An 8-Bit/Cycle 7.39×10^{31}-CRP/F^2 All-Digital Strong PUF Exploiting Thermal Noise for Improved Modeling Attack Resilience

Chongyao Xu[1], Jiale Li[1], Yaoqi Bao[1], Lichen Feng[2*], Man-Kay Law[3*] and Zhangming Zhu[1]
[1]Zhejiang Key Laboratory of Analog Integrated Circuits, Hangzhou Institute of Technology, Xidian University, Hangzhou 311231, China; [2]Key Laboratory of Analog Integrated Circuits and Systems, Ministry of Education, Xidian University, Xi'an 710071, China; [3]State Key Laboratory of Analog and Mixed-Signal VLSI, FST-ECE, University of Macau, Macau 999078, China. *Corresponding Author.

***Abstract*—This paper presents a light-weight high-throughput all-digital strong physical unclonable function (PUF) that breaks the learnability of strong PUFs by obfuscating responses using thermal noise. Exploiting both static process variations and dynamic thermal noise, the proposed design generates challenge-dependent, time-varying responses, which enhances security without complex nonlinear logic and enables single-cycle eight-bit response output. The entropy source consists of two identical delay blocks that are automatically placed and routed by EDA tools using standard digital cells to reduce the design complexity. The proposed spatial-redundancy scheme is capable of decreasing the worst-case bit error rate (BER) from 0.103 to 3.3×10^{-8}, without CRP loss. Experimental results show that the prediction accuracy of modeling attacks remains at ~50% even using 50M training CRPs. The achieved throughput is 160 Mbps (320 Mbps scalable) with only 4.6 MF^2 area. The area efficiency is up to 7.39×10^{31} CRP/F^2, over 7×10^6 times higher than prior arts.**

***Index Terms*—Strong PUF, Modeling Attack, Hardware Security, Light Weight, High Throughput**

I. Introduction

Strong Physical Unclonable Functions (SPUFs) offer a large number of Challenge-Response Pairs (CRPs) and are promising for both low-cost device authentication and high-security key generation. However, their vulnerability to machine learning (ML)-based modeling attacks remains a critical security concern. While recent works have enhanced resilience by employing large arrays of weak PUF cells combined with XOR architectures, these approaches often incur significant trade-offs: reduced throughput ($\leq$25Mbps [1]–[4]) or increased hardware overhead ($\geq$20MF^2 [4]–[6]). More fundamentally, these schemes rely on static responses—i.e. fixed challenge-response mappings—which inherently admit a learnable separating hyperplane (linear or nonlinear), posing a persistent risk of successful prediction.

In this work, we propose a noise-obfuscated (NOB) SPUF architecture. Unlike traditional designs that enhance the security via XOR operations, the proposed architecture leverages both static process variations and dynamic thermal noise of the entropy source to generate challenge-dependent, time-varying responses, as illustrated in Fig. 1(a). Due to the inherent randomness and unpredictability of noise, this architecture enables the generation of multi-bit responses within a single cycle and enhances security at the same time without complex nonlinear operations and greatly increases the area efficiency. As shown in Fig. 1(b), the proposed architecture achieves 8 response bits per cycle, representing a >116.2% improvement over previous work. Regarding security, this design maintains a prediction accuracy of approximately 50% even when subjected to state-of-the-art ML attacks utilizing a 50M training set. With a 128-bit challenge, it supports up to 3.4×10^{38} CRPs while occupying only 4.6 MF^2, yielding an area efficiency of 7.39×10^{31} CRP/F^2, which is 7.04×10^6 times higher than prior works. Furthermore, to meet the stringent reliability requirements of key generation, we introduce a spatial-redundancy-based key generation method, which enables error-free key regeneration (BER=0) under nominal operating conditions.

