Waseda University Doctoral Dissertation

Study on High Stability and Low Energy SRAM-Based Physically Unclonable Function for Hardware Security

Kunyang LIU

Graduate School of Information, Production and Systems
Waseda University
April 2021

ABSTRACT

The global market of Internet of Things (IoT) has been expanding rapidly in recent years. With increasing edge devices connected in the network, the security of IoT has become a great concern as many of these devices operate in the environments where human involvement is limited. This requires security to be considered not only at the software level, but also down from the integrated circuits (ICs) level. One critical issue is the secure generation of secret keys, which serve as the root of trust and a fundamental block of security applications like cryptography and authentication that is based on cryptography.

Physically unclonable functions (PUFs) are regarded as a promising solution for high-security and low-cost secret key generation. Leveraging the intrinsic random process variation of silicon devices, a PUF generates a chip-specific ID bitstream on-the-fly, and this bitstream can be used for secret key realization. Since process variation such as random dopant fluctuation is extremely difficult to observe, PUF is regarded as more resilient to reverse engineering than non-volatile memories (NVMs), and therefore, it is called physically unclonable.

However, there are many challenges in PUF circuit design. The most critical problem is that PUFs suffer from bit errors, because the process variation they leverage is usually in the millivolt (mV) range. Random noise, supply voltage/temperature (VT) variations, long-term aging, etc. cause errors in PUF data generation, which are not tolerable for cryptographic applications. In addition to stability problem, PUF should also be low-energy and small-area, because IoT edge devices are generally resource-constraint. Most importantly, robustness against physical attacks must be considered in PUF circuit design.

Static random access memory (SRAM)-based PUF or SRAM PUF is one of the most important categories of PUFs. It features small area, and it has a differential bitcell structure. The differential structure makes SRAM PUF naturally more resilient to power analysis attacks, which is a kind of popular physical attack, because the supply current waveforms of reading data “0” and data “1” have little difference. However, conventional SRAM PUFs suffer from higher bit error rate (BER) than low-native-BER mono-stable PUFs. In the

conventional solution, heavy error-correcting code (ECC) circuits are implemented to meet the zero-error requirement for cryptography. This results in substantial overhead in latency, power, and area, which are not desired for IoT edge devices.

Therefore, to meet the demands, an SRAM-based PUF with natively high-stability bitcells and efficient post-processing techniques are required to reduce BER so as to reduce or even eliminate the use of ECC.

Based on these considerations, in this dissertation, two SRAM-based PUF solutions with novel bitcells and post-processing techniques are presented.

The first work achieved more than 10 times lower native BER than conventional SRAM PUFs by using EE SRAM bitcells, and the BER is comparable with the state-of-the-art mono-stable PUFs. It also achieves “zero” error by using only circuit techniques for post-processing. The proposed dark-bit detection technique achieves “100%” reduction on BER by masking 67.4% detected bitcells, while the previous work only achieved 60% BER reduction.

The second work presents a hybrid SRAM PUF that achieves 61 times lower energy than the EE SRAM PUF while having the same level of stability. In this work, device characteristics modification by HCI burn-in is applied, and the bitcell is designed to be compatible with HCI burn-in with little overhead. Through the bitcell stabilization by HCI burn-in, “zero” BER is achieved without masking loss.

The dissertation is organized with four chapters as follows.

In Chapter 1, the background of IoT security system is first introduced using an example of edge device authentication. Then, the drawbacks of conventional solution are discussed and the motivation for a PUF is derived. Afterwards, the basic concept of silicon PUFs and PUF-based authentication are introduced.

In Chapter 2, the evaluation metrics of a PUF are shown. After that, prior art on PUF is introduced, and the reason why high-stability SRAM PUF is preferred is discussed. Also, prior post-processing techniques on PUF stabilization are shown.

In Chapter 3, an EE SRAM PUF with two-dimensional (2-D) power gating and V_{SS} bias-based dark-bit detection techniques are introduced. The presented EE SRAM PUF utilizes an EE bitcell structure to improve the stability of SRAM PUF. This reduces BER by 14 times compared with conventional SRAM-based designs (3.04% BER), resulting in 0.21% native BER. The EE SRAM bitcell also has a small footprint of 373 F^2, because only nMOS transistors are used, and the p-n boundary does not exist in its layout. By a novel 2-D power gating technique, the energy is reduced by approximately 64 times to 128 fJ/bit. The dark-bit detection using an integrated bias generator successfully detects “100%” dark bits ($<5.99\times10^{-7}$ BER) in the VT range across 0.8—1.4 V and −40—120 °C. Another merit of this technique is that costly temperature sweep is not required for testing. The above-mentioned metrics have been verified with prototype chips fabricated in 130-nm standard CMOS.

In Chapter 4, a hybrid SRAM PUF using hot carrier injection (HCI) burn-in to reinforce stability is presented. The SRAM bitcell is a hybrid of EE SRAM and conventional CMOS SRAM to benefit from both designs. During evaluation, it works in the EE SRAM mode for high stability, and after that, it is switched to the CMOS SRAM mode to take advantage of the low-power feature. Moreover, this mode transition enables the PUF to work under low supply voltages (V_{DD}s) down to 0.5 V. As a result, its energy is greatly reduced to 2.07 fJ/bit. This is 61 times reduction compared with EE SRAM PUF. The hybrid SRAM bitcell is compatible with HCI burn-in stabilization. HCI effect is utilized to enlarge the V_{th} mismatch of PUF cells for stability reinforcement. By this technique, all bit errors are eliminated ($<3.90\times10^{-7}$ BER) for the 5120 measured bits across 0.5—0.7 V and −40—120 °C after 10-min burn-in. Compared with error-free prior art, the hybrid SRAM PUF achieves the lowest energy while HCI burn-in is the only technique that has no requirements on additional fabrication processes, helper data, or visible oxide damages. The above-mentioned metrics have been verified with prototype chips fabricated in 130-nm standard CMOS.

In Chapter 5, the conclusions of the dissertation are drawn.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

γ	Voltage Exponent Factor.
AES	Advanced Encryption Standard.
AF	Anti-Fuse.
A-SSCC	IEEE Asian Solid-State Circuits Conference.
BCH	Bose-Chaudhuri-Hocquenghem.
BER	Bit Error Rate.
BL	Bit-Line.
CMOS	Complementary Metal-Oxide Semiconductor.
CRP	Challenge and Response Pair.
DUT	Device under Test.
E_a	Activation Energy.
ECC	Error-Correcting Code.
EE	Enhancement-Enhancement.
$\mathbf{F^2}$	Feature Size.
FHD	Fractional Hamming Distance.
$\mathbf{FHD_{Inter}}$	Inter-PUF Fractional Hamming Distance.
$\mathbf{FHD_{Intra}}$	Intra-PUF Fractional Hamming Distance.
HCI	Hot Carrier Injection.
$\mathbf{I_{Array}}$	Bitcell Array Current.
IC	Integrated Circuit.
IEEE	Institute of Electrical and Electronics Engineers.
IoT	Internet of Things.

$\mathbf{I_{SC}}$	Short-Circuit Current.
ISSCC	IEEE International Solid-State Circuits Conference.
JSSC	IEEE Journal of Solid-State Circuits
k	Boltzmann’s Constant.
KER	Key Error Rate.
LE	Load Enable.
mV	Millivolt.
NBTI	Negative Bias Temperature Instability.
NIST	National Institute of Standards and Technology.
NVM	Non-Volatile Memory.
PE	Power Enable.
ppm	Parts per Million.
PPMA	Privacy Preserving Mutual Application.
PUF	Physically Unclonable Function.
ReRAM	Resistive Random Access Memory.
RM	Reed-Muller.
$\mathbf{R_{ON}}$	On-Resistance.
SA	Sense Amplifier.
SRAM	Static Random Access Memory.
TAF	Thermal Acceleration Factor.
TMV	Temporal Majority Voting.
$\mathbf{T_{Operation}}$	Operation Temperature.
TRNG	True Random Number Generator.
$\mathbf{T_{Stress}}$	Stress Temperature.
VAF	Voltage Acceleration Factor.

V_{DD}	Supply Voltage.
$V_{Operation}$	Operation Voltage.
V_{Stress}	Stress Voltage.
VT	Supply Voltage and Temperature.
V_{th}	Threshold Voltage.
WL	Word-Line.

Chapter 1

Background

The Internet of Things (IoT) is playing a significant role in our daily life. From consumer applications such as smart home and wearable devices, to public applications such as medical monitoring and smart city, and further to Industry 4.0, it is forecasted that the worldwide number of active connected IoT devices would be more than 20 billion by 2025 [1]. Meanwhile, the massive connections of IoT devices raise challenges to security, especially since many edge devices, e.g., sensor nodes, operate in outdoor environments where protection is limited. This allows adversaries to perform physical attacks directly on the devices [2], and therefore, security must be considered not only at the protocol or software level, but also at the level of integrated circuits (IC) design [3]. Accordingly, hardware security [4] has been a popular research topic in both academia and industry.

Cryptography is a fundamental security application that is widely used in hardware security system to encrypt information, and it is also used to perform authentication. In lightweight IoT applications, block cipher or symmetric key cryptosystem is often selected as a cryptography solution. It is based on a symmetric secret key, which serves as the root of trust. The success and failure depend on the correctness of the secret key. On the other hand, the secret key is also an important attack target for adversaries, and this has become one of the main challenges for hardware security. In the remainder of this section, the basic concept of the IoT authentication system is introduced. Then, the vulnerabilities of conventional secret key solutions are explained. Finally, the PUF-based authentication solution is introduced.

1.1 Authentication in IoT network

One of the basic schemes for authentication is to use a challenge and response pair (CRP). Generally, it includes two phases: *enrollment* and *authentication*. Taking Alice and Bob as an example, in the enrollment phase, Alice and Bob share a common secret. It means that only Alice and Bob know the response (answer) to a challenge (question). In the

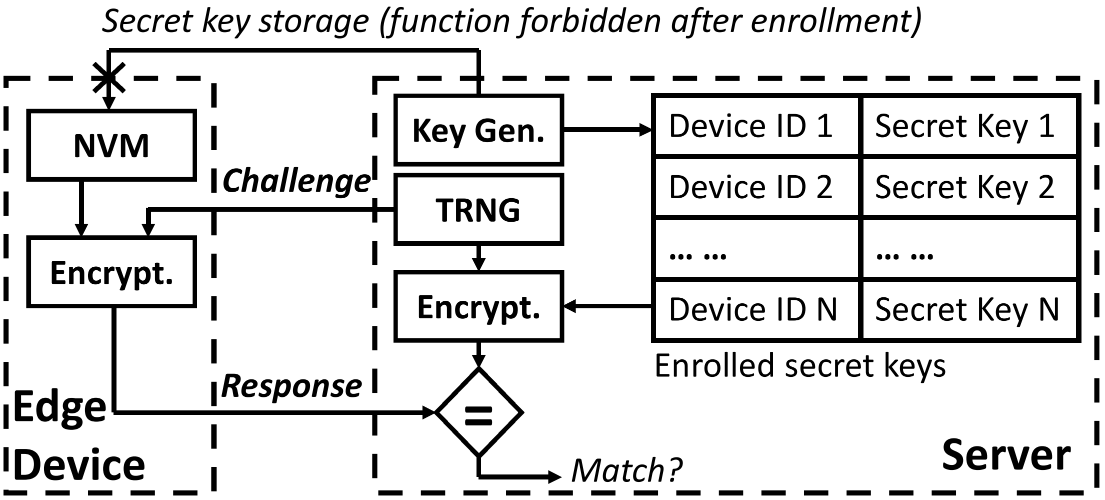


Figure 1.1: An example of conventional IoT edge device authentication based on NVM-stored secret keys.

authentication phase, Alice sends the challenge to Bob, and Bob sends back the response to the challenge. If the response is correct, the authentication is passed.

