Girish Vaidya, Chandramani Singh, T. V. Prabhakar
Indian Institute of Science, Bangalore, India
Email: {vaidyab, chandra, tvprabs}@iisc.ac.in

***Abstract*—Physically unclonable functions (PUFs) exploit the inherent manufacturing variations for generating a device identifier. However, different devices may map to the same identifier causing a "collision". It is imperative to determine the probability of such collisions before a PUF is deployed for an application. We present a framework that computes the collision probabilities based on its inter-device and intra-device variations. This framework could be used for determining the collision probabilities, tuning the PUF attributes as well as to compare different PUF implementations. We demonstrate the use of our framework for real-world applications by comparing the results from our analyses with data from experiments and numerical simulation.**

***Index Terms*—Physically unclonable function(PUF), Device ID, Collision probability**

2021 National Conference on Communications (NCC) | 978-1-6654-4177-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/NCC52529.2021.9530163

I. Introduction and Motivation

Physically unclonable functions (PUFs) are emerging as an alternative to conventional methods for generating a device identifier. PUFs exploit the random variations in the device parameters occurring due to manufacturing tolerances to generate distinct device identity. For instance, the random variations in the offset voltages present in the DAC and ADC are exploited to create a PUF using DAC-ADC back-to-back connection [1]. Ring oscillator PUFs (RO PUFs)[2], Arbiter PUFs[3], PUFs based on differential-pair mismatch[4], memory based PUFs[5] are a few other examples of PUFs described in the literature.

Device parameter measurements can be noisy. Variation of the parameter values across measurements for the same device is referred to as intra-device variation. On the other hand, variation of the mean parameter values across different devices is referred to as inter-device variation. Parameter values of the devices are mapped to identifiers for device identification. For correct identification, it is required that the device identifier be unique and repeatable, i.e. the identifiers of different devices must be different and repeated measurements for any device must give the same identifier. However, due to noisy parameter measurements and bounded manufacturing variations, parameters of two devices may map to the same identifier. This situation is referred to as "collision". We provide an alternate definition of collision that is amenable to analysis. To ensure that the collision probabilities are kept under check, we require that inter-device variation be large and intra-device variation be small. Clearly, for a large deployment, the mean parameter values of these devices come closer leading to higher collision probabilities.

A systematic estimation of collision probabilities is essential before using a PUF for any application. This analysis assists in determining the maximum number of devices to ensure that the collision probabilities are within permissible limits. We provide a framework to study the impact of inter-device and intra-device variations on collision probabilities.

A. Related work

Several approaches have been adopted to analyse and improve the uniqueness and repeatability of the PUF generated identifier. RO PUF[2] proposes the use of error-correcting syndrome like BCH code to generate redundant bits. The inter-device and intra-device parameter variations are computed experimentally, and the false-positive and the false-negative rates are computed probabilistically from these variations. Similarly, Arbiter PUF[3] estimates the number of distinguishable devices based on inter-device and intra-device variations. These approaches specific for digital PUFs can not be mapped directly for analysis of analog PUFs. Several analog PUFs convert the analog responses into digital codes and hence use measures to improve the robustness of the PUF identifier. A PCB-PUF[6] measures PCB impedance of traces and uses these as signatures. To minimise the impact of noise, averaging over multiple measurements is performed, and further few least significant bits are discarded. The PUF based on DAC-ADC connection[1] suggests binning of codes to improve the repeatability.

B. Our contribution

The requirement for a unique device identifier comes with a cost. A few applications can tolerate collisions, while PUF's infrastructure usage would require ideally zero collisions. e.g., when the PUF identifier is used for generation of authentication key, it is imperative that the identifier is collision-free. However, if the identifier is used for detecting a node replacement, a few collisions might be permissible. We provide a framework to estimate various collision probabilities for a generic PUF as a function of number of devices, inter-device and intra-device parameter variations. A few programmable attributes during PUF implementation, such as accumulation count [7] or averaging count [8], impact intra-device variation. For instance, the authors in [7] suggest the accumulation counts be increased as the number of devices in a deployment increases. Our analysis can lead to a mechanism to tune these attributes based on the number of devices in the deployment. Our framework is generic and can be used across PUF designs. Thus, this approach could be used to compare two different PUF implementations or to analyse any future design of PUFs.

We have also performed experiments with 38 devices, measuring three different parameters OSC, ADC_1 and ADC_2 for each device. We have 5000 measurements for each parameter for these devices. The experimental data validates our

This work was supported by Centre for Networked Intelligence (CNI), a Cisco CSR initiative, under the division of EECS and RBCCPS, IISc.

