Litho-Aware and Low Power Design of a Secure Current-Based Physically Unclonable Function

Raghavan Kumar, Wayne Burleson
Department of Electrical and Computer Engineering
University of Massachusetts Amherst, MA
{rkumar, burleson}@ecs.umass.edu

***Abstract*—Physically Unclonable Functions (PUFs) are lightweight cryptographic primitives for generating unique signatures from complex manufacturing variations. In this work, we present a current-based PUF designed using a generalized lithographic simulation framework for improving inter-die and inter-wafer uniqueness. The sensitivity of the circuit to manufacturing variations is enhanced by placing the gate structures at pitches closer to forbidden zone, where the sensitivity of Critical Dimension (CD) to the pitch variations is very high. Simulation results show that the litho-aware current based PUF has improved inter- and intra-distance over the conventional current-based PUF. The litho-aware PUF consumes about 0.034 pico joules of energy per response bit, which is substantially better than delay-based PUF implementations.**

***Keywords*—Physically unclonable functions, hardware security, sub-wavelength lithography, forbidden pitches**

I. Introduction

Security of integrated circuits (ICs) and embedded devices is a growing concern because of the ever-increasing usage of electronic devices for performing security related tasks. Majority of the electronic devices are equipped with memory to store and handle sensitive information. Typical examples include smart cards, mobile phones, etc. Consequently, they have become a target for adversaries. So, security must be ensured through various operations such as device identification/authentication and secure communication, while maintaining the confidentiality and integrity of the stored data. To store the secrets in an IC or an embedded device, non-volatile memory can be used. But, this technique suffers from high cost, implementation complexities and security vulnerabilities [1]. An alternative solution is to use a Physically Unclonable Function (PUF) to derive signatures and keys from the unpredictable and unique nature of silicon.

PUF circuits are cryptographic primitives that exploit inherent process variations in integrated circuits to generate unique signatures. A PUF circuit can be envisioned as a function that maps a set of challenges to a set of responses decided by the process variations [1].

One of the main applications of PUFs lies in signature/key generation for device authentication and identification. This application requires that any two responses from two separate PUFs for a particular challenge must have a significant difference. This property of PUFs is referred to as uniqueness, which determines the extent to which a signature is tied to a single device. To ensure stable and uninterrupted authentication, signatures must be stable across all operating conditions. In other words, a PUF is expected to produce the same response for a particular challenge under different operating conditions. This property of a PUF circuit is referred to as its reliability. Finally, signatures from a PUF device must be unpredictable such that even an adversary possessing a subset of challenge-response pairs should be unable to predict the response for a challenge outside the subset.

The process variations can be grouped into "systematic" and "random" variations. In a PUF circuit, it is desirable to have random variations dominant over systematic variations for better uniqueness. While applying design strategies to enhance random variations, care must be taken such that a PUF's reliability remains within the acceptable level. Inter-wafer variations are highly systematic [2]. So, the responses of any two PUF circuits at identical locations from two different wafers might be highly correlated.

In this paper, we present a current-based PUF designed using a generalized lithographic simulation framework adopted for PUF design. The main objective of the lithographic simulation framework is to enhance the sensitivity of the PUF circuit to process variations and improve uniqueness when viewed across dies and wafers. It is well known that forbidden pitches are more prominent in sub-wavelength lithography [3]. Forbidden pitches are often undesirable, as the polygons/structures at these pitches will not be printed to their maximum resolution. However, the sensitivity of critical dimension (CD) is very high to the pitch variations near the forbidden zone. This is used constructively to enhance the impact of process variations on PUF design. This is done by placing gate structures of the transistors at pitches closer to forbidden zone. Such a technique allows the circuit designers to amplify and effectively utilize various sources of lithographic variations including dose, resist thickness, lens imperfections, defocus, etc. in PUF designs. The proposed PUF has been designed using a standard 45nm CMOS technology node and compared to the current-based PUF designed using a conventional approach. To the authors' knowledge, there are only three other works that look at improving the performance metrics of PUFs from lithography level [4], [5], [6].