In IoT authentication, the idea is similar. However, one CRP is far not enough. For security reasons, one CRP should be used only once for one-time authentication. To match this requirement, many solutions introduce cryptography to the authentication system. One simple example [5] is depicted in Figure 1.1. On the server side, there are a true random number generator (TRNG), a key generation function, and a cryptographic engine for encryption. On the edge device side, there are a non-volatile memory (NVM) for secret key storage and a cryptographic engine for encryption. As for the crypto-engine realization, Advanced Encryption Standard (AES) [6] is the most widely used block cipher. However, for IoT lightweight purposes, more area- and energy-efficient block ciphers such as PRINCE [7], SIMON [8], and TWINE [9] are desired and widely researched.

This system also contains an enrollment phase and an authentication phase. In the enrollment phase, the server generates a secret key, and the secret key is stored in the memory of the edge device and on the server. The key generation function is applied only

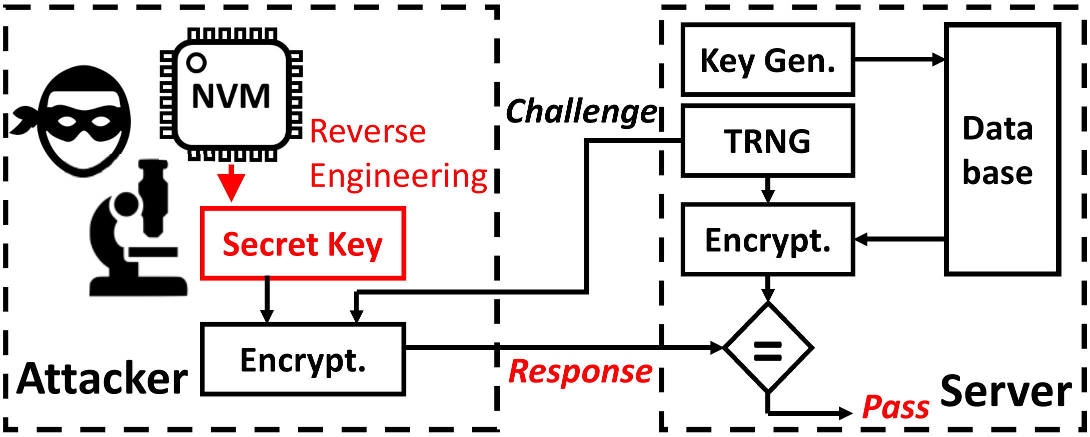


Figure 1.2: An example of an attacker hacking into the IoT network by recovering the secret key stored in NVM.

in the enrollment phase and will be disabled afterwards. In the authentication phase, the TRNG on the server generates a random bitstream as the challenge, and the challenge is sent to the edge device. Then, the cryptographic engine on the edge device encrypts the challenge based on the secret key stored in the NVM, and the encrypted bitstream is used as the response and sent back to the server. At the same time, the server also encrypts the challenge using the secret key stored on the server. In the final step, the server compares the response received from the edge device with the bitstream encrypted on the server. If they match, the edge device passes the authentication. By this scheme, any number of CRPs can be generated based on one shared secret key, and this secret key determines whether authentications succeed or not.

1.2 Vulnerability in conventional IoT authentication

Since the secret key is the root of trust, it becomes a major attack target for adversaries. In IoT system, it is even easier for attackers to recover the secret key in both invasive and non-invasive ways [10]. It has been reported that NVMs are vulnerable to these attacks [11, 12, 13, 14, 15]. Attackers can observe the secret key stored in an NVM even when the device is powered off. This allows them to gain access to sensitive information or even the permission to control the system, as shown in Figure 1.2. Consequently, a more secure solution for secret keys is demanded for IoT security.

1.3 PUF-based IoT edge device authentication

As a promising key generation solution, the concept of physically unclonable function (PUF) has emerged and become a popular research topic in IC design.

1.3.1 The concept of physically unclonable function

Physically unclonable function (PUF) is a kind of circuit that utilizes the random mismatch of silicon device to generate a process-variation-dependent bitstream [16, 17]. Since random mismatch or local variation is different between any two circuits [18, 19], the bitstream generated by PUFs are unique PUF-to-PUF. This is similar to the feature of the fingerprints of a human being, and therefore, for easy understanding, it could be regarded as the fingerprint of a silicon chip. Hence, PUF data could be used for security purposes such as identification and cryptographic key generation. One typical example of PUF is to compare the characteristics of a balanced differential pair to generate a digital "1" or digital "0".

Compared with NVMs, which typically require visible filaments [20] or oxide damage [21] to store digital data, process variation leveraged by PUF is hardly able to observe [22]. Thanks to this feature, PUF data only appear when the circuit is operating, which make it even harder to attack a PUF. Also, PUF-based solution enables on-the-fly (real-time) generation of secret keys. This eliminates the need for key generation function on the server and the corresponding key write-in function for the NVM storage shown in Figure 1.1, reducing potential attack points. In addition, a PUF can be realized with standard CMOS process, which reduces the cost caused by additional fabrication masks [22].

1.3.2 Strong PUFs and Weak PUFs

PUFs can be categorized into two kinds: Strong PUFs and Weak PUFs. Strong PUFs could generate exponentially increased CRPs compared with their size, and therefore, they can

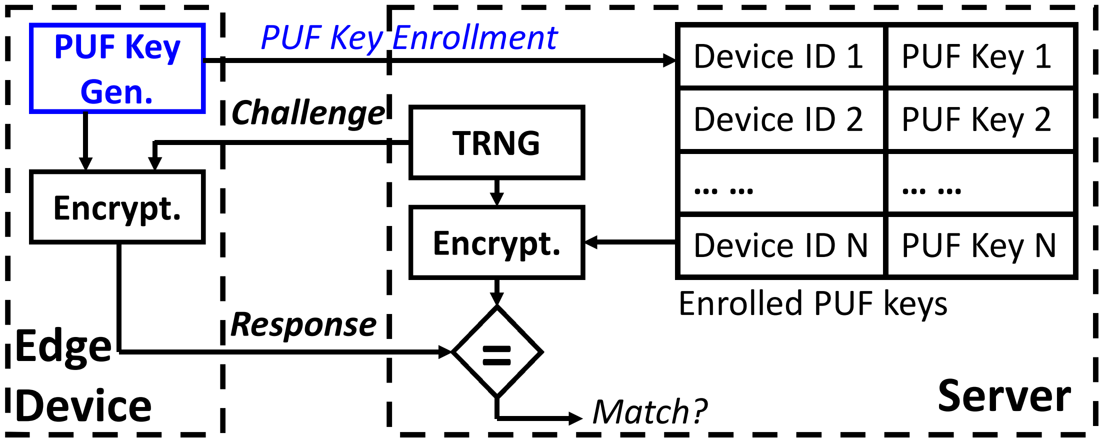


Figure 1.3: An example of IoT edge device authentication based on PUF key generation.

be directly used for authentication without cryptographic engines. However, Strong PUFs have a critical weak point, which is the vulnerability to modeling attacks [23, 24]. Consequently, Strong PUFs are still in the research phase and not widely used in real products.

Weak PUFs, in the contrary, have a CRP space that is linearly proportional to the PUF size. They are usually used for key generation. Weak PUFs are regarded as a promising solution to replace NVM for secret key storage. Compared with Strong PUFs, Weak PUFs are more mature and have been widely available in the market [25, 26, 27, 28].

In this dissertation, “PUFs” refer to the Weak PUFs, and Strong PUFs are beyond the range of this research.

1.3.3 PUF-based IoT edge device authentication

PUFs fit well with existing authentication protocols. A simple example is shown in Figure 1.3. On the server, there are a TRNG and a cryptographic engine for encryption. On the edge device, there are a PUF for key generation and a cryptographic engine. In the enrollment phase, the PUF on the edge device generates a golden key and sends it to the server. The server then stores the PUF key along with the ID of this device. As for the ID generation, there are many solutions and it could also be realized using a PUF. For conciseness, it is omitted in this dissertation. In the authentication phase, the TRNG on the server generates a random bitstream as the challenge and sends it to the edge device. Then, the edge device generates a PUF key on-the-fly and encrypts the challenge based on this PUF key. The encrypted challenge is used as the response and sent back to the server. At the same time, the server also encrypts the challenge based on the corresponding PUF key stored on the server. When the server receives the response, it compares the response with the encrypted challenge. If they match, the edge device passes the authentication. Compared with Figure 1.1, the NVM storage on the edge device and the key generation on the server are no longer required. There are also other protocols that provide better security. One example is the privacy preserving mutual authentication (PPMA) protocol shown in [29] and [30], where a TRNG is also required in edge devices to mask responses.

Despite the merits mentioned in Section 1.3.1, there are also challenges in PUF circuit design. One of the most critical issues is the stability of PUF. Although process variation leveraged by PUFs is beneficial to security, it is usually in the millivolt (mV) range, and therefore, circuit mismatch is easily influenced by noise and other factors, resulting in bit errors. There are two kinds of bit errors, *random errors* and *systematic errors*. Random errors are caused by random noise. When a PUF evaluate, if the amplitude of random noise is larger than the mismatch of a PUF cell, the PUF cell generates a random error. Systematic errors are caused by factors such as supply voltage/temperature (VT) variations and long-term aging. These factors influence the characteristics of transistors. If the polarity of a bitcell's mismatch is reversed by these factors, bit errors occur.

However, since PUFs need to work with cryptography, PUF keys must satisfy the key error rate (KER) requirement for cryptographic applications that is on the order of 10^{-6} [31]. This means the bit error rate (BER) of PUFs should be at least on the order of 10^{-7} to satisfy the KER requirement. As a result, the reduction of bit errors becomes a main challenge for PUF design.

In Chapter 2, the detailed metrics for a PUF and prior art are introduced.

Chapter 2

Preliminaries

2.1 Metrics for a PUF

2.1.1 Stability against random noise

As mentioned in Section 1.3.3, stability is a must for PUF design. Random noise is the source of random errors. The stability against random noise is generally evaluated at the nominal VT condition.

The percentage of unstable bits is usually used to evaluate the stability of a PUF. It is also called *instability*. The equation for instability calculation is shown as (2.1):

$$Instability = \frac{Number\ of\ unstable\ bitcells}{Number\ of\ bitcells} \tag{2.1}$$

In experiments, to measure instability, a PUF will be evaluated for many times, such as 500, 1000, 2000, etc. From these multi-evaluation data, the *golden data* of a PUF are decided by majority voting. When a PUF generates a datum that is different from the golden datum, it is defined as a *bit error*. When a bitcell generates even one-bit error in these multiple evaluations, this bitcell is regarded as an unstable bitcell.