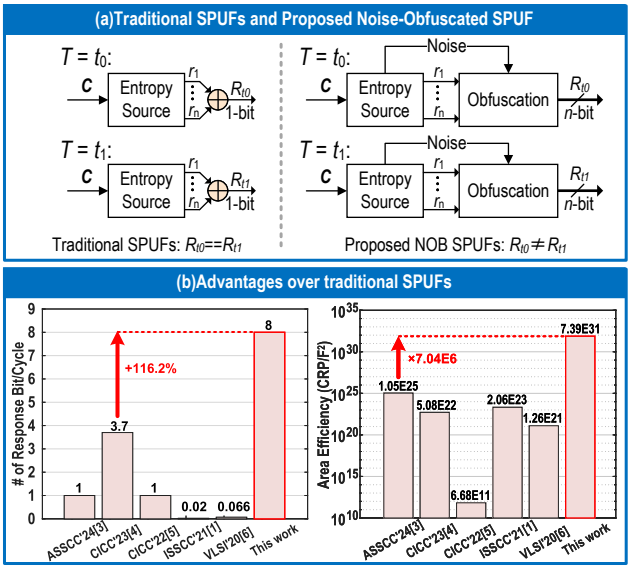


Fig. 1. (a) The theory of the proposed SPUF, (b) The comparisons with prior works.

This work is supported by the Fundamental Research Funds for the Central Universities under Grant QTZX26156.

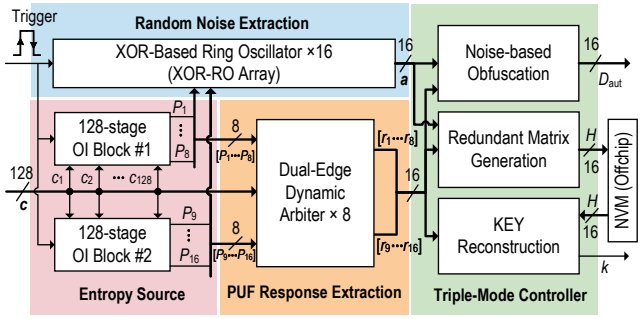


Fig. 2. Architecture of the proposed Triple-mode NOB SPUF.

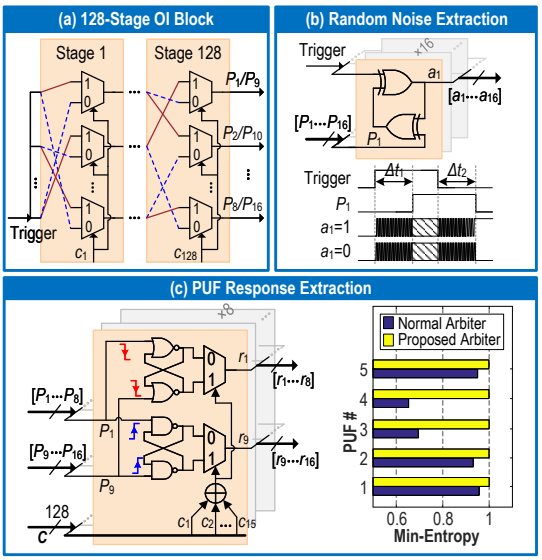


Fig. 3. The structure of (a) OI block, (b) RO-based random number generator and (c) dual-edge dynamic arbiter.

II. The Proposed Triple-Mode NOB SPUF

Fig. 2 illustrates the architecture of the proposed triple-mode NOB SPUF, which employs arbiters and XOR-based ring oscillators (ROs) to extract static and dynamic entropy, respectively, from two identical 128-stage obfuscated interconnection (OI) delay blocks. As shown in Fig. 3(a), each stage comprises eight 2-to-1 multiplexers. Inter-stage wiring is randomly permuted to enhance unpredictability. All 128 stages form eight distinct propagation paths $(P_1, \cdots, P_8)/(P_9, \cdots, P_{16})$. Unlike prior PUF designs [3], [7], the entropy source in this work does not require symmetry constraints or manual placement and routing (P&R), enabling full automation via standard EDA tools and significantly reducing design complexity.

