A Highly Reliable SRAM-Based PUF Retaining Memory Function

***Abstract*—In this paper, a novel highly reliable SRAM based Physical Unclonable Function (PUF) is proposed, which retains the memory function. The mismatch of NMOS is extracted during discharge process and amplified by the cross-coupled inverter to generate a response. At the beginning of discharge process, the NMOSs are biased at sub-threshold region, which can improve the reliability and stability. The proposed PUF is designed in a 40nm CMOS process and each bit cell only consumes 4.98 μm^2 (3112F^2). Post simulation shows that the bit error rate (BER) deterioration is 0.96$\%$ per 0.1V, 0.36$\%$ per 10°C with temperature changes from -40°C to 80°C temperature and supply voltage changes from 0.9V to 1.3V. It achieves 1.8$\%$ native instability through the simulation. Meanwhile, the proposed PUF can retains memory function after response generated.**

***Index Terms*—Physical unclonable function (PUF), SRAM, high reliability, subthreshold.**

I. Introduction

With the rapid development of the Internet of Things (IoTs), hardware security has become a critical issue. Traditional hardware encryption technology stores the key in NonVolatile Memory (NVM), which is vulnerable to attack such as side channel and reverse engineering [1]. As a hardware security primitive, physical unclonable function (PUF) can extract the random process variations during manufacturing process as a unique identification, which likes a fingerprint. Due to the manufacturing variations, it is to be measured and the PUF identification would disappear when it powers down. Therefore, PUF can be used as a emerging solution to protect important information from hackers [2].

However, the responses of PUF are sensitive to environmental noise, temperature variation and supply voltage, which limits PUF using as a key generator. Besides, with the increase of PUF using time, the aging effect leads to output data overturning. As a result, the bit error rate (BER) increases [3]. High BER needs complicated error correction module to correct the error, which consumes much hardware resources [4]. This is not desired for a low-cost PUF design.

SRAM PUF can generate the responses without additional circuit. However, the traditional SRAM PUF shows nearly 30$\%$ instability [5]. In order to reduce the area consumption, a novel SRAM PUF with two bits per PUF cell is proposed in [6], but the native instability is still high, which is 17$\%$. In [7], EE SRAM PUF is proposed to magnify the mismatch, thus greatly reducing the BER. However, the power consumption of EE SRAM PUF is high because the transistor works at the saturation region, resulting in large static power. A compact PUF is presented in [8], which can improve the reliability against to temperature variation, but the voltage reliability is not improved.

In order to reduce the BER and power consumption of SRAM PUF, this paper utilizes the difference of subthreshold discharge current to extract circuit mismatch. Specifically, the main improvements of this work are shown as follows:

- A new PUF based on SRAM with high stability and reliability is proposed. The stability is improved by starting up with an elaborately designed enabling signal instead of the supply voltage.
- The power consumption of the proposed SRAM PUF is decreased by controling the enabling signals sequence, which can reduce the current effectively.
- The memory function is retained by configuring the enabling signals to a specific state.

The rest of this paper is organized as follows: Section II analyzes characteristic of subthreshold discharge current; Section III presents the proposed SRAM PUF and the layout of overall PUF architecture. The results and comparison of proposed PUF is presented in Section IV. The conclusion is given in Section V.

II. Basis of Sub-threshold Current

When the MOS transistor works in the saturation and sub threshold region, the current are expressed as follows:

$$I_{sat} = \mu C_{ox}(\frac{W}{L})(V_{gs} - V_{th})^2 \tag{1}$$

$$I_{sub} = \mu C_{ox}(\frac{W}{L})\exp(\frac{V_{gs} - V_{th}}{V_t})(1 - \exp(-\frac{V_{ds}}{V_t})) \tag{2}$$

where μ is the carrier mobility of NMOS, C_{ox} is the sheet oxide capacitance density, W/L is the width to length ratio of NMOS transistor, V_{th} is the threshold voltage,and V_t is the thermal voltage. I_{sat} represents the current in the saturation region, while I_{sub} means the current at the subthreshold. Since V_{ds} starts discharging from V_{dd}, V_{ds} is much greater than V_t, thus $\exp(-V_{ds}/V_t)$ be ignored.