Another frequently used metric is the BER. The equation for BER calculation is shown as (2.2):

$$BER = \frac{Number\ of\ bit\ errors}{Number\ of\ bitcells\ \times Number\ of\ evaluations} \tag{2.2}$$

Usually, BER does not have obvious difference as the number of evaluations increases, but instability is positively related to evaluation number. Therefore, the number of evaluations should be large enough.

2.1.2 **Stability across VT variations**

In addition to random noise, VT variations influence the characteristics of transistors and lead to systematic errors. BERs at different VT conditions are used to evaluate the stability across VT variations. It should be noted that in this case, although the PUF is evaluated at different VT conditions, the golden data should always be the majority-voting data at the nominal VT condition.

2.1.3 **Long-term reliability**

Long-term aging changes the threshold voltage (V_{th}) of transistors and can cause bit errors. To simulate long-term usage, the accelerating aging test is usually applied [32]. To do this, a chip with a PUF is baked in a high temperature and a high supply voltage (V_{DD}) stress environment to accelerate aging effects. Then, BERs are calculated based on the golden data before accelerated aging. The voltage acceleration factor (VAF) and thermal acceleration factor are estimated as (2.3) and (2.4), respectively:

$$VAF = e^{\gamma(V_{Stress} - V_{Operation})} \tag{2.3}$$

$$TAF = e^{\frac{E_a}{k}(\frac{1}{T_{Operation}} - \frac{1}{T_{Stress}})} \tag{2.4}$$

where γ is the voltage exponent factor, V_{Stress} is the stress voltage, $V_{Operation}$ is the operation voltage, E_a is the activation energy, k is the Boltzmann's constant (8.62×10^{-5} eV/K), $T_{Operation}$ is the operation temperature, and T_{Stress} is the stress temperature. The total estimated acceleration factor is the product of VAF and TAF.

2.1.4 **Uniqueness and identifiability**

Each PUF data should be unique and different from each other. To evaluate uniqueness, the hamming distance (HD) between two different PUFs is used. It is also called inter-PUF HD (HD_{Inter}). HD is defined as (2.5):

$$HD = \frac{Sum\ of\ bitwise\ difference\ between\ two\ PUF\ bitstreams}{The\ length\ of\ a\ PUF\ bitstream} \tag{2.5}$$

where these two PUFs should be the same in length.

Identifiability defines how a PUF could be identifiable from another when bit errors are considered. It is calculated with the separation ratio of HD_{Inter} and intra-PUF HD (HD_{Intra}) (HD_{Inter}/HD_{Intra}), where HD_{Intra} is the HD between any two bitstreams of a same PUF with multiple evaluations. The higher HD_{Inter}/HD_{Intra} is, the more easily a PUF is identified from another PUF considering bit errors.

2.1.5 **Randomness**

For security reasons, randomness is a fundamental requirement for a PUF. One metric to evaluate randomness is the autocorrelation among bitcells, calculated with the correlation function R_{xx}. R_{xx} of bitcells with a lag (address distance between two bitcells) j is defined as (2.6):

$$R_{xx}(j) = \frac{\sum_{i=0}^{N} x_i x_{i-j}}{N} \tag{2.6}$$

where N is the length of the PUF bitstream, x is the golden datum of a bitcell. x equals to 1 when the datum is “1”, and x equals to -1 when the datum is “0”. The ideal value for R_{xx} is zero, representing that the bitcells with a certain address distance do not correlate with each other. If the lag j is swept from 1 to half of the PUF length and all the R_{xx} values are close to zero, it proves that the PUF has a good randomness.

There are also some standard tests to test the randomness of a PUF. The most frequently used one is the National Institute of Standards and Technology (NIST) SP 800-22 randomness tests set [33].

2.1.6 **Area and energy efficiency**

Area and energy efficiency are important to PUFs because IoT edge devices are resource-constraint. As for area, usually the area in feature size (F^2) of a bitcell is used to compare PUFs across different technology nodes. It is calculated as (2.7):

$$Area\ in\ feature\ size = \frac{Actual\ area\ of\ a\ PUF\ cell}{(Feature\ size)^2} \tag{2.7}$$

As for energy, usually only the energy of PUF cell array is compared. The reason is that the total energy depends on the read-out scheme, which might be different among designs and is difficult to compare. The energy of one bit is calculated as (2.8):

$$Energy = \frac{V_{DD} \times I_{Array}}{Throughput} \tag{2.8}$$

where I_{Array} is the measured current of the bitcell array and throughput is the number of bits that are read out in one second. Compared with TRNGs, the throughput of a PUF is less critical because the PUF keys do not need to be read so often.

2.1.7 **Attack resilience**

Attack resilience is one of the most important metrics for a PUF but sometimes ignored by designers. Although the security of PUFs is considered to be strong due to the volatile feature, there are several works reporting successful attacks on PUFs [34, 35]. Also, conventional side-channel attacks, such as power analysis attacks [36, 37], can be applied to PUFs. Therefore, attack resilience should be considered for PUF design.

2.2 Bit error countermeasures

The most conventional way to solve the bit error problem is to use error-correcting codes (ECCs) [38] such as Bose-Chaudhuri-Hocquenghem (BCH) code [39, 40], repetition code, Reed-Muller (RM) code [41, 42], and etc. (see Figure 2.1). However, the implementation of ECC circuits brings to great overheads in latency, power, and area, which makes it

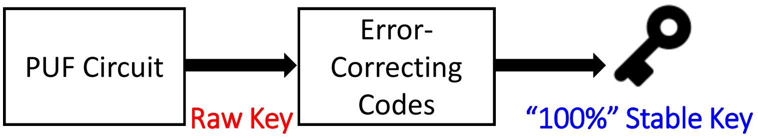


Figure 2.1: Conventional stable keys solution using only ECCs.

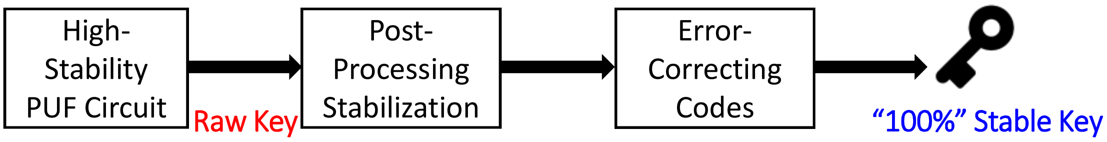


Figure 2.2: Recent stable keys solution using high-stability PUF, post-processing, and ECCs.

unsuitable for resource-constraint IoT applications. For example, it is reported in [43] that an ECC implementation using a combination of BCH codes and repetition codes requires almost 10 times of redundant bitcells compared with the key length. Also, some ECCs are reported to be vulnerable to power analysis attacks [44, 45, 46], elevating security risks.

Therefore, reducing BER and the scale of ECCs has been a target for recent researches. In order to achieve this goal, efforts have been reported in two categories. One is to design a PUF with lower native BER. The other is to use post-processing techniques to stabilize the PUF data. As shown in Figure 2.2, combining these two efforts could significantly reduce the need for ECCs, although not eliminating them. In the following two sections, previous works on PUF circuit design and post-processing techniques are introduced, respectively.

2.3 Previous PUF circuits

2.3.1 Early PUFs

The very first silicon PUF was presented in IEEE International Solid-State Circuits Conference (ISSCC) 2000 by Lofstrom et al. [47], but it was called an IC identification

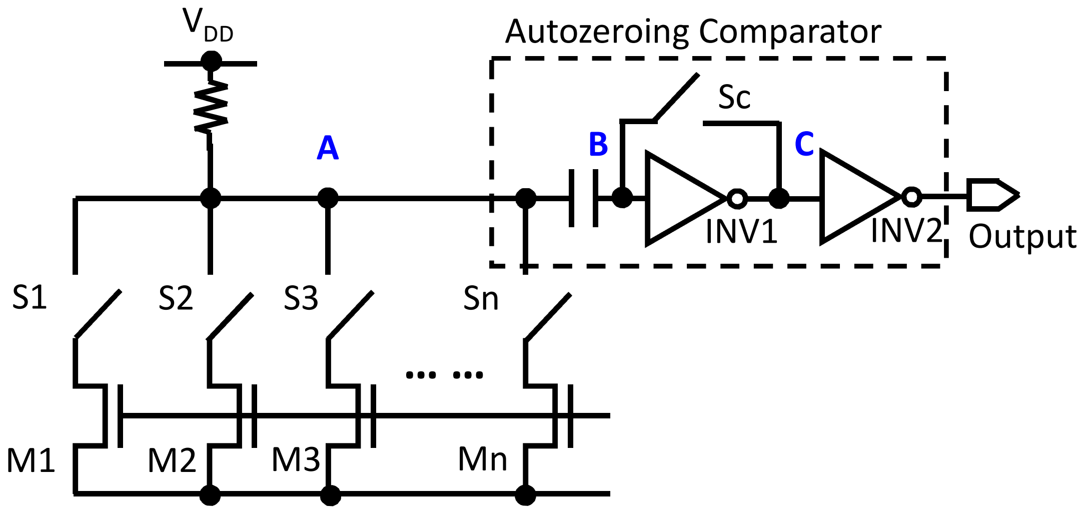


Figure 2.3: The first silicon PUF in ISSCC 2000 [47].

circuit at that time. The circuit is shown in Figure 2.3. It is composed of a transistor array controlled by switches, an autozeroing comparator, a pull-up resistor, and a coupling capacitor.

The PUF data are by comparing any two of the transistors M1—Mn. In phase 1, the switch in the comparator, Sc, is closed so that node B and node C are shorted, and INV1 and INV2 are set to a high-gain sensitive point. Then, one of M1—Mn is selected by controlling one of the switches S1—Sn. According to the V_{th} of the selected transistor, a drain voltage is generated at node A. In phase 2, Sc becomes open. Now, node B and node C keep the previous voltages, but they are in the high-impedance state. Afterwards, the circuit selects another transistor by opening the previous switch and closing another switch. According to the V_{th} mismatch between these two successively selected transistors, the voltage at node A either rises or falls, and generating a rising edge or falling edge at node B through the coupling capacitor. Then, this small difference at node B is amplified by INV1 and INV2, and one-bit output is generated at the final node. This circuit has large energy consumption due to the transistor short-circuit currents, and the comparator has a large footprint.

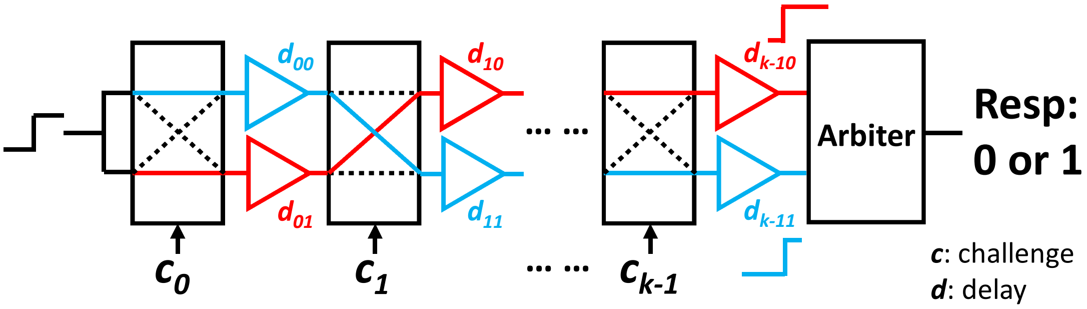


Figure 2.4: Basic structure of an arbiter PUF.