As shown in Fig. 3(b), the noise extraction circuit comprises 16 XOR-based ROs, whose operation is controlled by the $trigger$ and $(P_1, \cdots, P_{16})$ signals. As observed in the waveform, oscillation begins upon receiving the $trigger$ signal and ceases when the signal output from the OI block. We assume that the jitter at each logic gate (such as the multiplexer and XOR gate) follow $\mathcal{N}(0, \sigma^2)$. The ideal propagation delay of each gate is modeled as a constant μ. Consequently,

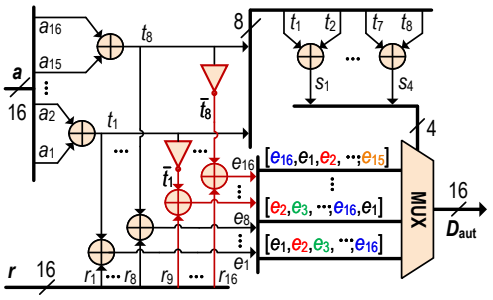


Fig. 4. The internal structure of noise-based obfuscation circuit.

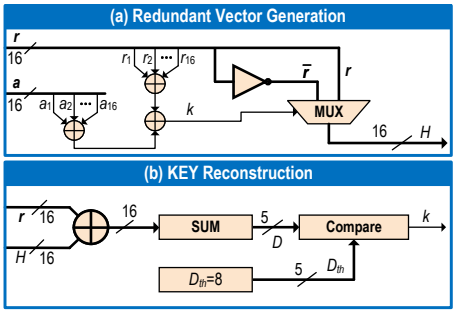


Fig. 5. The theory of (a) redundant vector generation and (b) key reconstruction.

the total propagation delay Δt_1 and Δt_2 in Fig. 3(b) follow $\mathcal{N}(n\mu, n\sigma^2)$ with $n = 128$ in this work. Similarly, the oscillation period T_{RO} follows $\mathcal{N}(2\mu, 2\sigma^2)$. Therefore, $a_i, 0 \le i \le 16$, is noise-dependent and $\text{Prob}(a_i = 1) = \text{Prob}(\text{mod}(\lfloor \frac{\Delta t_1 + \Delta t_2}{T_{RO}} \rfloor, 2) = 1)$.

Fig 3(c) shows the double-edge dynamic-select arbiters, which extract the process variation from both raising and falling edges for doubling the throughput without increasing extra entropy source. The response is dynamically selected from one of the two latch outputs according to the input challenge, effectively improving the randomness of the PUF responses according to our simulation.

Fig. 4 shows the procedure of generating the NOB responses, $D_{aut} = ((t_1, \cdots, t_8, \bar{t}_1, \cdots, \bar{t}_8) \oplus \boldsymbol{r}) \ggg \boldsymbol{s}$ with $t_i = a_{2i-1} \oplus a_{2i}$, $s_i = t_{2i-1} \oplus t_{2i}$. Due to the randomness of $\boldsymbol{a} = (a_1, \cdots, a_{16})$, D_{aut} is inherently dynamic and time-variant, enabling secure parallel output of all 16 response bits, which yields a significant improvement in throughput and area compared to the XOR compression methods, such as [1], [3], [5], [6].

For robust key generation, inspired by the majority correction [8], we introduce a spatial-redundancy-based scheme: a single key bit $k \in \{0, 1\}$ is encoded using a 16-bit static PUF response $\boldsymbol{r}$. As shown in Fig. 5(a), 1-bit key, k, is first generated by XOR-compressing the 16-bit $\boldsymbol{r}$ and $\boldsymbol{a}$, enhancing both randomness and security. The redundancy vector $\boldsymbol{H}$ is equal to $\boldsymbol{r}$ or its complement $\overline{\boldsymbol{r}}$, depending on the value of k. During key reconstruction, the original k is recovered by computing the Hamming distance (HD) between the measured response $\boldsymbol{r}$ and $\boldsymbol{H}$. A decoding error occurs only if more