Fig. 1 shows the circuit schematic of current mirror. The gate voltage V_{gs} controls the current passing through the transistor. When the switch S turns on, V_{ds} would be charged to V_{dd}. When the switch S turns off, NMOS starts to discharge and the V_{ds} decreases continuously. According to the discharge principle of RC circuit and the mismatch between M1 and M2, the voltage of V_{ds1} and V_{ds2} would be different over discharge time, as given by

$$\frac{V_{ds1}}{V_{ds2}} = \exp(-\frac{t}{C}(\frac{I_1}{V_{ds1}} - \frac{I_2}{V_{ds2}})) \quad (3)$$

where C is the capacitance of the circuit load connected to the drain of M1 or M2. Since the value is slightly affected by the mismatch between M1 and M2, it can be assumed that the load capacitance values are equal. I_1 and I_2 represent the discharge current over M1 and M2, respectively. When V_{gs} is different, the transistor will work in the saturation region or sub-threshold region. The discharge current will also change.

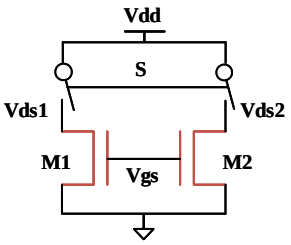


Fig. 1. Schematic of current mirror.

It can be noted in (1) that the sub-threshold current is exponentially related to the threshold voltage. Thus its value is more sensitive to the change of the threshold voltage compared with saturation value in (2). From (3), it can be seen that the value of V_{ds1}/V_{ds2} in sub-threshold region has more significant difference than it in saturation region with the same mismatch. Therefore, the stability of PUF circuit can be improved. Furthermore, the discharge current in subthreshold is very small, resulting in low power consumption.

Meanwhile, (3) shows that the ratio of V_{ds1}/V_{ds2} is not affected by the supply voltage, which can improve voltage reliability.

III. The Proposed PUF

A. PUF Cell Based on SRAM

The schematic of the proposed PUF cell is shown in Fig. 2. The circuit consists of a traditional SRAM cell and several switch transistors. The proposed PUF cell is controlled by EN and ENB. In the initial state, EN set to 0 and ENB set to 1. The gate voltages of M3 and M4 are pulled up to V_{dd}, and the gate voltages of M1 and M2 are pulled down to the ground. In this case, M7 and M8 are turned off to eliminate the static current. Then, the voltages V_a and V_b are discharged from V_{dd} through M1 and M2 while EN is set to 1 and ENB changes inversely. The values of V_{gs1} and V_{gs2} are increased from 0 in the process. Because of the voltage stabilizing effect of M7 and M8, V_{gs1} and V_{gs2} are always smaller than the threshold voltage, which keeps the whole discharge process in the sub-threshold region. The mismatch between M1 and M2 leads to the difference of V_a and V_b, and it can be amplified by the cross coupling inverters. It is worth nothing that the proposed PUF cell can work as a traditional SRAM when EN remains 1 and ENB is 0.

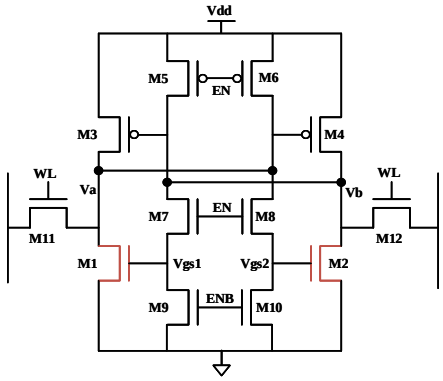


Fig. 2. Schematic of the proposed PUF cell.

The transient simulation waveforms of proposed PUF cell are shown in the Fig. 3. It can be found from the figure that when EN changes from 0 to 1, the voltage values of V_{gs1} and V_{gs2} increase gradually. However, due to the existence of M7 and M8, the voltage remains coincident. At the same time, the voltages of V_a and V_b decrease gradually due to the discharge. Furthermore, the voltage values of V_{gs1} and V_{gs2} are smaller than the threshold voltage, which keeps M1 and M2 working in the subthreshold region. The difference between V_a and V_b in the discharge process expands gradually. M1, M2, M3 and M4 amplify V_a and V_b to 0 or 1, respectively. Then the response can be read out through the word line, which controlled by WL.

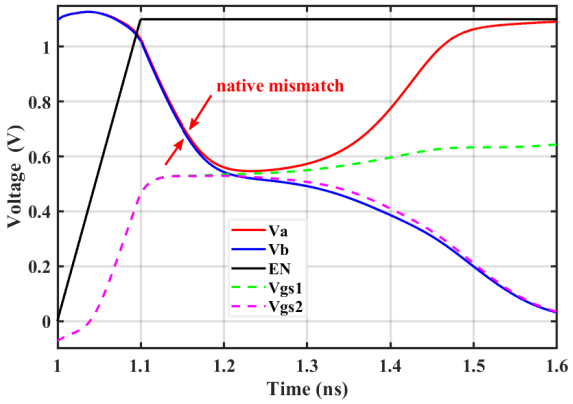


Fig. 3. Transient simulation waveforms of PUF cell.