The concept of PUF emerged in 2002 with silicon implementations [48] and an optical implementation [49], respectively. Since then, various kinds of PUFs have been presented in major conferences and journals.

2.3.2 Delay-based PUFs

Delay-based PUFs generate bitstreams by comparing the delay difference of two circuit paths. The delay difference results from process variation. One of the most famous delay based PUFs is the arbiter PUF presented in 2004 by J. W. Lee et al. [50]. The structure of an arbiter PUF is shown in Figure 2.4. It has two paths, the red path and the blue path, as well as many stages of switching components, which are controlled by the challenge inputs. According to the challenges, the two paths have different delay at each stage due to the mismatch of the two buffers. The delay of each path accumulates at the final stage. The final stage is an arbiter, which compares the rising time of the two paths and determines the final response output to be “0” or “1”.

An arbiter PUF requires a big circuit for one-bit response generation and hence it is not suitable for Weak PUF implementation. On the other hand, it has a very large CRP space and can be used as a Strong PUF. For example, in Figure 2.4, the area of the arbiter PUF is proportional to k while the CRP number is proportional to 2^k. However, as mentioned in Section 1.3.2, it is vulnerable to modelling attack so that it is not secure enough for security applications. In addition to arbiter PUFs, ring oscillators are also a popular candidate to build a delay-based PUF [51, 52].

2.3.3 SRAM-based / bi-stable PUFs

SRAM-based or SRAM-like PUFs are one of the most popular PUFs. Both a foundry SRAM [53] and a custom-designed SRAM [54] can be used to implement an SRAM PUF. Also, similar circuits like sense amplifiers (SAs) [55], latches [56], and bus-keepers [32] can be grouped into SRAM PUFs, as they share common features. Since an SRAM has two stable operating points, which can store both "0" and "1", an SRAM PUF is also called a bi-stable PUF.

The circuit of a conventional six-transistor (6T) complementary metal-oxide-semiconductor (CMOS) SRAM bitcell is depicted in Figure 2.5. It leverages the power-up state as the PUF data, which depends on the process variations among the four transistors P1, P2, N1, and N2. The access transistors A1 and A2 are for data read-out and have little influence on PUF evaluations. During power-up evaluation, A1 and A2 are turned off. Due to the large gain and the positive feedback of the cross-coupled inverters pair, a small mismatch can be amplified into full-swing differential voltages at the Q and QB nodes. Then, the differential voltages are read out through the access transistors and bit-lines.

SRAM PUFs have many advantages that make them so popular. First, the footprint is small as a bitcell can be built with only six or eight transistors. Second, SRAM techniques have been widely researched and they can also be utilized on SRAM PUFs. Third, SRAM has a differential structure, which is regarded to be more resilient to power analysis attacks, because the current waveforms of reading data "0" and data "1" are identical [21]. Finally, the differential structure is easier for post-processing. For example, burn-in stabilization can be applied on SRAM PUFs, but it is difficult to apply on other kinds of PUFs. The details of burn-in stabilization will be introduced in Section 2.4.4 and 2.4.5.

However, one big issue of conventional SRAM PUFs is that their native stability is low. The reason is explained with Figure 2.6. This figure shows the butterfly curves [57, 58, 59, 60] of a 6T SRAM. A butterfly curve is a combination of two direct-current (DC) transfer

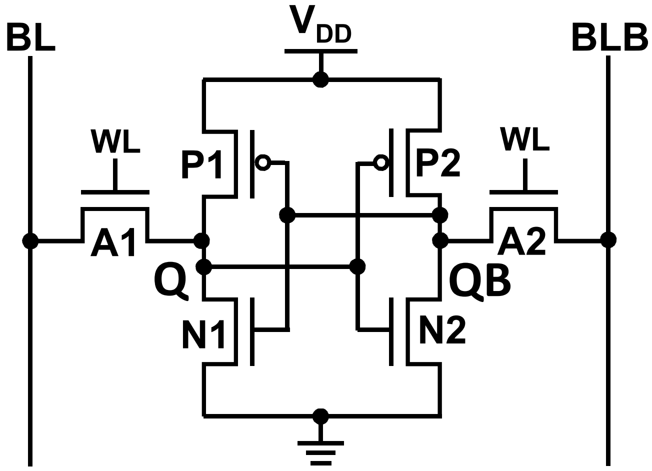


Figure 2.5: Circuit of a conventional SRAM PUF.

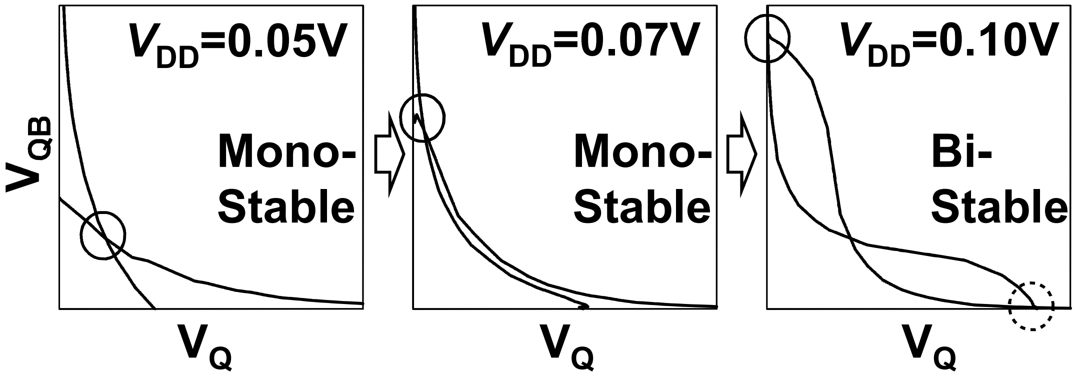


Figure 2.6: Butterfly curves of a conventional SRAM PUF during power-up, simulated with a 5-mV mismatch.

curves of the two cross-coupled inverters that compose an SRAM bitcell. The cross points of these two transfer curves represent the static operating points or the solutions of an SRAM. In Figure 2.6, it is shown that a CMOS SRAM PUF have two stable solutions even under a very low V_{DD} of 0.1 V. According to circuit mismatch, the correct solution should be the solid circle so that node QB should be charged to a higher voltage. However, in the real V_{DD} power-up transient process under 0.1 V, currents are too small to charge the node quickly, and the circuit fails to follow the DC trajectory. In this case, the actual solution can be easily influenced by random noise and flip into the wrong solution (i.e., the dotted circle in Figure 2.6). This leads to the poor native stability of conventional CMOS SRAM PUFs.

In addition to the power-up state, there exist other types of data generation for bi-stable PUFs. In [54, 61, 62], the authors present a hybrid of SRAM and delay PUF, whose data depend on a combined behaviour of the discharge of SRAM and the delay chains. Since the data do not simply rely on the power-up behaviour, this design could be more resilient to the attacks targeting the power-up behaviour of SRAM [35]. But as for stability, this hybrid design is similar to conventional CMOS SRAM PUFs.

2.3.4 Mono-stable PUFs

In order to achieve better native stability, mono-stable PUFs have emerged. The biggest difference of a mono-stable PUF is that it has only one stable solution according to its circuit mismatch, and therefore, it alleviates the solution flipping problem of bi-stable PUFs.

The first well-known mono-stable PUF was presented in ISSCC 2015 [63, 64]. Its circuit is shown in Figure 2.7, which is composed of a pair of complementary cascode current mirrors. This design also utilizes the power-up state as the PUF data. When we look at node X, the currents of M3 and M5 tend to be the same and they are mirrored to the other side (i.e., M4 and M6). If there is no mismatch in the circuit, the currents of M4 and M6 should be the same and V_X should be identical to V_{OUT}. However, the two currents are slightly different in practice because of mismatch. Due to the cascode structure, the output impedance at node OUT is very large. This large impedance then transforms the current mismatch into a voltage that is close to full swing (i.e., either GND or V_{DD}). This voltage is then further amplified and used as the PUF data.

In addition to the high stability feature due to the mono-stable characteristic, another merit of this design is that when a low V_{DD} is set, the V_{DD} can be lower than the V_{th} summation of the cascode transistors, biasing the PUF in sub-threshold region and featuring low

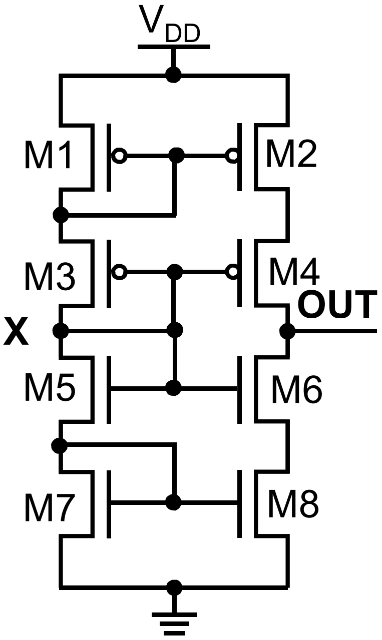


Figure 2.7: Circuit of the current-mirror-based mono-stable PUF in [63] and [64].

power. This low-power feature is preferable in IoT applications. An improved version of this design is published in IEEE Asian Solid-State Circuits Conference (A-SSCC) 2017 [65] and IEEE Journal of Solid-State Circuits (JSSC) 2018 [66].

Another major type of mono-stable PUFs utilizes a chain of amplifiers. The first work is presented in ISSCC 2016 by B. Karpinskyy et al. from Samsung Electronics [67]. The circuit is shown in Figure 2.8. The bitcell is composed of a chain of NAND gates. At the first stage, the input is shorted with the output. According to the mismatch between the first and the second stages, the voltage at node X is either slightly larger than the threshold or slightly smaller than the threshold. Then, this small voltage difference is amplified by the following stages into a full-swing voltage at node OUT. Being mono-stable, it has good stability, but the weak point is that the power consumption is quite large due to the short-circuit current (I_{SC}), especially at the first stage. Another work with reliability strategies was presented by the same group in ISSCC 2020 [68].

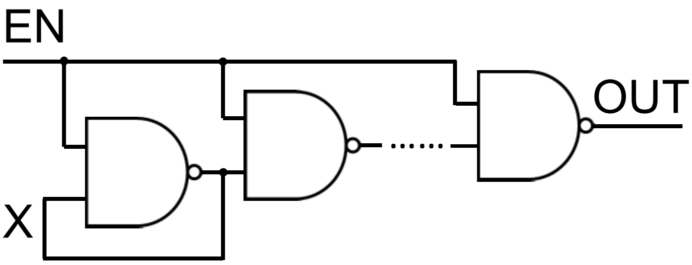


Figure 2.8: Circuit of the NAND-chain-based mono-stable PUF in [67].