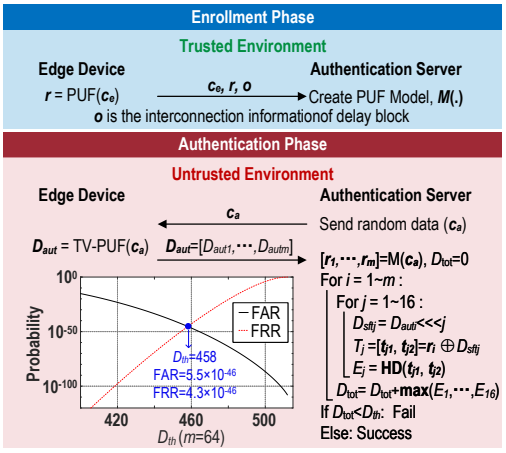


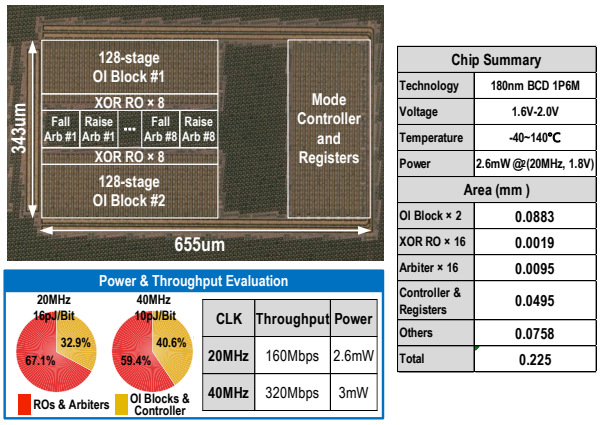
Fig. 6. The flow diagram of NOB-responses-based authentication, and the FAR and FRR under different threshold (D_{th}), m is the number CRPs used for authentication.

than half of the 16 response bits flip. In contrast to [8], the proposed design eliminates the requirement for a weak PUF array and enables dynamic key renewal by issuing new $\boldsymbol{H}$, resulting in lower overhead and greater security. Note that $\boldsymbol{H}$ is employed exclusively for reconstructing the root key of a system; consequently, its associated memory footprint remains minimal.

III. NOB-Response-Based Authentication

Fig. 6 shows the NOB-response-based authentication procedure. In the enrollment phase, the CRP set $\{\boldsymbol{c_e}, \boldsymbol{r}\}$ and interconnection information ($\boldsymbol{o}$) are employed to construct an accurate PUF model on server [9]. During the authentication phase, the server sends challenges to the edge device and receives the corresponding D_{aut} arrays. The server then rotates each D_{aut} 16 times to reconstruct all possible shift positions, and each rotated version is bitwise XORed with the model-generated $\boldsymbol{r_i}$, producing 16 candidate vectors $(T_1, \cdots, T_j, \cdots, T_{16})$, $T_j = (t_{j1}, t_{j2})$. The correct vector satisfies the complementary-pair condition, i.e. $E_j = \mathrm{HD}(t_{j1}, t_{j2}) = 8$. To improve the robustness of the scheme, a tolerance threshold D_{th} is introduced to accommodate near-matching responses, and multiple CRPs should be employed in each authentication session.

Assuming that each authentication section utilizes m CRPs, the false acceptance rate (FAR) and false rejection rate (FRR) are given by $\mathrm{FAR} = \sum_{i=D_{th}}^{8m} \binom{8m}{i} \left(p_u^2 + (1-p_u)^2\right)^i \left(2p_u(1-p_u)\right)^{8m-i}$, and $\mathrm{FRR} = \sum_{i=8m-D_{th}+1}^{8m} \binom{8m}{i} \left(p_r^2 + (1-p_r)^2\right)^{8m-i} \left(2p_r(1-p_r)\right)^i$, respectively. Here, p_u and p_r denote the inter-die and intra-die Hamming distances (HDs) of the static PUF responses, respectively. In this work, p_u=26.21% and p_r=0.35%. Therefore, for an authentication session employing 64 CRPs, we obtain $\mathrm{FAR} = 5.5 \times 10^{-46}$ and $\mathrm{FRR} = 4.3 \times 10^{-46}$ with a decision threshold $D_{th} = 458$. In comparison, a traditional arbiter PUF [7] requires at least 10× CRPs to achieve similar performance.