The rest of this paper is structured as follows. Section II describes the prior work related to this paper. The litho-aware current-based PUF is described in section III. Experimental methodology and results are described in sections IV and V, respectively. Concluding remarks are presented in section VI.

II. Related Work

In this section, we describe the existing works on improving a PUF's quality from lithography level. In [5], novel PUF circuits known as litho-PUFs have been proposed that consider proximity effects, density effects and formation of non-rectangular gates printed during the lithography process. Though the proposed scheme in [5] also uses forbidden pitches to improve the uniqueness of PUFs, the evidence of improving inter-wafer uniqueness was not presented. Also, the impacts of different lithographic variations on polygons placed near the forbidden pitch are not evaluated in [5]. Traditionally, Optical Proximity Correction (OPC) tries to reduce variations, both systematic & random by using the geometric error between the simulated contour and the target as a cost function. In [4], a PUF-aware OPC scheme was described, that tries to reduce the systematic variations and increase random variations in the regions of the mask that contain the PUF circuit. The scheme continues to work as a traditional OPC in non-PUF regions of the mask. The proposed scheme in [4] tries to maximize the variance of the mean edge-placement error (EPE). An improvement in uniqueness of 5% and reliability of 70% compared to conventional OPC was reported in [4]. Similar work on enhancing random variations and suppressing systematic variations was presented in [6]. In order to suppress systematic variations, several layout techniques have been used. However, all the existing works in the literature focus only on improving inter-die variations. Wafer-to-wafer variations are known to be more correlated [2]. This affects the uniqueness of PUFs and has not been extensively explored in the literature.

III. Litho-aware Current-based PUF

The litho-aware current-based PUF is based on the current PUF architecture proposed by Majzoobi et al. in [7]. However, the transistors in the circuit were modified as per the lithographic simulation framework to increase the sensitivity of the PUF circuit to process variations. The overall litho-aware current-based PUF architecture is shown in Figure 1. The basic operation of the current-based PUF is presented in section III-A. The modified current-based PUF as per the lithographic simulation framework is described in section III-B.

A. Description of PUF operation

The operation of current-based PUF is almost similar to an arbiter PUF, as the current PUF is based on linear addition of process variation sensitive currents generated by the circuit. However, in arbiter PUFs, all the stages contribute to the final response by introducing a delay to the propagating signal. However, in a current-based PUF, only the stages selected by the challenge vector contribute to the final response. As shown in Figure 1, the process variation sensitive currents are generated by the current-generation transistors. Based on the external challenge, a subset of currents are selected and combined using select-and-combine (SC) transistors. The two SC transistors fed by a current generation transistor is referred to as the select-and-combine (SC) module. The combined currents are then compared using a current sense-amplifier, which produces either '1' or '0' based on the input currents. The currents to be compared are denoted as I_a and I_b respectively, as shown in Figure 1. For some challenge vectors, both the inputs of a SC module may be at logic '0'. In such a condition, no current will be generated. If only one of the inputs of a SC module is at logic '1', then the generated current directly flows through the SC transistor whose gate voltage is at logic '1'. If both the inputs of a SC module are at logic '1', then the generated current will be split between the SC transistors. Ideally, the current will be split in half. However, the current split ratio may slightly depart from 0.5, if the SC transistors have process variations. It is also essential to size the sense amplifier accordingly to enable fair evaluation. So, the sense amplifier must be designed such that it is tolerant to process variations.

B. Litho-aware design

In the litho-aware current-based PUF design, the transistors that are responsible for generating currents were modified using the lithographic simulation framework. The transistors were fractured to n segments of minimum dimension. Then the fractured segments were placed such that the pitch between any two structures is closer to forbidden zone, where the sensitivity of critical dimension is very high to the pitch variations. The pitch value was obtained by running an extensive set of lithographic simulations under various sources of variations. A detailed explanation on forbidden pitches is presented below.