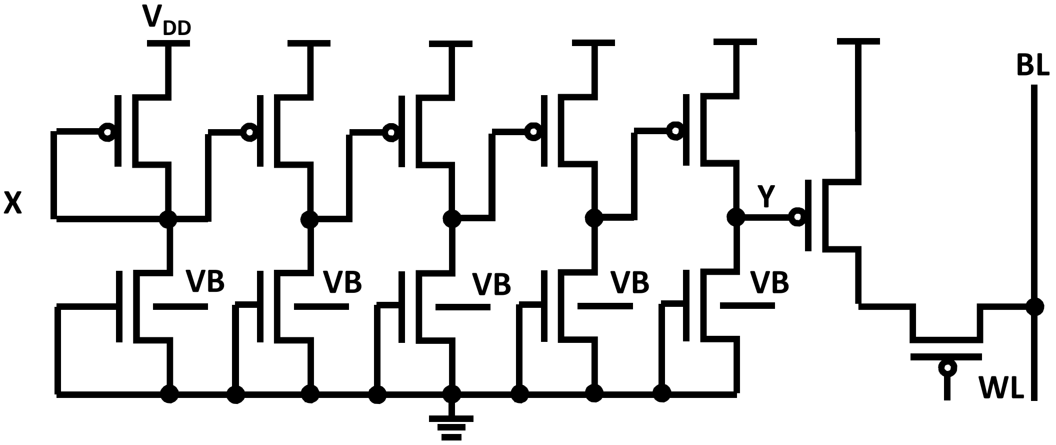


Figure 2.9: Circuit of the 2T-amplifier-chain-based mono-stable PUF in [69].

In ISSCC 2017, K. Yang et al. presented a mono-stable PUF with similar structure as [67], but the energy efficiency was improved [69]. The circuit is shown in Figure 2.9. In this work, the authors use two-transistor (2T) amplifiers instead of NAND gates. To make it low-power, the amplifiers are biased in sub-threshold region by connecting the gate of the nMOS to its source (GND). To prevent the nMOS from totally turning off, a body bias is applied. An improved version of this work was presented in ISSCC 2019 [70] and JSSC 2020 [71].

Although mono-stable PUFs achieve high native stability against random noise, it still suffers from systematic errors caused by VT variations. Also, the single-ended structure

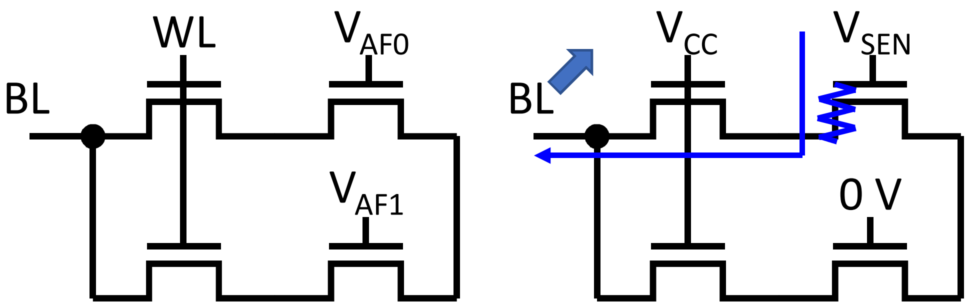


Figure 2.10: Circuit of the anti-fuse PUF in [72] and an example when AF0 has a breakdown.

might be more vulnerable to power analysis attacks, compared with PUFs with a differential structure.

2.3.5 NVM-based / NVM-like PUFs

In recent years, to realize ultra-low BER, several works have utilized NVMs or oxide breakdown to generate on-chip PUF key. These PUFs are quite unique because they don't have PUF data natively and require a programming process for data generation. The generated data are then stored in a way similar to NVMs. Therefore, they are not actually "physically unclonable" and there are arguments whether they should be considered as PUFs or not [31, 68].

One famous work was presented in ISSCC 2018 using a pair of anti-fuses (AFs) to form a bitcell [72]. The circuit is shown in Figure 2.10. Before actual usage, the anti-fuse PUF needs a self-programming mechanism using competing oxide breakdown to form PUF data. In this phase, the bit-line (BL) is set to 0 V, the word-line (WL) turns on, and a high voltage of 5.5 V is applied to both AF0 and AF1. According to the slight difference in characteristics, oxide breakdown occurs on one of AF0 and AF1, and a low-resistance path is formed between the gate of the breakdown AF and its source. Due to this path, BL is charged to a high voltage, and the other AF is relieved from the stress. During read-out, a sense voltage V_{SEN} is applied to the gate AF0 and WL turns on. If AF0 has a breakdown, BL will be charged up. In other words, if AF1 has a breakdown and AF0 does not, BL will not be charged up. Since the ON/OFF ratio of a high resistance path and a low resistance path is large (e.g., ~100), the BER of the AF PUF is in the parts per million (ppm) range.

In another work [73], the authors use a similar structure to [72], but instead of applying 5.5 V to form a hard oxide breakdown, this work applies a lower stress voltage so that only soft oxide breakdown happens. Since soft oxide breakdown is more difficult to observe compared with hard oxide breakdown, it is considered to be more secure. Also, the energy consumption becomes lower due to the lower gate current of soft oxide breakdown. However, the risk of reverse engineering is still higher than normal PUFs based on process variations.

There are also other works using resistive random access memory (ReRAM) [74] and contact failure [75] to implement a PUF. Their common features are that they have low BERs but might be more vulnerable to reverse engineering. Also, some NVMs [74] require additional fabrication processes, resulting in higher fabrication costs.

2.4 Previous post-processing techniques

In the Section 2.3, previous efforts on improving the native stability of PUFs have been introduced. However, random errors are still not eliminated, and systematic errors caused by VT variations are hardly mitigated. Therefore, post-processing is required to further reduce bit errors. In this section, previous works on post-processing will be introduced.

2.4.1 Temporal majority voting

Temporal majority voting (TMV) is one of the simplest schemes to reduce bit errors. The basic idea is to evaluate the same bitcell repeatedly for N times and take a majority voting on these N-bit data to get a final 1-bit output. An example of the 5-to-1 TMV (TMV5) is shown in Figure 2.11. In this example, four bitcells, whose golden data are "0" "0" "1" "1", respectively, need to be read. For each bitcell, it is evaluated for 5 times. The PUF evaluation data in black mean that they are correct outputs, and the evaluation data in red mean that they are erroneous data. These evaluation data are not directly used for the final outputs. Before that, a majority voting is applied on the 5-time evaluation data, and their

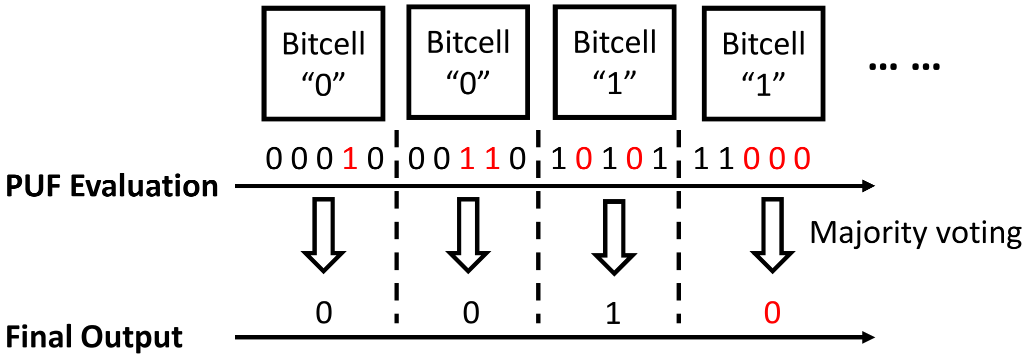


Figure 2.11: An example of TMV5.

majority value is used as the final output of this bitcell. For the first 3 bitcells, there are bit errors before majority voting. However, since the number of bit errors is smaller than the number of correct outputs, these bit errors are successfully screened out by the majority voting. In this way, TMV works as a filter for BER reduction. However, if the erroneous data happen to be the majority, TMV would fail, as the fourth bitcell in Figure 2.11 shows.

Since the golden data of PUF are determined by the majority values of multiple evaluations at the nominal VT condition, statistically the number of bit errors caused by random noise should be smaller than the number of correct outputs, and therefore, TMV could reduce bit errors induced by random noise. The cost is that with multiple evaluation, the energy for 1-bit final output becomes N times for N-to-1 TMV, so as the latency.

A main weak point of TMV is that it is not so effective on systematic bit errors caused by VT variations, because they often result in majority value flipping.

2.4.2 Dark-bit masking

Dark-bit masking is a direct method to reduce BER. Dark bits refer to unstable bits. Dark-bit masking is to screen out the unstable bits to reduce the probability of bit errors.

The key point is how dark bits are found. As for dark bits induced by random noise, they can be found easily by repeatedly evaluating the PUF. However, as for those bitcells which are potentially unstable in a different VT condition, they might appear to be stable at the nominal VT condition in which golden data are determined. One straightforward method to find out these potential dark bits is to sweep the test V_{DD} and temperature[67]. However, temperature sweep largely increases test cost.

To find out dark bits without costly temperature sweeps, a method using body bias voltage to emulate temperature variation was proposed in [69]. However, this method only achieved 60% reduction in BER at most, and the body bias needed to be applied off-chip.

Another problem of dark-bit masking is that it requires memories to store the helper data which indicate the address of dark bits. Although the helper data do not reveal sensitive information, they cause overheads in storage. These overheads could be even heavier than the PUF itself.

To alleviate the helper data storage issue, S. Satpathy et al. from Intel proposed a soft dark-bit masking technique in [61] and [62]. This technique regenerates the dark-bit mask through repeated evaluations at every time the PUF is powered up. In this way, the mask does not need to be pre-stored in a memory. However, this technique is not efficient in finding out VT variations-induced dark bits so that the BER reduction is limited.

In summary, a low-cost dark-bit detection technique which can detect all the potential dark bits without temperature sweep is desired.