978-1-6654-4177-3/21/$31.00 ©2021 IEEE

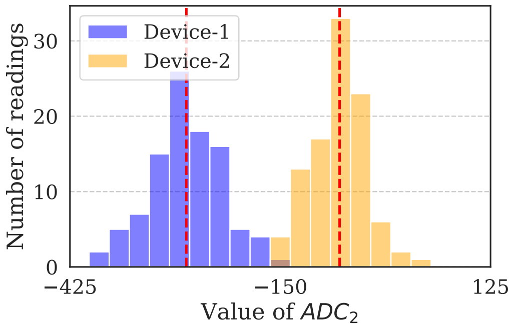


Fig. 1: Histogram of ADC_2 for two devices showing normal distribution for intra-device variation

modeling assumptions and the subsequent analysis. We make the following unique contributions:

- We derive the probability of there being no collision for a generic PUF as a function of number of devices and standard-deviation of inter-device variation.
- We derive an upper bound on probability of there being k distinguishable (collision-free) devices.
- We also provide an expression for the expected number of distinguishable devices.
- We develop an experimental setup and collect data from it to validate our analytical findings. We also use simulation to verify analytical results for large number of devices.

II. Background

A. Physically unclonable functions

Due to process variations, two devices, though fabricated from the same fab and same process are not identical. A PUF exploits this inherent randomness introduced during manufacturing and creates a unique 'fingerprint', an identity signature or trust anchor for the device[9][10]. Unlike other identification methods, the identity generated by the PUF is inherent to the device and is not externally attached. The potential applications for PUFs include anti-counterfeiting, identification, authentication, and key-generation[10].

PUF properties: We discuss two properties necessary for using PUF based fingerprints as device identifiers, viz. *repeatability* and *uniqueness*.

- **Repeatability:** Due to the inherent noise present in all electronic systems, the device parameters have random variations. The environmental factors like changes in temperature and voltage further add to the variations. These intra-device variations must be within acceptable limit so that the device signatures are repeatable.
- **Uniqueness:** Each device should generate a signature which is different from other devices. In other words, two different devices though having same hardware configuration and executing same software code, should generate distinct signatures. As measured values of device parameters are noisy, we require that their means be as-far-as-possible from one another.

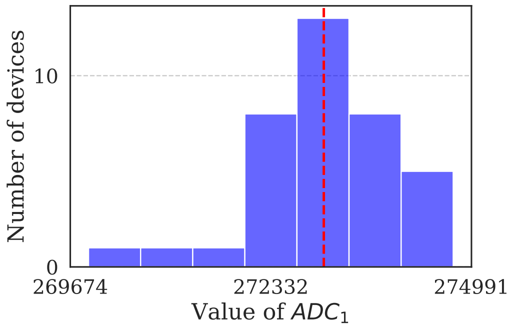


Fig. 2: Histogram of ADC_1 showing normal distribution for inter-device variation

B. Assumptions

We assume that means of parameter values of different devices follow a normal distribution. As reported by several semiconductor manufacturers, for many parameters, their variations across devices exhibit normal distribution[11][12]. The monte-carlo analysis for parameters like offset voltage of a comparator also shows a normal distribution [8]. We also confirm this assumption through our experiments.

We also assume that measurement noises for all the devices have identical distribution. i.e. intra-device variations for all devices are identically distributed. In particular, parameter values for two devices may have different mean but will have same variance. The authors in [13] observe that intra-device variations are also normally distributed. However, our analysis does not assume intra-device variations being normal.

We illustrate intra-device and inter-device variations through Figure 1 and Figure 2 respectively. These figures are based on our measurements of ADC_1 and ADC_2 of 38 devices.In Figure 1, we plot the histograms of 100 measurements of ADC_2 for two devices. The parameter values for device 1 are in the range of (-378, -145), whereas those for device 2 are in the range of (-158, 44). The means of parameter values for device 1 and device 2 are -273 and -72, respectively, and the standard deviations are 47 and 36 respectively. We also see that intra-device variations of parameter values of the two devices resemble normal distributions.

In Figure 2, we plot histogram of mean parameter values of ADC_1 for all the 38 devices. We see that these means also have approximately normal distribution. We can attribute poor resemblance with normal distribution to the small number of devices in our experiment.

C. Collisions

We first illustrate the phenomenon of collision between two devices. We then provide an alternate definition of collision that is amenable to analysis.

PUF based device identification consists of two phases:

- Device binning: For each device, the designated parameter is measured n_B ($n_B \gg 1$) times. Using the measured values of all the devices as training data and supervised learning(classification), the range of potential values is divided into bins corresponding to each of the devices.
- Device identification: Whenever a device is required to be identified, the parameter is measured $n_I (n_I \ll n_B; n_I$

can be one) times. The device is then identified to correspond to the bin that contains the empirical average of the n_I measured values.

If during device identification, a device is mapped to the bin corresponding to another device, the former one is said to collide with the latter one. A device is said to be in collision if it collides with any other device.