Forbidden Pitches: In sub-wavelength lithography, forbidden pitches are more prominent. The structures/polygons at some pitches may not print to their complete resolution due to diffraction effects. Those pitches are referred to as forbidden pitches [3]. At forbidden pitches, the printed gate structure will either be highly constricted or broken. Based on simulations as per the methodology described in section IV, we found that the forbidden pitch for the 45nm technology node is around 190nm. Figure 2 shows the presence of forbidden pitches. To determine forbidden pitches, the polygons were created in polysilicon layer, as the gate length often determines the critical dimension.

Near the forbidden pitch, the gate structures are printed with lower resolution. But, they are also highly sensitive to pitch variations and even a small variation in pitch can lead to high variations in the critical dimension. Figure 3 shows the critical dimension sensitivity with respect to pitch variations.

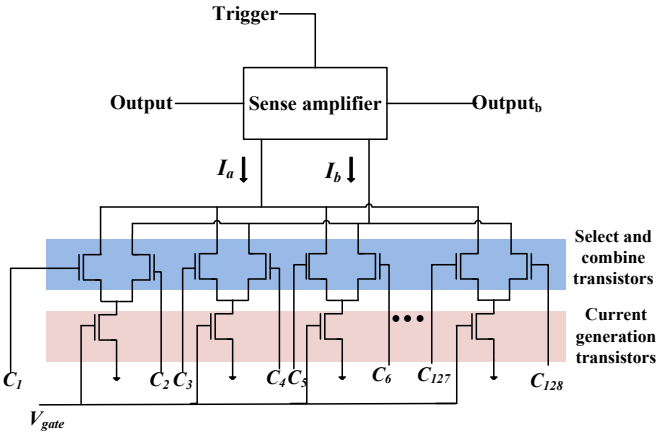

Fig. 1: Current-based PUF architecture [7]. C_i represents i^{th} challenge bit. The inputs to the sense amplifier are the currents I_a and I_b. $output_b$ refers to the complimentary form of the output bit.

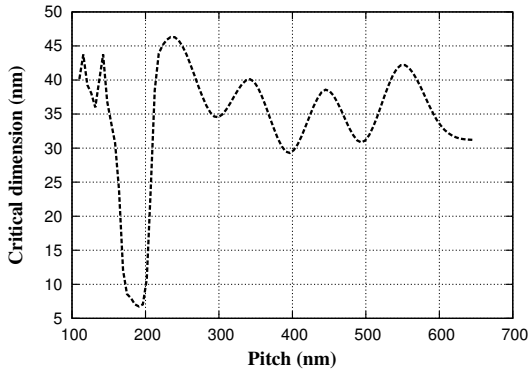

Fig. 2: Impact of pitch on the critical dimension in 45nm technology node. Forbidden pitch can be seen around 190nm.

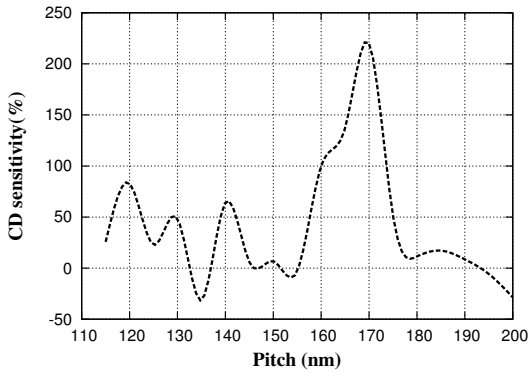

Fig. 3: CD sensitivity with respect to pitch variations

The pitch range is reduced to 200nm, as it falls under the region of interest for this work.

To increase the sensitivity of the circuit to process variations, the layout of the current-based PUF was fed into the lithographic simulation framework. The framework then identified the current generation transistors, fractured them and aligned the fractured segments at a pitch closer to forbidden zone. The other transistors in the SC module and sense amplifier were not modified by the lithographic simulation framework. Our method requires that the transistors must be larger than the minimum dimension so that they can be fractured. This might impose area overheads. However, the percentage increase in area due to our method is very minimal as demonstrated in section V-D.