2.4.3 **Dark-bit reconfiguration**

Though dark-bit masking effectively reduces bit errors, it results in bitcell loss as the masked bitcells are discarded. To stabilize PUFs without causing bitcell loss, the dark-bit reconfiguration technique was first presented in [76], which was called remapping. This PUF generates data by comparing the leakage current of a pair of bitcells. If the leakage currents are close, this bitcell pair might be unstable. If the leakage difference is large, this bitcell pair is stable. The basic concept of remapping is shown in Figure 2.12. In this example, before remapping, these two pairs of bitcells are both unstable because their leakage currents are both large and both small, respectively. However, if the combination

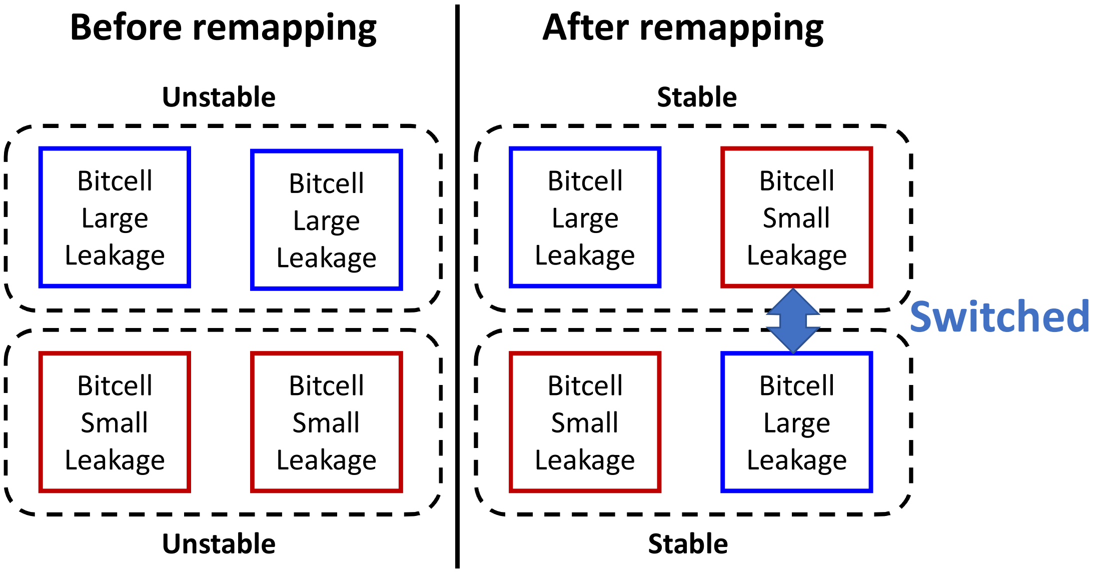


Figure 2.12: Basic concept of remapping in [76].

of the bitcells is switched to form two new bitcell pairs, they can be both stable, and no bitcell is lost. In this way, BER is reduced without bitcell loss.

Another work on dark-bit reconfiguration was presented in [70] and [71]. This PUF is a mono-stable PUF based on a chain of sub-threshold inverters. The reconfiguration technique is shown in Figure 2.13. Without reconfiguration, the stability of this PUF is determined by the mismatch between INV1 and INV2. If the mismatch is small, the bitcell is unstable. Instead of masking it, this PUF is reconfigured by closing the switch between the input and the output of INV2. After reconfiguration, INV1 and INV2 becomes parallel and can be seen as a new inverter which is a combination of INV1 and INV2. Now, its stability no longer depends on the mismatch between INV1 and INV2, but instead depending on the mismatch between the new inverter and INV3. After reconfiguration, although the bitcell can still be unstable, the probability of instability both before and after reconfiguration is squared.

Although dark-bit reconfiguration successfully reduces BER without bitcell loss, it still needs memories to store helper data. Also, its efficiency depends on the how many dark

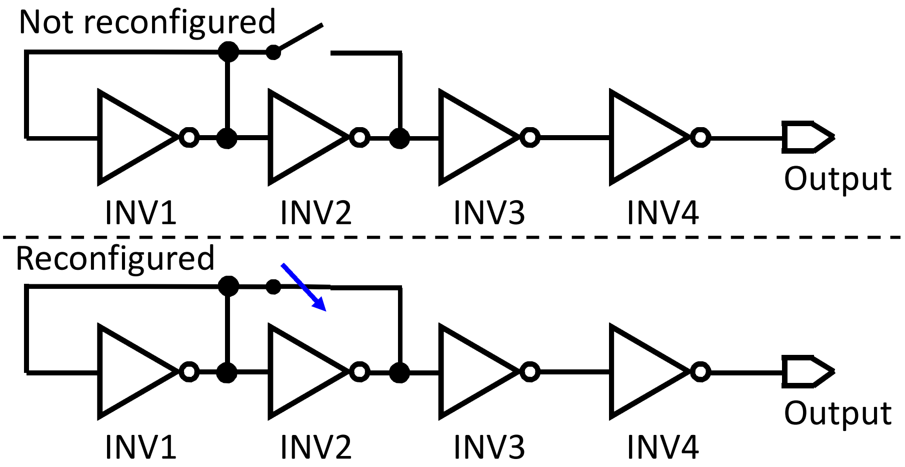


Figure 2.13: Dark-bit reconfiguration in [70] and [71].

bits are found, and hence it also requires an efficient dark-bit detection technique. In addition, its effectiveness in BER reduction is lower than dark-bit masking because the reconfigured bitcells can be unstable.

2.4.4 Negative bias temperature instability burn-in

Different from the previously introduced post-processing techniques that are based on circuit techniques, burn-in is a kind of technique that requires device technique assistance.

Negative bias temperature instability (NBTI) is a major aging effect on pMOS transistors [77]. It is usually observed when a same potential is given to the drain and the source of a pMOS transistor while the gate is negatively biased. NBTI effect causes the absolute V_{th} value of pMOS transistors to be larger, degrading performance. Even in the steady CMOS state where there is no current in the channel, NBTI can happen. Therefore, it is a crucial issue in circuit reliability.

However, the V_{th} shift caused by NBTI can be used to improve circuit metrics, especially the stability of PUFs. Since the stability of PUFs depends on their process mismatch, it can be reinforced by enlarging the V_{th} mismatch through aging. As an advantage of the differential structure, NBTI burn-in stabilization scheme can be applied on SRAM-based

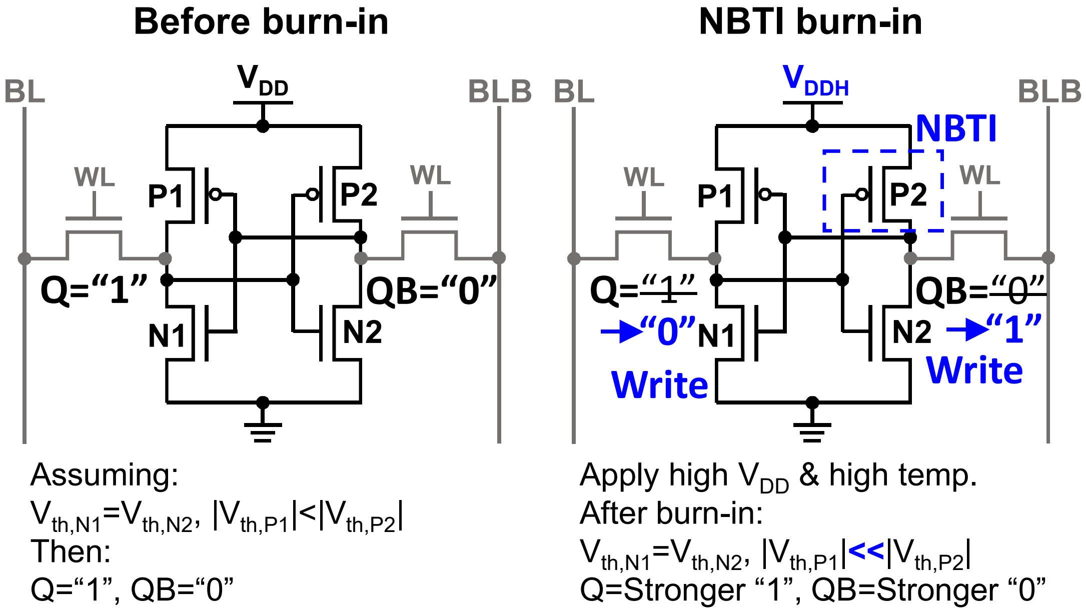


Figure 2.14: A typical example of NBTI burn-in on SRAM PUFs.

or bi-stable PUFs [54, 62, 78, 79]. Burn-in is originally applied for reliability tests, which give a high V_{DD} and high temperature condition to a device under test (DUT) to accelerate the aging effects and test its long-term performance. For the PUF stability reinforcement, burn-in is used to accelerate the aging effects which enlarge the V_{th} mismatch.

Figure 2.14 shows an example of how NBTI burn-in works on an SRAM PUF. In this example, assuming originally $V_{th,N1}$ equals to $V_{th,N2}$, and $|V_{th,P1}|$ is smaller than $|V_{th,P2}|$, the datum of this bitcell should be Q equalling to "1" and QB equalling to "0", because the pull-up ability of P1 is higher than P2. In order to reinforce the stability of this bitcell, the V_{th} mismatch of the two pMOS transistors should be enlarged. NBTI could be applied on P2 to enlarge its V_{th} and correspondingly enlarge the mismatch. To apply NBTI on P2, first, the inverse data need to be written. Now, Q becomes "0" and QB becomes "1". In this case, both drain voltage and the source voltage of P2 are high voltage and only its gate voltage is negatively biased to 0 V so that NBTI effect works on P2. Afterwards, V_{DD} is set to a

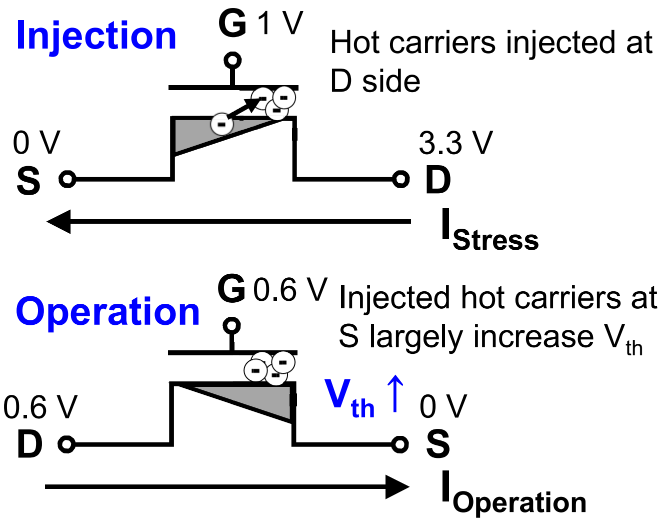

Injection
G 1 V
Hot carriers injected at D side
0 V
S
3.3 V
D
I_Stress
Operation
G 0.6 V
Injected hot carriers at S largely increase V_th
0.6 V
D
V_th ↑
0 V
S
I_Operation

Figure 2.15: An example of HCI effect on a single nMOS transistor.

high voltage and temperature is raised to accelerate NBTI, and $|V_{th,P2}|$ is enlarged. After burn-in, $|V_{th,P1}|<<|V_{th,P2}|$ is achieved and the stability is reinforced.

Although NBTI burn-in could improve PUF stability, its efficiency is limited, and it suffers from recovery [80]. For instance, in [78], the BER reduction remains at 40% after 120 hours of NBTI burn-in.

2.4.5 Hot carrier injection burn-in

Hot carrier injection (HCI) [81] is another major aging effect, in addition to NBTI. Figure 2.15 shows an example of HCI effect on an nMOS transistor. When a large drain voltage (e.g., 3.3 V) is applied to the transistor, and the transistor is biased into the saturation region, the channel electrons (carriers) are accelerated due to the high electrical field near the drain. This causes a very small portion of lucky carriers (hot carriers) to gain high enough energy and get injected into the gate oxide. After injection, they generate surface states and result in a positive V_{th} shift and degradation of I_{DS} on this nMOS transistor.

In distribution, most of the hot carriers are injected to the drain side. Therefore, the transistor becomes asymmetric after HCI. When the operation direction is the same as the HCI stress direction, the shift in V_{th} is relatively low, because the effect of hot carriers is alleviated by the depletion region near the drain. However, if the transistor operates in the reverse direction, now the hot carriers are distributed at the source side, as shown in Figure 2.15. In this case, they largely influence the V_{GS} characteristic and lead to significant V_{th} shift even with a short stress time.