Chip Summary	
Technology	180nm BCD 1P6M
Voltage	1.6V-2.0V
Temperature	-40~140℃
Power	2.6mW @(20MHz, 1.8V)
Area (mm)	
OI Block × 2	0.0883
XOR RO × 16	0.0019
Arbiter × 16	0.0095
Controller & Registers	0.0495
Others	0.0758
Total	0.225

Fig. 7. Die photograph, area breakdown, and power-efficiency evaluation of the fabricated 180 nm SPUF chip.

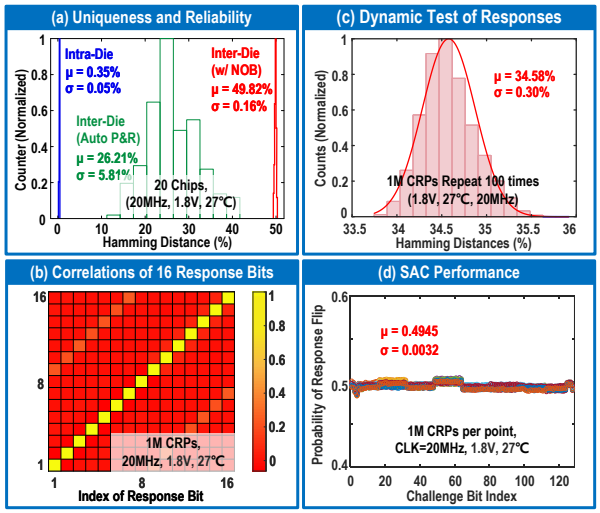


Fig. 8. Measured results of (a) Intra/Inter-Die HDs, (b) cross-correlation of 16 response bits, (c) dynamic performance and (d) SAC performance.

IV. Experimental Results

Fig. 7 shows the proposed triple-mode NOB SPUF fabricated in 180nm CMOS process with two OI blocks automatically P&R by the EDA tools using standard digital cells.

Fig. 8(a) shows a measured average intra-die HD of 0.35% across 20 chips under nominal condition. Due to the limited process variation of standard digital cells, the uniqueness of PUF is 26.21%—higher than the 23% of traditional APUF [7], and further increases to 49.82% in authentication mode due to the incorporation of random vector $\boldsymbol{a}$. As observed in Fig. 8(b), the cross-correlation among the 16 NOB responses is minimal, with 93% of the coefficients falling within the range [-0.1, 0.1], confirming a weak statistical dependence even under identical challenges. The dynamic behavior of D_{aut} was evaluated by applying the same 1M challenges repeatedly over 100 trials. On average, 34.58% of response bits changed across any two iterations, shown in Fig. 8(c), demonstrating strong response volatility. As shown in Fig. 8(d), the strict avalanche criterion (SAC) is satisfied with ∼50% NOB response bits toggling per

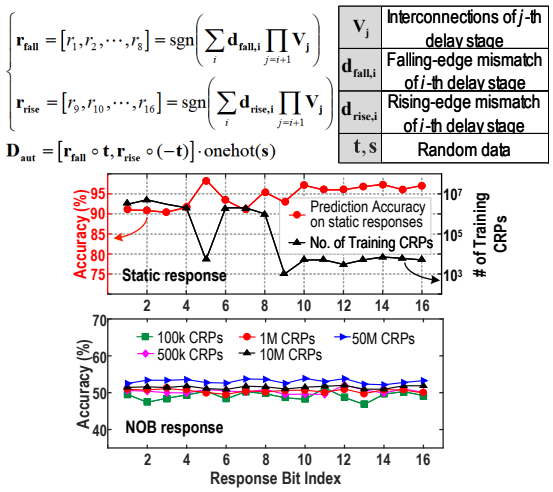
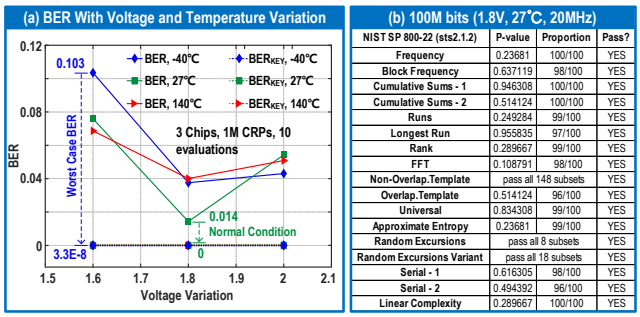


Fig. 9. The security model of this work and the ML attack results of static PUF responses and NOB PUF responses.