IV. Experimental Methodology and Framework

In this section, we briefly describe the methodology used in lithographic and circuit simulations to compute the performance metrics of the litho-aware current PUF. All simulations and analyses were based on an industrial standard 45nm CMOS technology node.

Lithographic simulation framework: As described in section III, the lithographic simulation framework modifies the current generation transistors, so that their sensitivity to process variations is enhanced. Apart from the layout modification, the framework also triggers an aerial image simulation over the modified layout using Calibre©Workbench™, an industrial standard environment for generating process models. The simulation was carried out under the presence of various lithographic sources of variations including dose, defocus, resist thickness and lens aberrations. The sources of variations were faithfully modeled so as to depict both inter-die and inter-wafer variations. To emulate inter-wafer variations, wafer-tilt effect was captured and modeled in Calibre Workbench. The amount of variations considered is shown in Table I. Due to the absence of industrial standard data depicting the extent of lithographic variations, the values were chosen conservatively. For aerial imaging simulations, a dipole light source was considered with the wavelength fixed at 193nm. The numerical aperture (NA) of the imaging system was fixed at 1.35.

TABLE I: Lithographic variations considered

	Inter-die	Inter-wafer
Dose	$\pm$ 5%	$\pm$ 2%
Defocus	$\pm$ 5nm	$\pm$ 2nm
Resist thickness	$\pm$ 5%	$\pm$ 5%
Wafer-tilt angle	0	$\pm$ 5 degrees

Once the simulated contour of polygons is obtained from Calibre Workbench, the lithographic simulation framework takes the layout as input and obtains the printed gate length of all the transistors using gate-slicing approach [8]. By applying an extensive set of lithographic sources of variations to the modified layout, the critical dimension values were collected. The obtained critical dimension values were fit within a distribution to allow a smooth transition from lithographic simulations to circuit simulations. The overall simulation flow is shown in Figure 4. The mean (μ) and standard deviation (σ) of the critical dimension distribution obtained by applying inter-die and inter-wafer variation statistics are shown in Table II.

TABLE II: Statistical parameters of the critical dimension distribution

Mean (μ_{CD}^{die})	44.5nm
Std. Deviation (σ_{CD}^{die})	10nm
Mean (μ_{CD}^{wafer})	44.8nm
Std. Deviation (σ_{CD}^{wafer})	8nm

Circuit simulations: To model random variations, threshold voltage (V_{th}) variations were assigned from a Gaussian distribution with a 3σ deviation of 150mV. This value was chosen to be consistent with ITRS specifications.

V. PUF validation Results

This section describes in detail the validation results of the litho-aware current-based PUF. The performance metrics namely uniqueness, reliability and security of the PUF were computed through circuit simulations. To measure the performance metrics, a 128-stage PUF was built and used. An ideal PUF is expected to have 100% uniqueness and 100% reliability [1].

To obtain the best value of n, which corresponds to the number of fractured segments of current generation transistors, experiments were conducted with n={2,3,4,5}. Results indicated that n=3 is an optimal choice while considering the trade-off between performance and area overheads. All the analyses described below assume that the current generation transistors