Let there be N devices with mean parameter values (for the designated parameter) $X_1, \ldots, X_N$. Following the discussion in section II-B, X_n's are i.i.d. Gaussian random variables say $X_n \sim N(\mu, \sigma^2)\ \forall n$; our analysis and results are insensitive to μ, and so, we assume $\mu = 0$ in simulation. Let $X_{(1)}, \ldots, X_{(N)}$ be the order statistics of $X_1, \ldots, X_N$; $X_{(1)}, \ldots, X_{(N)}$ is a permutation of $X_1, \ldots, X_N$ such that $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(N)}$. In particular,

$$X_{(1)} = \min_{1 \leq n \leq N} X_n \text{ and } X_{(N)} = \max_{1 \leq n \leq N} X_n.$$

Let intra-device variations (measurement noises) for the devices, which are symmetric, zero-mean random variables, have distribution G. Also, let $Y_1, \ldots, Y_N$ denote the empirical means of n_B measurements for the devices during the device binning process. Then

$$\mathbb{P}(Y_n - X_n \leq y) = G^{n_B}(n_B y) \coloneqq \overline{G}^{n_B}(y),$$

where G^{n_B} denotes the distribution of sum of n_B i.i.d. random variables, each with distribution G. Let $Y_{(1)}, \ldots, Y_{(N)}$ be the ordered statistics of $Y_1, \ldots, Y_N$. We make the following two assumptions regarding the data collection and binning process.

i) For $1 \leq n \leq N$, $\left[\frac{Y_{(n-1)}+Y_{(n)}}{2}, \frac{Y_{(n)}+Y_{(n+1)}}{2}\right]$ is the bin corresponding to the device with empirical mean $Y_{(n)}$; we assume $X_{(0)} = Y_{(0)} = -\infty$ and $X_{(N+1)} = Y_{(N+1)} = +\infty$.

ii) For any two devices n_1 and n_2, $Y_{n_1} < Y_{n_2}$ if and only if $X_{n_1} < X_{n_2}$. This is a reasonable assumption since typically $n_B \gg 1$.

Now let $Z_1, Z_2 ... Z_N$ be the empirical means of n_I measurements for the devices during the identification process. Clearly, for all n, $\mathbb{P}(Z_n - X_n \leq z) = \overline{G}^{n_I}(z)$. Let k be the index of the device with mean parameter value $X_{(n)}$ and empirical mean (during binning process) $Y_{(n)}$. The probability that k^{th} device is in collision can be expressed as follows:

$$P_{coll} = \mathbb{P}\left(Z_k \leq \frac{Y_{(n-1)} + Y_{(n)}}{2}\right) + \mathbb{P}\left(Z_k > \frac{Y_{(n)} + Y_{(n+1)}}{2}\right).$$

Let us see the two terms on the right hand side separately.

$$\begin{aligned}
&\mathbb{P}\left(Z_k \leq \frac{Y_{(n-1)} + Y_{(n)}}{2}\right) \\
&= \mathbb{P}((Z_k - X_{(n)}) \leq \frac{Y_{(n-1)} - X_{(n-1)} + Y_{(n)} - X_{(n)}}{2} \\
&\qquad - \frac{X_{(n)} - X_{(n-1)}}{2}) \\
&= \mathbb{P}\left(Z'_k \leq Y'_{(n)} - \frac{\Delta_{n-1}}{2}\right)
\end{aligned}$$

where, for $i = 1, \ldots, N$, $Y'_{(n)} \sim \overline{G}^{2n_B}$, $Z'_k \sim \overline{G}^{n_I}$ and $\Delta_n, n = 0, \ldots, N$ are spacings between order statistics $X_{(0)}, \ldots, X_{(N+1)}$;

$$\Delta_n \coloneqq X_{(n+1)} - X_{(n)}, n = 0, \ldots, N.$$

Continuing, given a constant $d_{th} > 0$,

$$\begin{aligned}
&\mathbb{P}\left(Z_k \leq \frac{Y_{(n-1)} + Y_{(n)}}{2}\right) \\
&= \mathbb{P}\left(Z'_k \leq Y'_{(n)} - \frac{\Delta_{n-1}}{2} \middle| \Delta_{n-1} > d_{th}\right) \mathbb{P}\left(\Delta_{n-1} > d_{th}\right) \\
&\quad + \mathbb{P}\left(Z'_k \leq Y'_{(n)} - \frac{\Delta_{n-1}}{2} \middle| \Delta_{n-1} \leq d_{th}\right) \mathbb{P}\left(\Delta_{n-1} \leq d_{th}\right) \\
&\leq \mathbb{P}\left(Z'_k \leq Y'_{(n)} - \frac{d_{th}}{2}\right)\left(1 - \mathbb{P}(\Delta_{n-1} \leq d_{th})\right) \\
&\quad + \mathbb{P}\left(Z'_k \leq Y'_{(n)}\right) \mathbb{P}\left(\Delta_{n-1} \leq d_{th}\right) \\
&= \mathbb{P}\left(Z'_k - Y'_{(n)} \leq -\frac{d_{th}}{2}\right) \\
&\quad + \left(\frac{1}{2} - \mathbb{P}\left(Z'_k - Y'_{(n)} \leq -\frac{d_{th}}{2}\right)\right) \mathbb{P}\left(\Delta_{n-1} \leq d_{th}\right) \\
&= Q(d_{th}) + \left(\frac{1}{2} - Q(d_{th})\right) \mathbb{P}(\Delta_{n-1} \leq d_{th}),
\end{aligned}$$