(b) 100M bits (1.8V, 27°C, 20MHz)

NIST SP 800-22 (sts2.1.2)	P-value	Proportion	Pass?
Frequency	0.23681	100/100	YES
Block Frequency	0.637119	98/100	YES
Cumulative Sums - 1	0.946308	100/100	YES
Cumulative Sums - 2	0.514124	100/100	YES
Runs	0.249284	99/100	YES
Longest Run	0.955835	97/100	YES
Rank	0.289667	99/100	YES
FFT	0.108791	98/100	YES
Non-Overlap.Template	pass all 148 subsets		YES
Overlap.Template	0.514124	96/100	YES
Universal	0.834308	99/100	YES
Approximate Entropy	0.23681	99/100	YES
Random Excursions	pass all 8 subsets		YES
Random Excursions Variant	pass all 18 subsets		YES
Serial - 1	0.616305	98/100	YES
Serial - 2	0.494392	96/100	YES
Linear Complexity	0.289667	100/100	YES

Fig. 10. Measured results of (a) bit error rate of static responses and keys and (b) 15 NIST statistical tests.

challenge bit flip, indicating excellent resistance to statistical analysis.

Using the pypuf library—a powerful artificial neural network (ANN) framework shown to effectively break the arbiter-PUF variants [10]—we trained models using millions of CRPs. As shown in Fig. 9, the prediction accuracy exceeded 90% for all 16 static response bits using <10M training CRPs. In contrast, the proposed NOB response D_{aut} exhibits ∼50% prediction accuracy even with 50M CRPs as training set. The rationale of security on resisting ML attacks is also provided in the equation of Fig. 9.

Fig. 10(a) presents BER under voltage $(1.6 - 2.0V)$ and temperature $(-40°\mathrm{C}$ to $85°\mathrm{C})$ variations. The native BER is 0.014 under nominal conditions. With spatial redundancy, the BER of key drops to 0 across three chips, 1M CRPs, and 10 evaluations. Even under worst-case conditions $(-40°\mathrm{C}$, 1.6 V), $\mathrm{BER_{KEY}}$ is as low as $3.3{\times}10^{-8}$. With respect to randomness, as shown in Fig. 10(b), 100 million consecutively key bits pass all 15 NIST SP 800-22 statistical randomness tests.

Table I compares this work with prior arts. Our design achieves 3.4×10^{38} usable CRPs, a typical throughput of 160 Mbps (scalable to 320 Mbps at 40 MHz), and resilience

TABLE I
OVERALL COMPARISON

		This work	ASSCC'24[3]	CICC'23[4]	CICC'22[5]	ISSCC'21[1]	VLSI'20[6]
Technology		180nm	65nm	130nm	65nm	130nm	14nm
Usable CRPs		3.40E+38	4.00E+31	1.27E+30	1.84E+19	1.27E+30	3.00E+28
Nonlinear Type		OI+Random	OI+XOR Network	Feedback SPN	XOR Network	Feedback SPN	AES Sbox
Area (MF^2)		4.6	3.8	25	27.55	6.17	23.8
KEY Generation							
Stability	Worst BER	<3.3E-8	<4E-6	<2E-8	<4E-8	0.73%	0.26%
	Temp. (°C)	-40~140	-20~100	-20~125	-40~125	0~120	0-100
	Voltage Variation (V)	1.6-2.0	1.08-1.32	0.75-1.00	0.8-1.4	1.4-1.8	0.65-0.85 (±13%)
	CRP Loss	No	Yes (99%)	No	No	No	Yes (99%)
Stabilization Method		Spatial Redundancy	Metastability Detection & Masking	Hot Carrier Injection	ASCH	Hot Carrier Injection	Masking
Authentication							
Modeling Attack	Accuracy	~50%	~50%	~50%	~50%	~50%	~50%
	Training CRPs	50M	20M	10M	10M	20M	6M
Dynamic Response		Yes	No	No	No	No	No
Throughput (Mbps)	Bit/Cycle	8	1	3.7	1	0.02	0.066
	Mbps	160@1.8V	25@1.2V	17.64@0.85V	150@1.2V	0.02@0.5V	106@0.75V
Energy Efficiency (pJ/bit)		16@1.8V	4.54@1.2V	4.67@0.85V	4.28@1.2V	54.9@0.5V	97@0.75V
Area Efficiency1 (CRP/F^2)		7.39E+31	1.05E+25	5.08E+22	6.68E+11	2.06E+23	1.26E+21
Area Efficiency2 (bps/F^2)		34.78	6.58	0.71	5.44	0.0032	4.45