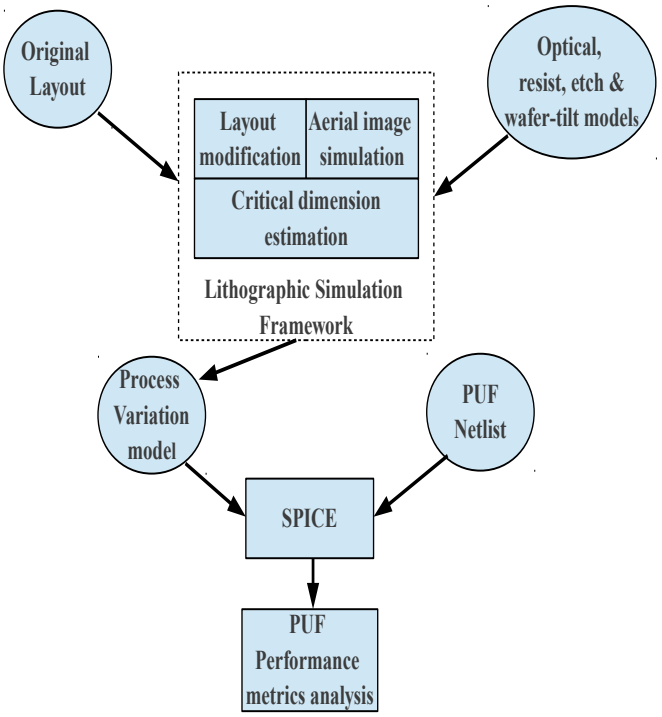

Fig. 4: Overall design flow

were fractured into 3 segments. Also, the performance metrics of litho-aware current-based PUF is compared with respect to the current-based PUF designed using a conventional approach, which we refer to as base PUF in the further sections. The gate voltage of current generation transistors were set at two discrete values (0.5 V_{DD}, 0.7 V_{DD}). Below $0.5V_{DD}$, the responses were highly biased.

A. Uniqueness

As explained in section I, uniqueness refers to the ability of a PUF to produce unique responses for a challenge across different PUF instances. Since a challenge produces a single bit response, 128 challenges were grouped together to form a challenge set. The responses of 128 challenges were grouped together to form a 128-bit response. In order to quantify uniqueness, we use the metric inter-distance (d_{inter}), given by equation 1 [9].

$$d_{inter} = \frac{2}{m(m-1)} \sum_{p=1}^{m-1} \sum_{q=p+1}^{m} \frac{HD(R_p, R_q)}{k} * 100\%. \quad (1)$$

In the above equation, k is the number of bits per responses, m is the number of PUF instances considered and $HD(R_p, R_q)$ refers to the hamming distance between any two responses R_p and R_q from two different instances for the same challenge. In our work, we set $\{m,k\}$ = $\{128,128\}$. Separate experiments were conducted to obtain uniqueness measured across the dies and wafers. The inter-distance results are tabulated in Table III. We can observe that the best case improvement in the uniqueness of litho-aware PUF design is as much as 15% when compared to the base PUF design.

TABLE III: Comparison of litho-aware and base PUF's inter-distance

Type of simulation	Gate voltage	Type of PUF	d_{inter}
die	$0.5V_{DD}$	Base design	0.46
		Litho-aware design	0.49
	$0.7V_{DD}$	Base design	0.47
		Litho-aware design	0.49
wafer	$0.5V_{DD}$	Base design	0.34
		Litho-aware design	0.38
	$0.7V_{DD}$	Base design	0.33
		Litho-aware design	0.38

B. Reliability

Reliability is a measure of consistency of PUF responses over different operating conditions. Temperature and supply voltage fluctuations can degrade the reliability of PUF circuits. A measure of reliability is intra-distance, which is given by equation 2 [9],

$$d_{intra} = \frac{1}{s} \sum_{j=1}^{s} \frac{\text{HD}(R_i, R'_{i,j})}{m} \text{ x } 100\% \quad (2)$$

where R_i is the response of a PUF to a challenge under nominal operating conditions ($T = 25°C$, V_{DD} = $1.1V$), s is the number of samples of response R_i obtained at different operating conditions by changing temperature and supply voltage, $R'_{i,j}$ refers to j^{th} sample of response R_i for a challenge and m is the number of bits in the response.

Intra-distance of litho-aware PUF was computed in the presence of both temperature and supply voltage fluctuations. Temperature was varied in fine steps from $0°C$ to $100°C$, while supply voltage was varied from 1V to 1.2V. The impact of supply voltage and temperature fluctuations on the intra-distance of both litho-aware and base PUF designs are shown in Figures 5(a) and 5(b), respectively. The results are shown for $V_{gate} = 0.5V_{DD}$. Similar results were observed for the other gate voltage (V_{gate}) as well.