B. Time Strategy

As shown in Fig. 2, the following situation will occur if ENB is enabled after EN. The conduction state of M9 and M10 before ENB is enabled causes the voltages of V_a and V_b to drop and makes M3 and M4 in the saturation region, which generates a DC path from the supply voltage to the ground. After ENB is enabled, M9 and M10 turns off, and V_a and V_b are charged gradually through M3 and M4. However, the

high energy consumption situation is undesired and is not the sub-threshold discharge state we need.

In order to reduce the power consumption and stabilize the output, a time strategy schematic of controling enable signal is proposed. As shown in the Fig. 4. The schematic consists of an array of inverter pairs. Each pair of inverters can ensure the correct sequence of signals and control 2×4 PUF cells. A total of 8×8 PUF cells can be controlled.

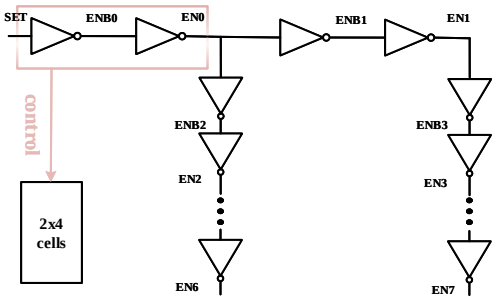


Fig. 4. Schematic of the enable module controling 8×8 cells.

C. Overall Architecture and Layout

The proposed PUF employs a array architecture, as shown in the Fig. 5(a), which is similar as SRAM. The decoder can choose the position of the PUF cell through the controling word line. In order to reduce the area consumption, hierarchical decoding is adopted. When reading out the PUF enabled response, the circuit should be precharged, which is controlled by the signal PC. At the end of the precharge, the word line WL selects the PUF cell output, which is amplified by the amplifier. The enable module provides correct timing control for the whole circuit.

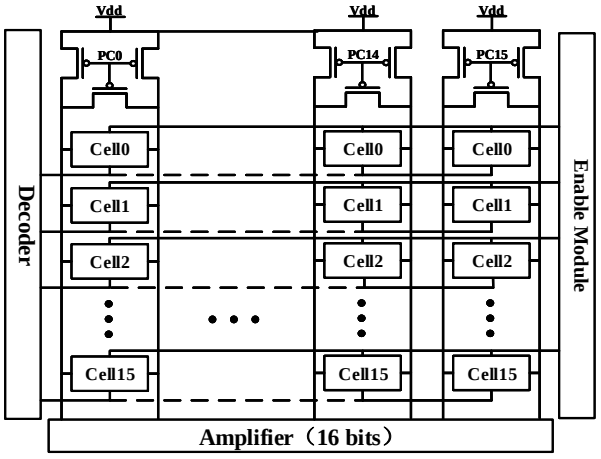
The proposed implementation was designed using a standard 40 nm CMOS process. Fig. 5(b) shows both the layout of enable module and PUF core, which features a compact overall silicon area of 1,540μm^2. Each PUF cell occupies an area of 4.98μm^2. The peripheral circuitries such as the decoders, SA, and the per-charge circuit were accomplished by schematic, which did not affected the results of responses.

IV. Simulation Results and Comparison

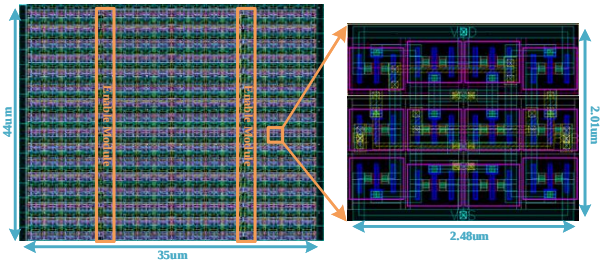
In order to better evaluate the performance of the proposed PUF, the PUF layout Monte Carlo simulation results is exhibited. The uniqueness, stability and reliability of the proposed PUF are analyzed by post-simulation. Finally, the advantages of the proposed PUF are discussed.

A. Uniqueness

Uniqueness can be expressed by hamming distance between chips. The inter-chip Hamming distance (Inter-HD) is expressed as the HD of responses from different PUF chips in the same environment. In order to evaluate the uniqueness of PUF, we simulated 50 times of Monte Carlo simulation on the



(a)



(b)

Fig. 5. (a) Overall archiecture of the proposed PUF; (b) Layout of the proposed PUF core.

proposed PUF layout at 27°C and 1.1V. The result is shown in Fig. 6 that Inter-HD exhibits a Gaussian distribution with μ=0.49 and σ=0.0676.