against ML attacks with up to 50M training CRPs while only consuming $4.6MF^2$ area. The area efficiency reaches 7.39×10^{31} CRP/F^2 and 34.78 bps/F^2, obviously exceeding the previous work.

REFERENCES

[1] K. Liu, Z. Fu, G. Li, H. Pu, Z. Guan, X. Wang, X. Chen, and H. Shinohara, "36.3 A Modeling Attack Resilient Strong PUF with Feedback-SPN Structure Having <0.73% Bit Error Rate Through In-Cell Hot-Carrier Injection Burn-In," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, vol. 64, San Francisco, CA, USA, 2021, pp. 502–504.

[2] K. Liu, G. Li, Z. Fu, X. Wang, and H. Shinohara, "A 2.17-pJ/b 5b-Response Attack-Resistant Strong PUF with Enhanced Statistical Performance," in *ESSCIRC 2022- IEEE 48th European Solid State Circuits Conference (ESSCIRC)*, 2022, pp. 513–516.

[3] L. Zhang, C. Xu, and M. K. Law, "Compact Hybrid Strong SOIPUF Exploiting Secret Key Generation and Dual-Edge Response Extraction for Enhanced Modeling Attack Resistance," in *2024 IEEE Asian Solid-State Circuits Conference (A-SSCC)*, 2024, pp. 1–3.

[4] K. Liu, Y. Tang, S. Xu, R. Zhang, and H. Shinohara, "A 100-Bit-Output Modeling Attack-Resistant SPN Strong PUF with Uniform and High-Randomness Response," in *2023 IEEE Custom Integrated Circuits Conference (CICC)*, 2023, pp. 1–2.

[5] Y. He, Q. Yu, and K. Yang, "A lossless and modeling attack-resistant strong puf with <4e-8 bit error rate," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Newport Beach, CA, USA, Apr. 24-27 2022, pp. 1–2.

[6] V. Suresh, R. Kumar, M. Anders, H. Kaul, V. De, and S. Mathew, "A 0.26% BER, 10^{28} challenge-response machine-learning resistant strong-PUF in 14nm CMOS featuring stability-aware adversarial challenge selection," in *Proc. IEEE Symp. VLSI Circuits*. Honolulu, HI, USA: IEEE, Jun. 2020, pp. 1–2.

[7] D. Lim, J. W. Lee, B. Gassend, G. E. Suh, M. Van Dijk, and S. Devadas, "Extracting secret keys from integrated circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 10, pp. 1200–1205, Dec. 2005.

[8] B. Karpinskyy, Y. Lee, Y. Choi, Y. Kim, M. Noh, and S. Lee, "8.7 physically unclonable function for secure key generation with a key error rate of 2e-38 in 45nm smart-card chips," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, 2016, pp. 158–160.

[9] C. Xu, L. Zhang, M.-K. Law, X. Zhao, P.-I. Mak, and R. P. Martins, "Modeling Attack Resistant Strong PUF Exploiting Stagewise Obfuscated Interconnections With Improved Reliability," *IEEE Internet Things J.*, vol. 10, no. 18, pp. 16 300–16 315, 2023.

[10] N. Wisiol, B. Thapaliya, K. T. Mursi, J.-P. Seifert, and Y. Zhuang, "Neural Network Modeling Attacks on Arbiter-PUF-Based Designs," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 2719–2731, Aug. 2022.