We can observe that litho-aware current-based PUF performs better than the base PUF design under different operating conditions. As the current generation transistors have been upsized in our design, the impact of process variations becomes dominant over temperature and supply voltage fluctuations.

C. Unpredictability

Unpredictability is a measure of the security level of a PUF and determines the extent to which a PUF is unclonable. There are different methods to estimate the unpredictability of a PUF. One method is to estimate the tolerance of a PUF to modeling attacks, as majority of the PUFs have been broken using machine learning algorithms [10]. The other method is to estimate the amount of randomness in PUF responses using NIST tests [11]. Since the litho-aware design looks at increasing the sensitivity of the PUF circuit to process variations, it is interesting to look at the impact of process variations on the differential current ($\Delta I = I_a - I_b$), which determines the final response. If the extent of process variations is too

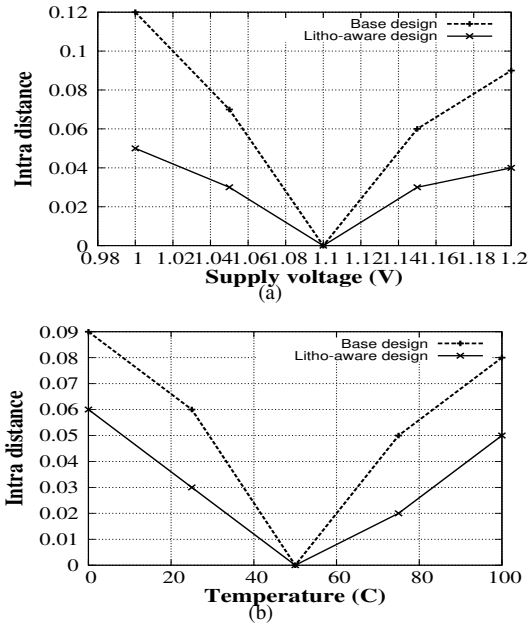

Fig. 5: Intra-distance evaluation under, (a) Supply voltage fluctuations and (b) Temperature fluctuations.

large, then the circuit might be biased towards either '0' or '1'. So, we conducted experiments on about 1000 litho-aware PUF instances and determined the impact of litho-aware design technique on the differential current. The differential current distribution is shown in Figure 6. It can be clearly observed that the differential current is often very small and also the mean of the distribution is around 0. This clearly shows that the litho-aware design technique introduces a negligible bias in the PUF design. Again, we show results only for $V_{gate} = 0.5V_{DD}$.

Modeling attacks: To estimate the tolerance of litho-aware current-based PUF to modeling attacks, we employed Support Vector Machines (SVM) based attack on PUF responses from the base and litho-aware PUF designs. The feature vectors used in SVM were derived as per the framework described by Lim et al. in [12]. However, the framework was modified to incorporate the operation of current-based PUF, as the stages contribute to currents I_a and I_b only if at-least one of the gate inputs of the SC module is set to logic '1'. SVM^{light} was employed for modeling attack purposes. Based on our attack strategy, we observed that both the PUF designs can be attacked using fewer than 5000 CRPs to achieve a prediction accuracy of around 95%. The number of CRPs goes down even further when leakage currents are not considered. Figure 7 shows the classification error obtained from SVM^{light} for our PUF responses. This shows that current-based PUF design cannot be used as such in secure system design. To improve modeling attack resistance, some sort of logic circuitry has to be employed which will mask the linear behavior of PUFs.