where $Q(d_{th}) \coloneqq \mathbb{P}\left(Z'_k - Y'_{(n)} \leq -\frac{d_{th}}{2}\right)$. The second last equality holds since $\mathbb{P}\left(Z'_k - Y'_{(n)}\right) = \frac{1}{2}$, which in turn follows from the fact that Z'_k and $Y'_{(n)}$ are independent, symmetric, zero-mean random variables. Similarly, using symmetry of intra-device variations,

$$\begin{aligned}
&\mathbb{P}\left(Z_k > \frac{Y_{(n)} + Y_{(n+1)}}{2}\right) \\
&\quad \leq Q(d_{th}) + \left(\frac{1}{2} - Q(d_{th})\right) \mathbb{P}(\Delta_n \leq d_{th})
\end{aligned}$$

and so,

$$\begin{aligned}
P_{coll} &\leq 2Q(d_{th}) + \left(\frac{1}{2} - Q(d_{th})\right) (\mathbb{P}(\Delta_{n-1} \leq d_{th}) \\
&\quad + \mathbb{P}(\Delta_n \leq d_{th})) \\
&\leq 2Q(d_{th}) + (1 - 2Q(d_{th}))\overline{P}_{coll}(d_{th}),
\end{aligned}$$

where $\overline{P}_{coll}(d_{th}) \coloneqq \mathbb{P}(\min(\Delta_{n-1}, \Delta_n) \leq d_{th})$. Hence, given parameters $\mu, \sigma^2, G, n_B, n_I$ and d_{th}, an upper bound on P_{coll} can be obtained via computing $\overline{P}_{coll}(d_{th})$. Clearly, the tightness of this bound is affected by the choice of d_{th}. We do not attempt to determine the optimal d_{th}, which depends on other system parameters. Rather, we analyse $\overline{P}_{coll}(d_{th})$ for arbitrary $d_{th} > 0$. In particular, we say that a device collides with another device if their mean parameter values are closer than d_{th} and analyse the probability of a device being in collision.

III. Collision Probabilities

We begin with estimating the asymptotic probability of none of the devices being in collision. We then derive upper bounds on asymptotic probability of at most K devices being in collision. Finally, we provide asymptote of expected number of devices in collision. Throughout this section, we use f and F to denote the density and the distribution of inter-device variation, respectively;

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Let N_{coll} denote the number of devices in collision. We also define $\Delta_{(0)},\ldots,\Delta_{(N)}$ to be the order statistics of the spacings $\Delta_0,\ldots,\Delta_N$. Note that the order statistics $X_{(i)},\Delta_{(i)}, i = 0,\ldots,N$ depend on N though we do not show this dependence explicitly. Note that none of the devices is in collision if and only if $\Delta_{(1)} > d_{th}$. Also observe that $\int f^2(x)dx = \frac{1}{2\sqrt{\pi}\sigma}$, and so $f \in L^2(\mathbb{R})$.Hence, defining $\lambda = \frac{1}{2\sqrt{\pi}\sigma}$ and following [14, Theorem 1], for any $\overline{d}_{th} > 0$,

$$\lim_{N\to\infty} \mathbb{P}(N^2\Delta_{(1)} > \overline{d}_{th}) = \exp\left(-\lambda \overline{d}_{th}\right).$$

Hence, for large N, we have the following approximation:

$$\mathbb{P}(N_{coll} = 0) = \mathbb{P}(\Delta_{(1)} > d_{th}) \approx \exp\left(-\lambda d_{th} N^2\right). \quad (1)$$

Note that if N is increased while keeping d_{th} fixed, both the sides in the above approximation approach zero. The relative error in the approximation may grow substantially as N is increased. However, we are interested in a regime where probability of no device being in collision is bounded away from zero, i.e., $N = O\left(\frac{1}{\sqrt{d_{th}}}\right)$. We compare the results from approximate formula (1) with simulation and experimental data in Figure 4.

Many applications permit a non-zero number of collisions without compromising their performance. Note that for at least K devices to be in collision, $\Delta_{(\lceil \frac{K}{2} \rceil)} \le d_{th}$, the two events being identical if none of the devices is associated with more than one of the lower spacings $\Delta_{(1)},\ldots,\Delta_{(\lceil \frac{K}{2} \rceil)}$. On the other hand, $\Delta_{(K-1)} \le d_{th}$ implies that at least K device will be in collision, the two events being identical if all of the lower spacings $\Delta_{(1)},\ldots,\Delta_{(K-1)}$ are contiguous. Figure 3 shows two scenarios that illustrate these assertions. We thus see that

$$\mathbb{P}(\Delta_{(\lceil \frac{K+1}{2} \rceil)} > d_{th}) \le \mathbb{P}(N_{coll} \le K) \le \mathbb{P}(\Delta_{(K)} > d_{th}). \quad (2)$$

In a typical realisation with a large number of devices it is unlikely that any device will be associated with more than one of the smallest $\lceil \frac{K+1}{2} \rceil$ spacings. We thus expect the first inequality to be tighter. We confirm the same in Figure 8 through simulations. In the following we provide approximations for $\mathbb{P}(\Delta_{(K)} > d_{th})$ for $K \ge 2$.