Due to the large V_{th} shift, reverse HCI effect has been used in several works to improve different circuit metrics. In [82], A. Kawasumi et al. from Toshiba used HCI effect to trim the sense amplifier (SA) for a low V_{DD} SRAM. In [83] and [84], the authors used HCI to trim the access transistors in order to improve the read margin of SRAMs. HCI effect has also been used to improve the stability of bi-stable PUFs. However, different from NBTI, HCI requires channel currents and hence cannot be directly applied to conventional 6T SRAM PUFs. In [85], the authors used a custom SA PUF for HCI burn-in implementation. Although it proved that HCI burn-in could successfully eliminate bit errors even at VT corners within a short time (e.g., several minutes), the circuit of the SA PUF was very large (containing about 23 transistors), and its energy consumption was not reported.

Compared with dark-bit masking, the merits of HCI burn-in are that HCI does not need to store helper data, and it does not rely on the efficiency of dark-bit detection. Compared with NVM-based and oxide-breakdown solutions, the effect of HCI is difficult to observe while it can achieve “zero” bit errors.

2.5 Motivation and concept of this research

As introduced in Section 2.3 and 2.4, there is lack of lightweight PUF solutions which could eliminate bit errors with high-stability process-variation-based PUF bitcells and efficient post-processing techniques. As for the PUF circuit, SRAM PUFs would be a good candidate due to the small footprint and their differential structure, which could provide better resilience to power analysis attack. However, previous SRAM PUFs or bi-stable PUFs suffer from high native BER, and they need to be improved.

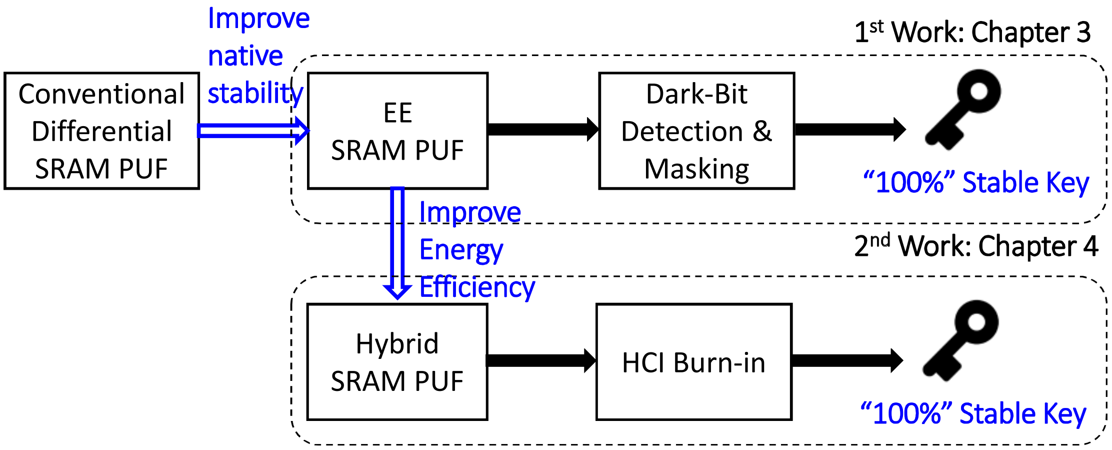


Figure 2.16: Two PUF solutions in this research, where ECCs are not used.

Therefore, in this research, my goal is to design a PUF solution that combines a high-stability SRAM PUF and an efficient post-processing technique to achieve "zero" BER without using ECCs. The concept is shown in Figure 2.16. Two works have been done to achieve this target.

In the first work introduced in Chapter 3, a new type of SRAM PUF, called enhancement-enhancement (EE) SRAM PUF, is designed. This work is based on a conference paper in A-SSCC 2018 [86] and a journal article in JSSC 2020 [87]. It is the first SRAM PUF that achieves comparable native stability with the state-of-the-art mono-stable PUFs [66, 71]. Compared with the conventional SRAM PUF in [56], EE SRAM PUF has a 14× lower native BER. To reduce the energy consumption caused by short-circuit currents, a two-dimensional (2-D) power gating technique is designed and first introduced to the bitcell array. By this technique, only one bitcell in a block is powered-up in one read-out cycle, and the energy is reduced by approximately 64× to 128 fJ/bit.

For the post-processing, a pure circuit approach is applied. A dark-bit detection technique based on V_{SS} biases generated by a lightweight integrated bias generator is implemented. Compared with the prior art [69] which could only achieve 60% BER reduction by masking the detected dark bits, the proposed technique successfully detects "100%" dark bits across all the measured VT conditions, and it is the first work which achieves “zero” BER ($<5.99\times10^{-7}$, assuming one-bit error in 3339 bitcells × 501 evaluations) with only circuit techniques.

In the second work introduced in Chapter 4, a hybrid SRAM PUF is presented. This work is based on the previous papers presented in IEEE Custom Integrated Circuits Conference (CICC) 2020 [88] and published in IEEE Journal of Solid-State Circuits [89]. In the EE SRAM PUF, it suffers from I_{SC} during operation, and the energy consumption is still larger than the state of the art even with 2-D power gating. The hybrid SRAM PUF is designed to improve the energy efficiency. It has a hybrid-mode operation. During evaluation, it works in the EE SRAM mode to achieve high native stability. After resolving the native data, it switches to CMOS SRAM for low power and to get rid of I_{SC}. Also, the stable data latching scheme in CMOS SRAM enables low-voltage operation. The lowest V_{DD} is down to 0.5 V and the minimum energy is only 2.07 fJ/bit, which is 62× smaller than EE SRAM PUF.

For the post-processing, an aid of device characteristics modification through burn-in is applied. This hybrid SRAM PUF is compatible with HCI burn-in without adding any transistor in the bitcell. After HCI burn-in, it achieves “zero” bit errors at all the VT conditions. Compared with other “zero” error PUFs including the NVM-based ones, this solution achieves the lowest energy as well as the smallest energy-area product, while HCI is the only solution that does not require visible oxide damages or filaments, helper data storage, or additional fabrication processes.

Chapter 3

EE SRAM PUF with 2-D Power Gating and V_{SS} Bias-Based Dark-Bit Detection Technique

3.1 Introduction

In this chapter, a combination of high-stability EE SRAM PUF bitcell and V_{SS} bias based dark-bit detection technique is presented. To my best knowledge, it is the first SRAM PUF that achieves high stability that is comparable with state-of-the-art mono-stable PUFs, and it is the first work that achieves “zero” error with only circuit techniques. It also includes a 2-D power gating technique to reduce the energy consumption of EE SRAM brought by the short-circuit currents, while enhancing the security through normally-off bitcells and a remanent charges clearance scheme.

First, the bitcell design of EE SRAM PUF is introduced, which shows how the EE SRAM structure can achieve higher stability than the previous SRAM PUFs. Next, the power consumption problem of EE SRAM is introduced, and its solution, the 2-D power gating technique, is described. Then, the array architecture and operations are shown. It explains how the PUF data are read out after evaluation. Afterwards, the V_{SS} bias based dark-bit detection is introduced. It shows how the V_{SS} bias detects hidden dark bits and how the bias voltage is generated using a lightweight bias generator. By masking the detected dark bits, bit errors are reduced or eliminated. After all the techniques are exhibited, experimental results based on test chips in a 130-nm standard CMOS process are shown. It proves that the proposed EE SRAM PUF successfully achieves 14× better stability than the conventional SRAM PUFs and the dark-bit detection technique can screen out all the dark bits even in extreme VT corners. After masking the detected dark bits, “zero” error is achieved without ECCs. Finally, a conclusion is drawn.

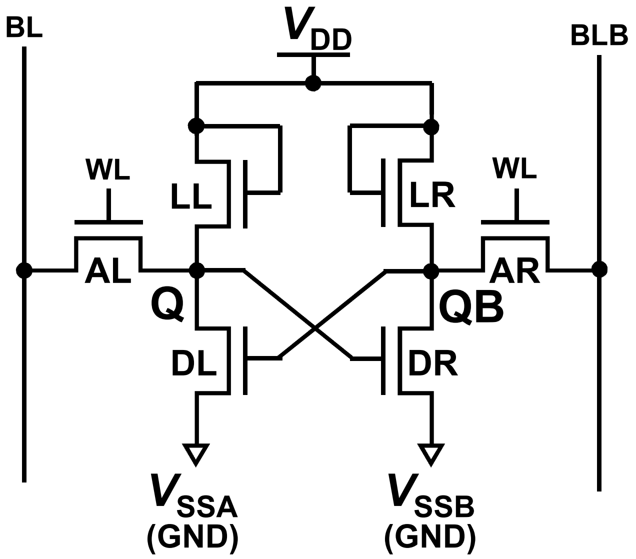


Figure 3.1: EE SRAM PUF bitcell.

3.2 EE SRAM PUF

In this section, the concept and bitcell design of EE SRAM PUF is introduced. By designing EE SRAM PUF, I intend to realize an SRAM PUF with high native stability while previous SRAM PUFs generally have poor native stability.

3.2.1 Considerations to achieve high stability in SRAM PUF

As introduced in Section 2.3.3, the reason why conventional SRAM PUFs suffer from poor stability is that they switch from the mono-stable to the bi-stable state too suddenly. When the state transition happens, the voltages of the data nodes are far from the target solution (i.e., V_Q and V_{QB} do not separate enough). Also, the state transition occurs under a very low V_{DD} that is smaller than 0.1 V. As a result, the currents in deep sub-threshold region are not large enough to charge or discharge the data nodes quickly so that they fail to follow the solution change. In this case, the data can easily flip into the wrong solution due to random noise, resulting in bit errors.

According to the analysis above, if V_Q and V_{QB} have a large separation when the PUF switches from the mono-stable state to the bi-stable state, and the state transition occurs more smoothly and under a higher V_{DD}, the stability of SRAM PUF could be improved. Under this consideration, the EE SRAM PUF is designed.

3.2.2 Bitcell Design of EE SRAM PUF

The CMOS inverter is the reason why a conventional CMOS SRAM has a sudden transition from the mono-stable to the bi-stable state. A CMOS inverter has a very steep voltage transfer curve even in the sub-threshold region. When two CMOS inverters are cross-coupled and form a latch, the positive feedback between them amplifies a small voltage difference into a large differential voltage. Due to the same reason, even under 0.1 V, a CMOS SRAM bitcell can store both “0” and “1” and works in the bi-stable state.

In order to make the state transition occur at a higher V_{DD}, I take an approach of replacing the CMOS inverter with another type of inverter with a smaller gain. In this work, the EE inverter is designed. Figure 3.1 shows the circuit of an EE SRAM PUF bitcell. Compared with a conventional 6T CMOS SRAM depicted in Figure 2.5, the two pMOS load transistors are replaced with two diode-connected enhancement nMOS transistors (i.e., LL and LR). Since each of the cross-coupled inverters is composed of two enhancement nMOS transistors, the inverter is called EE inverter and the SRAM is called EE SRAM.

Same as a CMOS SRAM, an EE SRAM PUF cell also has two driver transistors, i.e., DL and DR, for cross-coupling and pull-down, and two access transistors, i.e., AL and AR, for data read-out.