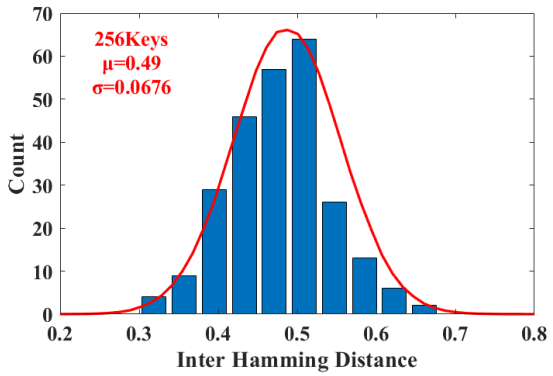


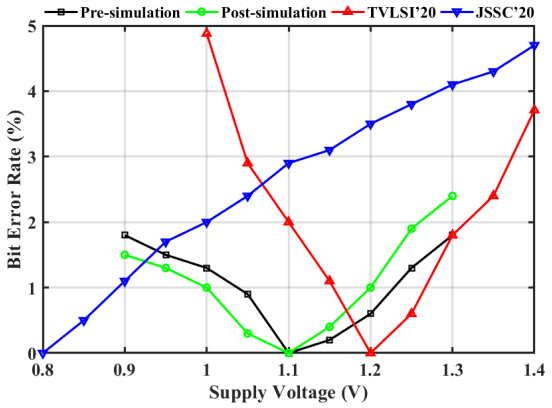
Fig. 6. Layout simulation of inter-HDs across 256 keys.

B. Stability

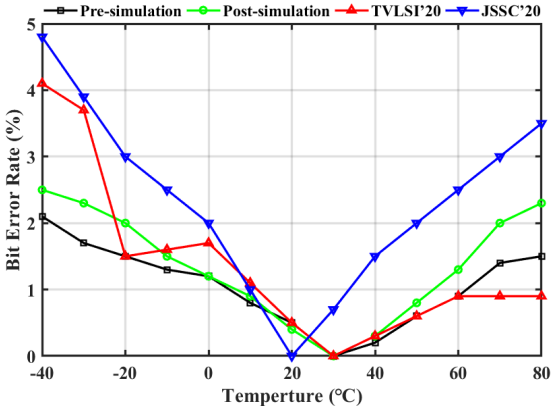
Due to the influence of environmental factors such as noise, PUF responses cannot remain constant under multiple measurements even if the same chip. Stability can be expressed by BER. In order to better simulate the actual PUF chip environment, we performed 2,000 noise simulations at a noise frequency between 100KHz to 500MHz. The results show that the native BER of the pre-simulation is 0.8%, and the post-simulation is 1.8%. After traditional temporal majority voting (TMV) error correction, the post-simitation instability is reduced to 0.3%.

C. Reliability

The proposed PUF's reliability is affected by environmental temperature and supply voltage variations. We simulated the temperature reliability of -40°C∼80°C and the voltage reliability of 0.9V∼1.3V. Furthermore, we compared with the results of native BER in [7] and [8], as shown in the Fig. 7. The results show that the proposed PUF has obvious advantages in voltage and temperature reliability. For per 0.1V voltage change, the average BER deteriorates by 0.98%, and the temperature changes per 10°C, the BER increases by 0.36%.



(a)



(b)

Fig. 7. (a) Voltage relibility between 0.8V to 1.3V; (b) Temperature relibility between -40°C to 80°C.

TABLE I
COMPARISON WITH THE STATE-OF-THE-ART WEAK PUF

	This Work	[7]	[8]	[9]	[10]
Technology	40nm	130nm	65nm	22nm	14nm
Structure	SRAM	EE SRAM	Cross-coupled	SRAM	SRAM
bitcell size(F^2)	3112	373	1036	9628	9387
Unstable Bit(%)	**1.8**	2.14	2.64	30	26.38
Temp. Range(°C)	-40∼80	-40∼120	-50∼150	N/A	N/A
Volt. Range(V)	0.9∼1.3	0.8∼1.4	1.0∼1.4	N/A	N/A
BER per 10°C(%)	**0.36**	0.75	0.26	N/A	N/A
BER per 0.1V(%)	**0.98**	0.78	2.34	N/A	N/A
Uniqueness	49	∼ 0.5	49.53	50.001	48.6
Core Energy(fJ/bit)	**13.25**	128	2980	13	4
Memory Function	**Yes**	No	No	No	No