Following the analysis in [14], $N^2(\Delta_{(2)} - \Delta_{(1)}),\ldots,N^2(\Delta_{(K)} - \Delta_{(K-1)})$ are asymptotically independent exponential random variables with parameters $\lambda, \lambda(1-\frac{1}{N}),\ldots,\lambda(1-\frac{K-1}{N})$, respectively. Hence from [15], $N^2\Delta_{(K)}$ is asymptotically a hypoexponential random variable. In particular, for large N, we have the following approximation:

$$\mathbb{P}(\Delta_{(K)} > d_{th}) \approx \prod_{i=0}^{K-1} \lambda_i \sum_{j=0}^{K-1} \frac{\exp(-\lambda_j d_{th} N^2)}{\lambda_j \prod\limits_{\substack{l=0\\ l\neq j}}^{K-1} (\lambda_l - \lambda_j)} \quad (3)$$

where $\lambda_i = \lambda\left(1 - \frac{i}{N}\right), i = 0,\ldots,K-1$.

We can obtain a simpler, albeit crude, approximation by assuming that $N^2\Delta_{(1)}, N^2(\Delta_{(2)} - \Delta_{(1)}),\ldots,N^2(\Delta_{(K)} - \Delta_{(K-1)})$ are asymptotically i.i.d. exponential random variables with parameter λ. In this case,

$$\mathbb{P}(\Delta_{(K)} > d_{th}) \approx \sum_{i=0}^{K-1} \frac{(\lambda d_{th} N^2)^i}{i!} \exp(-\lambda d_{th} N^2). \quad (4)$$

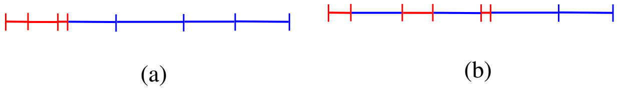


Fig. 3: Relation between number of short spacings ($\le d_{th}$) and number of devices in collision. Three short spacings, depending on how they are arranged, may cause at least four devices (Subfigure (a)) and at most six devices (Subfigure (b)) to be in collision.

We use both (3) and (4) to approximate the probabilities of at most K devices being in collision. We illustrate our observations in Figure 6.

Finally, we estimate the expected number of distinguishable (collision-free) devices based on joint distribution of spacings around various devices. Let us define $N_{coll-free}$ to be the number of such devices. Then

$$\begin{aligned}\mathbb{E}[N_{coll-free}] &= \sum_{k=1}^{N} \mathbb{1}\{\text{the device with parameter value } X_{(k)} \\ &\qquad \text{is collision free}\} \\ &= \sum_{k=1}^{N} \mathbb{P}(\Delta_{k-1} > d_{th}, \Delta_k > d_{th}).\end{aligned}$$

Following [16], the joint asymptotic distributions of adjacent spacings around $X_{(k)}$ depend on scaling of k vis-a-vis N. If (a) $\frac{k}{N} \to p \in (0,1)$ or (b) $\frac{k}{N} \to 0$ and $k \to \infty$ or (c)$\frac{k}{N} \to 1$ and $N-k \to \infty$, the left and the right spacings are asymptotically i.i.d. exponential random variables. The asymptotic joint distributions of spacings are more complex for the extreme values of k (fixed k or fixed $N-k$). Asymptotically all the devices are covered in cases (a), (b) and (c). For our estimate of $\mathbb{E}[N_{coll-free}]$, we assume that k=2,…,$N-1$, the two adjacent spacings around X_k are i.i.d. exponential random variables. In particular, from [16]

$$\begin{aligned}&\mathbb{P}\left(Nf\left(F^{-1}\left(\frac{k}{N}\right)\right)\Delta_{k-1} > \overline{d}_{th},\right.\\ &\qquad\left. Nf\left(F^{-1}\left(\frac{k}{N}\right)\right)\Delta_k > \overline{d}_{th}\right) \approx \exp(-2\overline{d}_{th}).\end{aligned}$$