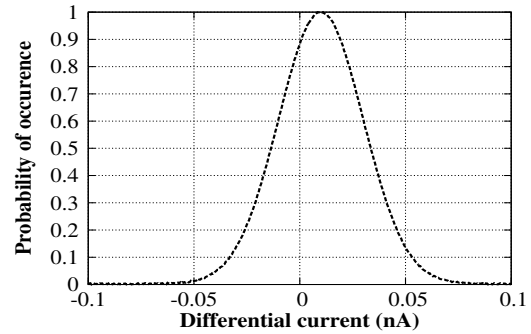

Fig. 6: Impact of process variations on the differential current

NIST Tests: The amount of randomness present in PUF responses was tested using NIST test suite. It is important to note that randomness is different from uniqueness. Uniqueness is computed across the responses from different PUF instances, whereas randomness is measured for a single PUF device. The results of the NIST tests are not shown in detail due to space limitations. However, the amount of randomness as measured using *ApproximateEntropy* test is shown in Table IV. For evaluation purposes, more than 100,000 response bits were generated from a single PUF device. In general, both litho-aware and base PUF had enough randomness and passed all the tests.

TABLE IV: Approximate Entropy results

Base design	Litho-aware design
0.96	0.98

D. Implementation Details

Area: The modified layout obtained from the lithographic simulation framework was used for area and power measurements. A 128-stage PUF has approximately 100 NAND gate equivalents (around 400 transistors). Since our design does not add any extra transistors, the number of NAND gate equivalents remains the same between the base and litho-aware designs. However, increasing the size of current generation transistors imposes a small area overhead. The area of the

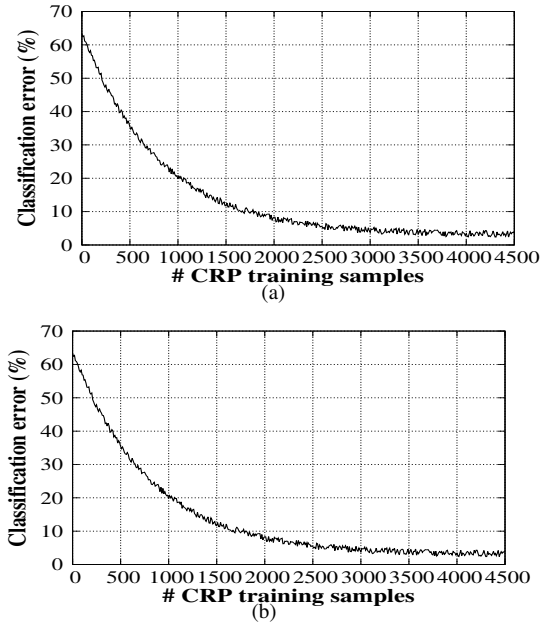

Fig. 7: Classification error of SVM based attack on current PUFs. (a) Base PUF and (b) Litho-aware PUF

die footprint required to fit in both the designs is shown in Table V. It can be observed that litho-aware design introduces an area overhead of approximately 3% when compared to base design. Despite a minimal area overhead, it is important to note that litho-aware current-based PUF fares better than other type of PUF implementations such as delay-based PUFs. For example, a 64-stage arbiter PUF required a die footprint area of $1800\mu m^2$ [13].

TABLE V: Die footprint area requirements

Base design	Litho-aware design	% increase in area
475 μm^2	490 μm^2	3

Energy consumption: To enable fair comparison, instead of computing power, we computed the energy consumption per response bit. Since, litho-aware design generates currents slightly higher than the base design, there is a minimal overhead on energy consumption. The time taken to generate a single bit of response still remains the same, as it is partially decided by the resolution time of sense amplifier. The current-based PUF approximately takes about 250ps to generate a single bit of response. The energy consumption details are shown in Table VI. It can be observed that the energy overhead imposed by litho-aware design is very minimal ($< 7\%$).

TABLE VI: Energy consumption details

Base design	Litho-aware design	% increase in Energy
0.032pJ	0.034pJ	6.25

When compared with other low power designs [13], the litho-aware design consumes lesser energy per response bit. For example, a 64-stage sub-threshold delay based arbiter PUF consumes about 0.047 pJ of energy. However, litho-aware design consumes only 0.034 pJ of energy, resulting in about 27% improvement.