D. Comparison

Table I provides a comprehensive comparison of our proposed PUF against the reported previous works. It shows that the stability of the proposed PUF is the best compared to [7], [8], [9], [10]. The temperature reliability of the proposed PUF is not much different from the optimal value of [8]. Meanwhile, the voltage reliability is improved by more than 50%. Compared with [7], the temperature reliability has obvious advantages, while the voltage reliability has not been reduced much. It is particularly pointed out that the power consumption of [7] is measured at 0.8V supply voltage, and other comparative works is measured at the standard voltage of the corresponding process. On the other hand, considering the characteristics of less energy consumption in the more advanced process, the energy consumption of our proposed PUF is the smallest, which is a very competitive circuit. One of the most important contributions of this work is that our improved SRAM-based PUF still has memory function, so our PUF can not only be used as a key generator, but also can be used with traditional SRAM as a cache. Furthermore, the proposed PUF area is compact comparing with [9], [10].

V. CONCLUSION

In this paper, a SRAM-based PUF cell is proposed, which utilizes subthreshold discharge to improve voltage reliability and temperature reliability. In addition, a new time strategy is designed to reduce PUF power consumption. A super-symmetrical PUF cell layout is designed and the post-simulation performance is not reduced. Meanwhile, the proposed PUF retains the memory function, which has a wide range of application scenarios.

REFERENCES

[1] Yijun Cui, Qingqing Ma, Chongyan Gu, Yue Fang, Chenghua Wang, Maire O'Neill and Weiqiang Liu, "Lightweight Modeling Attacks Resistant Multiplexer Based Multi-PUF (MM-PUF) Design on FPGA," *Electronics*, vol. 9, no. 5,815, 2020.

[2] Chongyan Gu, Weiqiang Liu, Yijun Cui, Neil Hanley, Maire O'Neill, and Fabrizio Lombardi, "A Flip-Flop Based Arbiter Physical Unclonable Function (APUF) Design with High Entropy and Uniqueness for FPGA Implementation," *IEEE Transactions on Emerging Topics in Computing*, pp. 1–1, 2019.

[3] R. Maes and V. van der Leest, "Countering the Effects of Silicon Aging on SRAM PUFs," in *Proc. IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*, pp. 148–153, 2014.

[4] W. Liu, L. Zhang, Z. Zhang, C. Gu, C. Wang, M. O'neill, and F. Lombardi, "XOR-Based Low-Cost Reconfigurable PUFs for IoT Security," *ACM Trans. Embed. Comput. Syst.*, vol. 18, no. 3, 2019.

[5] S. Chellappa and L. T. Clark, "SRAM-Based Unique Chip Identifier Techniques," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 24, no. 4, pp. 1213–1222, 2016.

[6] Y. Shifman, A. Miller, Y. Weizman, A. Fish, and J. Shor, "An SRAM PUF with 2 Independent Bits/Cell in 65nm," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–5, 2019.

[7] K. Liu, Y. Min, X. Yang, H. Sun, and H. Shinohara, "A 373-F^2 0.21%-Native-BER EE SRAM Physically Unclonable Function With 2-D Power-Gated Bit Cells and V_{SS} Bias-Based Dark-Bit Detection," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 6, pp. 1719–1732, 2020.

[8] Q. Zhao, Y. Wu, X. Zhao, Y. Cao, and C. Chang, "A 1036-F^2/Bit High Reliability Temperature Compensated Cross-Coupled Comparator-Based PUF," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 28, no. 6, pp. 1449–1460, 2020.

[9] S. K. Mathew, S. K. Satpathy, M. A. Anders, H. Kaul, S. K. Hsu, A. Agarwal, G. K. Chen, R. J. Parker, R. K. Krishnamurthy, and V. De, "A 0.19pJ/b PVT-Variation-Tolerant Hybrid Physically Unclonable Function Circuit for 100% Stable Secure Key Generation in 22nm CMOS," in *Proc. IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pp. 278–279, 2014.

[10] S. Satpathy, S. K. Mathew, V. Suresh, M. A. Anders, H. Kaul, A. Agarwal, S. K. Hsu, G. Chen, R. K. Krishnamurthy, and V. K. De, "A 4-fJ/b Delay-Hardened Physically Unclonable Function Circuit With Selective Bit Destabilization in 14-nm Trigate CMOS," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 4, pp. 940–949, 2017.