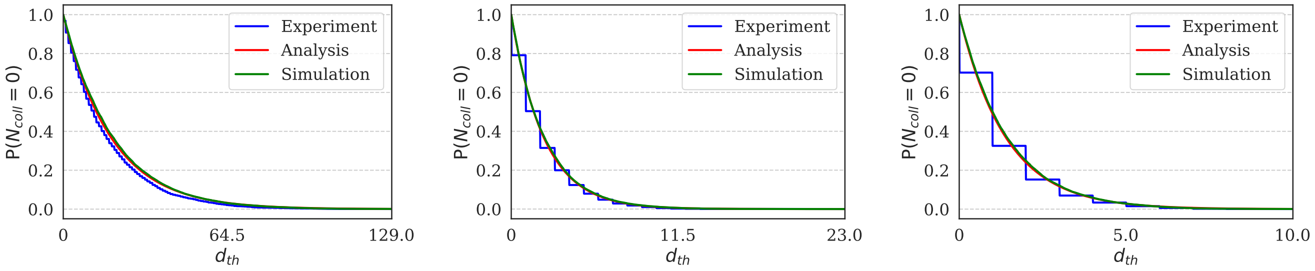
Also,

$$\begin{aligned}&\mathbb{P}\left(Nf\left(F^{-1}\left(\frac{1}{N}\right)\right)\Delta_1 > \overline{d}_{th}\right)\\ &= \mathbb{P}\left(Nf\left(F^{-1}\left(1-\frac{1}{N}\right)\right)\Delta_{N-1} > \overline{d}_{th}\right) \approx \exp(-\overline{d}_{th}).\end{aligned}$$

Accordingly, substituting $\overline{d}_{th} = Nf\left(F^{-1}\left(\frac{k}{N}\right)\right)$, $k = 1,\ldots,N-1$, in the above equations,

$$\begin{aligned}\mathbb{E}[N_{coll-free}] \approx &\sum_{k=2}^{N} \exp\left(-2Nf\left(F^{-1}\left(\frac{k}{N}\right)\right)d_{th}\right)\\ &+2\exp\left(-2Nf\left(F^{-1}\left(\frac{1}{N}\right)\right)d_{th}\right). \quad (5)\end{aligned}$$

We compare the mean number of collision-free devices from analysis and from simulation in Figure 9.



(a) OSC

(b) ADC_1

(c) ADC_2

Fig. 4: Comparison of analytical values with experimental and simulation results for $\mathbb{P}(N_{coll} = 0)$ across 3 parameters with $N = 38$.

IV. Evaluation and Results

We installed 38 devices for collecting the signatures. These devices were deployed for over a month during which 5000 values for three different parameters viz. OSC, ADC_1 and ADC_2 were collected. While OSC exploits the variation in clock oscillators, ADC_1 and ADC_2 exploit the variations in ADC offset across the devices. Please see [7] for more details about the setup and the parameters.

A. Probability of all devices being collision-free

Variation of $\mathbb{P}(N_{coll} = 0)$ with d_{th}: In order to compute $\mathbb{P}(N_{coll} = 0)$, we measure 5000 sets of mean parameter values (each set containing 38 values across devices) for each of the three parameters viz. OSC, ADC_1 and ADC_2. For each parameter, the fraction of sets in which $\Delta_{(1)} > d_{th}$ gives an estimate of $\mathbb{P}(N_{coll} = 0)$. We also use this experimental data to compute inter-device standard deviation, σ, for each of the three parameters. We find these standard deviations to be σ=7642, 912 and 569, respectively. We then use these standard deviations to simulate parameter means of 38 devices. We generate 10000 sets of mean parameter values for each of the parameters and compute $\mathbb{P}(N_{coll} = 0)$ as explained above for experimental data. Finally, we use the same standard deviations in (1) to obtain analytical values of $\mathbb{P}(N_{coll} = 0)$. We see in Figure 4 that both simulation and analytical values of $\mathbb{P}(N_{coll} = 0)$ closely match the experimental values. The step nature of experimental plots in Figures 4b and 4c can be attributed to quantized measurements of ADC_1 and ADC_2 - number of quantization levels is less in case of ADC_2. However, as expected, $\mathbb{P}(N_{coll} = 0)$ decreases with d_{th} in all the three cases.

Variation of $\mathbb{P}(N_{coll} = 0)$ with number of devices: We show variation of $\mathbb{P}(N_{coll} = 0)$ with N in Figure 5. We use $\sigma = 100$ and $d_{th} = 10^{-5}, 2\times10^{-5}, 5\times10^{-5}$ in this simulation. As expected, for each value of d_{th}, $\mathbb{P}(N_{coll} = 0)$ decreases with the number of devices.

B. Probability of at-most K devices being in collision

Comparison of Hypoexponential and Erlang approximations: We compute $\mathbb{P}(\Delta_{(K)} > d_{th})$ using simulation in the same way as we computed $\mathbb{P}(N_{coll} = 0)$ in Section IV-A. We use $N = 200$, $\sigma = 10$, $d_{th} = 0.01$ and vary K from 1 to 20. To estimate $\mathbb{P}(\Delta_{(K)} > d_{th})$, we use 10000 samples of mean parameter values for each of the above sets of parameters. We then compute $\mathbb{P}(\Delta_{(K)} > d_{th})$ using the Hypoexponential and

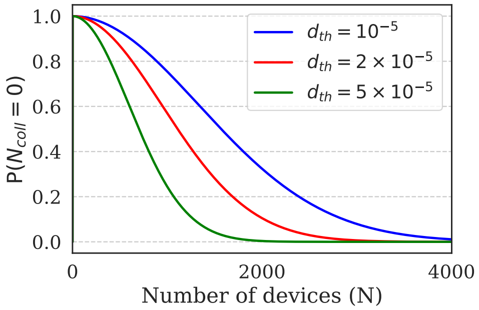


Fig. 5: $\mathbb{P}(N_{coll} = 0)$ obtained from analysis for different values of d_{th}. We use $\sigma = 100$.