VI. Conclusion

In this paper, we presented a current-based PUF designed using a lithographic simulation framework for improving the sensitivity of a PUF circuit to process variations. By effectively fracturing transistors to multiple segments and spacing them apart closer to forbidden pitch zone, the circuit can be made sensitive to process variations. The litho-aware PUF shows better inter-wafer uniqueness with respect to conventional current-based PUF, while achieving acceptable levels of reliability and unpredictability. Moreover, the area and power/energy overheads imposed by the litho-aware design are very minimal and substantially better than delay-based PUFs. Thus, the litho-aware design technique can be used for achieving better uniqueness while allowing a marginal increase in power consumption. Future work includes rigorous modeling of inter-wafer variations to analyze the impact on uniqueness of PUFs.

Acknowledgment

The authors would like to thank Dr. Sandip Kundu for providing constructive suggestions related to lithography. This work is supported by Semiconductor Research Corporation (SRC) task # 1836.074 through the Texas Analog Center of Excellence (TxACE) and NSF grant 0964379.

References

[1] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Proc. Design Automation Conference*, 2007, pp. 9–14.

[2] D. Kim, C. Cho, J. Kim, J.-O. Plouchart, R. Trzcinski, and D. Ahlgren, "Cmos mixed-signal circuit process variation sensitivity characterization for yield improvement," in *Proc. IEEE Custom Integrated Circuits Conference*, Sept. 2006, pp. 365 –368.

[3] S. Kundu, A. Sreedhar, and A. Sanyal, "Forbidden pitches in sub-wavelength lithography and their implications on design," *Journal of Computer-Aided Materials Design*, vol. 14, pp. 79–89, Apr. 2007.

[4] D. Forte and A. Srivastava, "On improving the uniqueness of silicon-based physically unclonable functions via optical proximity correction," in *Proc. Design Automation Conference*, June 2012, pp. 96 –105.

[5] A. Sreedhar and S. Kundu, "Physically unclonable functions for embeded security based on lithographic variation," in *Design, Automation Test in Europe Conference Exhibition (DATE)*, March 2011, pp. 1 –6.

[6] D. Forte and A. Srivastava, "Manipulating manufacturing variations for better silicon-based physically unclonable functions," in *Proc. IEEE Computer Society Annual Symposium on VLSI*, Aug. 2012, pp. 171 – 176.

[7] M. Majzoobi, G. Ghiaasi, F. Koushanfar, and S. Nassif, "Ultra-low power current-based puf," in *Proc. IEEE International Symposium on Circuits and Systems*, 2011, pp. 2071–2074.

[8] A. Sreedhar and S. Kundu, "On modeling impact of sub-wavelength lithography on transistors," in *Proc. International Conference on Computer Design*, Oct. 2007, pp. 84 –90.

[9] I. Kim, A. Maiti, L. Nazhandali, P. Schaumont, V. Vivekraja, and H. Zhang, "From statistics to circuits: Foundations for future physical unclonable functions," in *Towards Hardware-Intrinsic Security*, ser. Information Security and Cryptography, A.-R. Sadeghi and D. Naccache, Eds. Springer Berlin Heidelberg, 2010, pp. 55–78.

[10] U. Rührmair, F. Sehnke, J. Sölter, G. Dror, S. Devadas, and J. Schmidhuber, "Modeling attacks on physical unclonable functions," in *Proc. ACM conference on Computer and communications security*, 2010, pp. 237–249.

[11] A. Rukhin, J. Soto, J. Nechvatal, E. Barker, S. Leigh, M. Levenson, D. Banks, A. Heckert, J. Dray, S. Vo, A. Rukhin, J. Soto, M. Smid, S. Leigh, M. Vangel, A. Heckert, J. Dray, and L. E. B. III, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," 2001.

[12] D. Lim, "Extracting secret keys from integrated circuits," Master's thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, 2004.

[13] L. Lin, D. Holcomb, D. K. Krishnappa, P. Shabadi, and W. Burleson, "Low-power sub-threshold design of secure physical unclonable functions," in *Proc. International Symposium on Low Power Electronics Design*, Aug. 2010, pp. 43 –48.