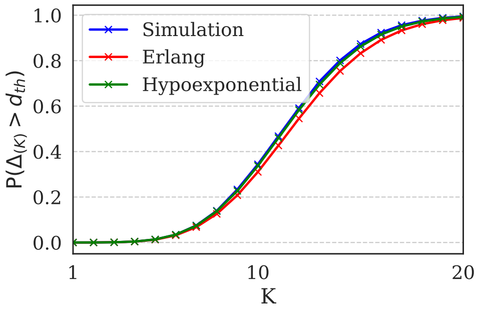


Fig. 6: Comparison of Erlang and Hypoexponential approximations for $\mathbb{P}(\Delta_{(K)} > d_{th})$ with simulation values. We use N=200, σ=10 and $d_{th} = 0.01$.

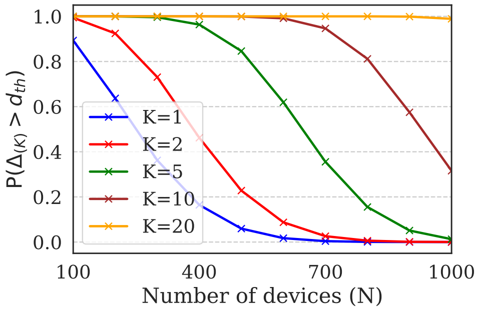


Fig. 7: $\mathbb{P}(\Delta_{(K)} > d_{th})$ obtained through Hypoexponential approximation for different values of K. We use d_{th}=0.004 and σ=100.

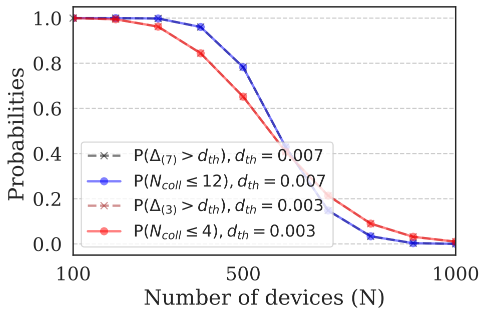


Fig. 8: Comparison of $\mathbb{P}(N_{coll} \leq K)$ with $\mathbb{P}(\Delta_{(\lceil \frac{K+1}{2} \rceil)} \leq d_{th})$ for two different cases. We use σ=100.

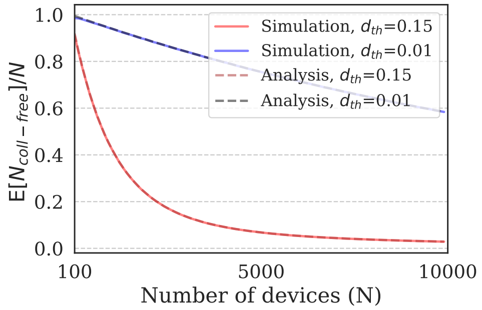


Fig. 9: Comparison of analytical and simulation values of $\mathbb{E}[N_{coll-free}]/N$. We use $\sigma = 100$.

Erlang approximations as in (3) and (4), respectively, for the same set of parameters. We show variation of $\mathbb{P}(\Delta_{(K)} > d_{th})$ with K in Figure 6. As we see, $\mathbb{P}(\Delta_{(K)} > d_{th})$ increases with K. Also Hypoexponential approximation is much more accurate than Erlang approximation. So, we use Hypoexponential approximation in the following.

Variation of $\mathbb{P}(\Delta_{(K)} > d_{th})$ *with number of devices:* We show variation of $\mathbb{P}(\Delta_{(K)} > d_{th})$ with N in Figure 7. Here we use $\sigma = 100$, $d_{th} = 0.004$ and five different values of K viz. K = 1,2,5,10, and 20. We vary N from 100 to 1000. Expectedly, $\mathbb{P}(\Delta_{(K)} > d_{th})$ decreases with N.

Comparison of $\mathbb{P}(N_{coll} \leq K)$ *with* $\mathbb{P}(\Delta_{(\lceil \frac{K+1}{2} \rceil)} \leq d_{th})$*:* We compare these two probabilities in Figure 8 for two different sets of parameters viz. $\sigma = 100$, $K = 12$, $d_{th} = 0.007$ and $\sigma = 100$, $K = 4$, $d_{th} = 0.003$. We again vary N from 100 to 1000. We also use 10000 samples of mean parameter values for each of the above sets of parameters to estimate $\mathbb{P}(N_{coll} \leq K)$. We use (3) for a bound on $\mathbb{P}(\Delta_{(\lceil \frac{K+1}{2} \rceil)} \leq d_{th})$. As discussed in Section III, $\mathbb{P}(N_{coll} \leq K)$ is well approximated by $\mathbb{P}(\Delta_{(\lceil \frac{K+1}{2} \rceil)} \leq d_{th})$ for all the considered sets of parameters.

C. Expected fraction of collision-free devices

We compute the expected fraction of collision-free devices, $\mathbb{E}[N_{coll-free}]/N$, via both simulation and analysis. We use $\sigma = 100$ and two different values of d_{th}, 0.1 and 0.15. We vary N from 100 to 10000. For simulation, we again generate 10000 samples of parameter means for each combination of σ, d_{th} and N, and compute mean for fraction of collision-free for each case. We use (5) for analytical estimates. As can be seen in Figure 9, simulation and analytical results closely match. As expected, fraction of collision-free devices decreases with N. Also, for a given N, this fraction is lower for the higher value of d_{th}.

V. Conclusion and Future work

Through this work, we have presented an analytical framework to estimate the collision probabilities of a generic PUF as a function of number of devices, inter-device and intra-device variations. Our framework could be used for analysing the collision probabilities and tuning the PUF attributes. The comparison of our analytical results with the experimental and simulation results validates our framework.

We plan to enhance our current work by modeling intra-device variations as normal distribution, thus relating d_{th} with intra-device variations. We also plan to study multi-parameter PUF by analysing how the co-relation amongst different parameters impacts the collision probability.

References

[1] A. Duncan, L. Jiang, and M. Swany, "Repurposing SoC analog circuitry for additional COTS hardware security," in *2018 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*. IEEE, 2018, pp. 201–204.

[2] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Design Automation Conference, 2007. DAC'07. 44th ACM/IEEE*. IEEE, 2007, pp. 9–14.

[3] J. W. Lee, D. Lim, B. Gassend, G. E. Suh, M. Van Dijk, and S. Devadas, "A technique to build a secret key in integrated circuits for identification and authentication applications," in *VLSI Circuits, 2004. Digest of Technical Papers. 2004 Symposium on*. IEEE, 2004, pp. 176–179.

[4] S. Deyati, B. Muldrey, A. Singh, and A. Chatterjee, "Design of efficient analog physically unclonable functions using alternative test principles," in *Mixed Signals Testing Workshop (IMSTW), 2017 International*. IEEE, 2017, pp. 1–4.

[5] D. E. Holcomb, W. P. Burleson, K. Fu *et al.*, "Initial SRAM state as a fingerprint and source of true random numbers for RFID tags," in *Proceedings of the Conference on RFID Security*, vol. 7, no. 2, 2007, p. 01.

[6] F. Zhang, A. Hennessy, and S. Bhunia, "Robust counterfeit PCB detection exploiting intrinsic trace impedance variations," in *2015 IEEE 33rd VLSI Test Symposium (VTS)*. IEEE, 2015, pp. 1–6.

[7] G. Vaidya, A. Nambi, T. Prabhakar, S. Sudhakara *et al.*, "IoT-ID: A novel device-specific identifier based on unique hardware fingerprints," in *2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI)*. IEEE, 2020, pp. 189–202.

[8] T. Bryant, S. Chowdhury, D. Forte, M. Tehranipoor, and N. Maghari, "A stochastic approach to analog physical unclonable function," in *Circuits and Systems (MWSCAS), 2016 IEEE 59th International Midwest Symposium on*. IEEE, 2016, pp. 1–4.

[9] S. Joshi, S. P. Mohanty, and E. Kougianos, "Everything you wanted to know about PUFs," *IEEE Potentials*, vol. 36, no. 6, pp. 38–46, 2017.

[10] Y. Gao, S. F. Al-Sarawi, and D. Abbott, "Physical unclonable functions," *Nature Electronics*, vol. 3, no. 2, pp. 81–91, 2020.

[11] "Cascading Precision Op Amp Stages for Optimal AC and DC Performance," https://www.ti.com/lit/an/sboa356/sboa356.pdf, Texas Instruments, 2020.

[12] *Datasheet 8657/8659*, Analog Devices, Feb. 2011.

[13] "Measurement Uncertainty," https://en.wikipedia.org/wiki/Measurement_uncertainty, Wikipedia, [Online; accessed Jan-2021].

[14] S. Molchanov and A. Y. Reznikova, "Limit theorems for random partitions," *Theory of Probability & Its Applications*, vol. 27, no. 2, pp. 310–323, 1983.

[15] S. V. Amari and R. B. Misra, "Closed-form expressions for distribution of sum of exponential random variables," *IEEE Transactions on Reliability*, vol. 46, no. 4, pp. 519–522, 1997.

[16] H. N. Nagaraja, K. Bharath, and F. Zhang, "Spacings around an order statistic," *Annals of the Institute of Statistical Mathematics*, vol. 67, no. 3, pp. 515–540, 2015.