3.2.3 Operation of EE SRAM PUF

Like conventional CMOS SRAM PUFs, an EE SRAM PUF utilizes the power-up state of the bitcells as the PUF data. The data are determined by the mismatch between the cross-coupled EE inverters, or in other words, the two load transistors LL and LR, and the two driver transistors DL and DR. The two access transistors, i.e., AL and AR, are turned off during PUF evaluation and have no influence on the PUF data. After evaluation, PUF data

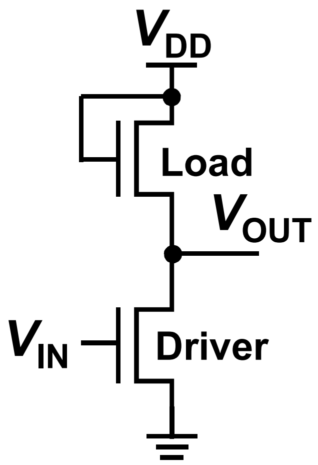


Figure 3.2: Circuit of an EE inverter.

in the bitcells are read out through the access transistors and the differential bit-lines, i.e., BL and BLB, with a complementary sensing scheme.

With a different structure, the DC transfer characteristic of an EE inverter is different from a CMOS inverter. Figure 3.2 shows the circuit of an EE inverter. Since both transistors work in the saturation region when V_{IN} is close to V_{OUT}, from the long-channel model [90], we get (3.1) and the DC transfer function (3.2), ignoring body effect:

$$\frac{\beta_{Load}}{2}(V_{DD}-V_{OUT}-V_{th,Load})^2=\frac{\beta_{Driver}}{2}(V_{IN}-V_{th,Driver})^2 \tag{3.1}$$

$$V_{OUT}=V_{DD}-V_{th,Load}-\sqrt{\frac{\beta_{Driver}}{\beta_{Load}}}(V_{IN}-V_{th,Driver}) \tag{3.2}$$

where $\beta=\mu C_{OX}(W/L)$, μ is the carrier mobility, C_{OX} is the gate capacitance of a unit area, W is the channel width, and L is the channel length. From (3.2), we can get (3.3):

$$Gain=\left|\frac{dV_{OUT}}{dV_{IN}}\right|=\sqrt{\frac{\beta_{Driver}}{\beta_{Load}}} \tag{3.3}$$

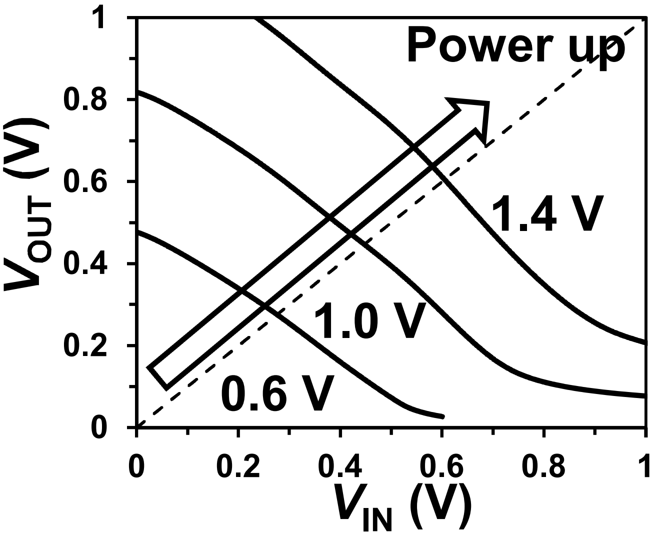


Figure 3.3: Simulated DC transfer curves when V_{DD} = 0.6 V, 1.0 V, and 1.4 V.

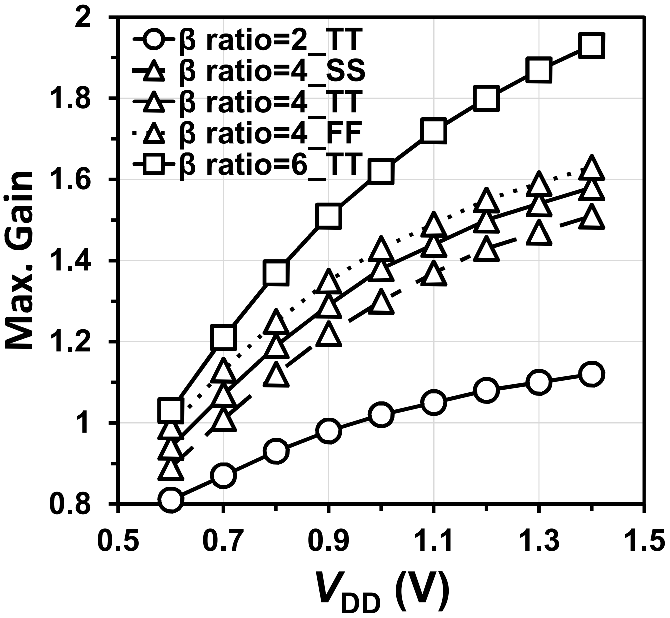


Figure 3.4: Simulated maximum voltage gains of an EE inverter in the saturation region with respect to V_{DD} and β ratio.

It reveals that the gain of the EE inverter depends on the β ratio between the driver transistor and the load transistor. Since the voltage gain is constant when body effect is ignored, the DC transfer curve of the EE inverter is more linear than a CMOS inverter. The simulated DC transfer curves at the three V_{DD}s of 0.6 V, 1.0 V, and 1.4 V are shown in Figure 3.3. They show that these curves are in a linear shape.

In addition to the β ratio, body effect also has influence on the DC transfer curves. In Figure 3.2, the body of the load transistor is connected to GND so that it is applied with a reverse body bias as its V_{BS} equals to $-V_{OUT}$. Considering body bias, we have:

$$V_{th,Load} = V_{th,Load0} + \gamma(\sqrt{2\phi_F + V_{OUT}} - \sqrt{2\phi_F}) \tag{3.4}$$

where $V_{th,Load0}$ is the V_{th} of load transistor without body effect, γ is the body effect coefficient, and $2\phi_F$ is the surface potential. From (3.2), we get:

$$\frac{dV_{OUT}}{dV_{IN}} = -\frac{dV_{th.Load}}{dV_{OUT}} \times \frac{dV_{OUT}}{dV_{IN}} - \sqrt{\frac{\beta_{Driver}}{\beta_{Load}}} \tag{3.5}$$

From (3.4), we have:

$$\frac{dV_{th.Load}}{dV_{OUT}} = \frac{\gamma}{2}(2\phi_F + V_{OUT})^{-\frac{1}{2}} \tag{3.6}$$

Combining (3.5) and (3.6):

$$\left(1 + \frac{\gamma}{2\sqrt{2\phi_F + V_{OUT}}}\right)\frac{dV_{OUT}}{dV_{IN}} = -\sqrt{\frac{\beta_{Driver}}{\beta_{Load}}} \tag{3.7}$$

$$Gain = \left|\frac{dV_{OUT}}{dV_{IN}}\right| = \frac{1}{1 + \frac{\gamma}{2\sqrt{2\phi_F + V_{OUT}}}}\sqrt{\frac{\beta_{Driver}}{\beta_{Load}}} \tag{3.8}$$

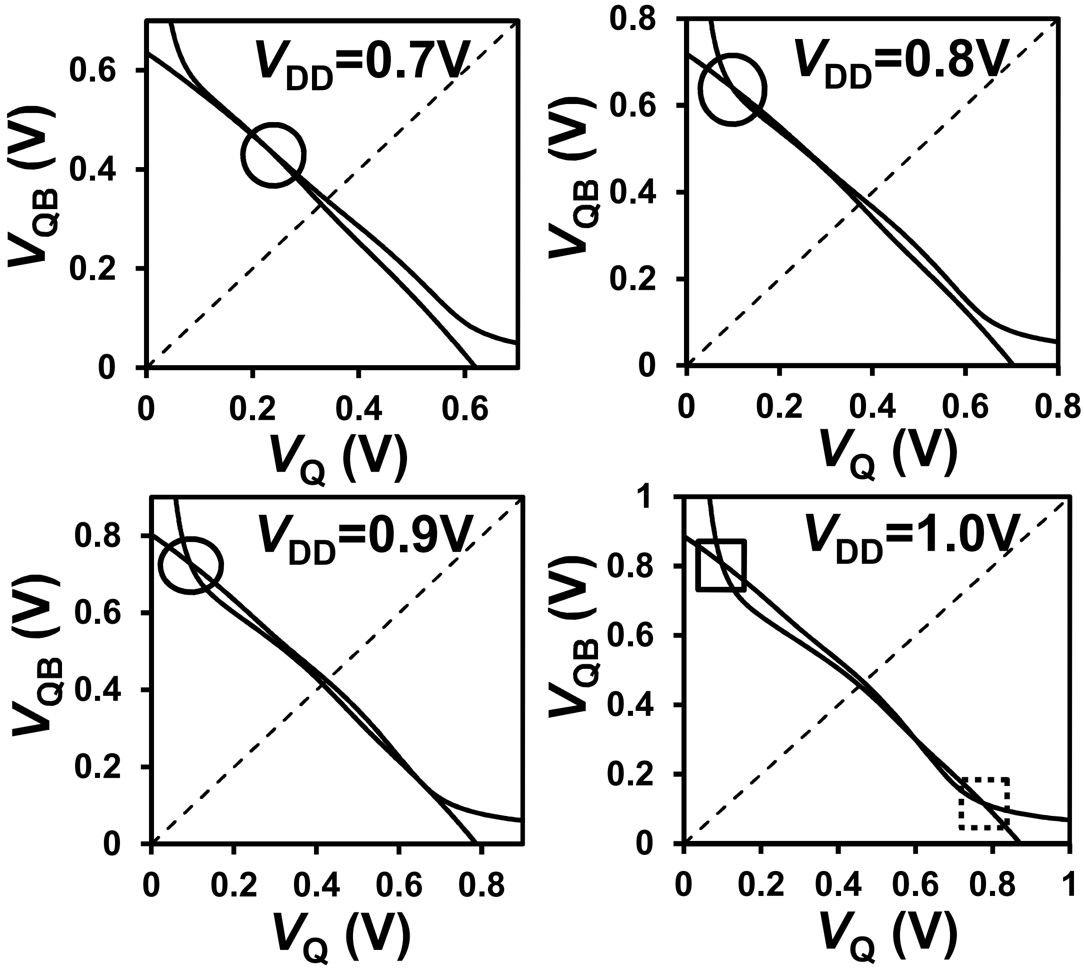


Figure 3.5: Butterfly curves of an EE SRAM PUF cell during power-up evaluation. The curves are simulated with a typical 20-mV mismatch.

It shows that the gain of an EE inverter has a positive relation with V_{OUT} in addition to the β ratio. As V_{OUT} increases following V_{DD}, the voltage gain of an EE inverter is also positively related to V_{DD}. It is depicted in Figure 3.3 that the slope of the DC transfer curve becomes steeper as V_{DD} goes up. Figure 3.4 shows the simulated gains of an EE inverter with respect to V_{DD} and β ratio. It verifies that the gain is positively related to both factors.

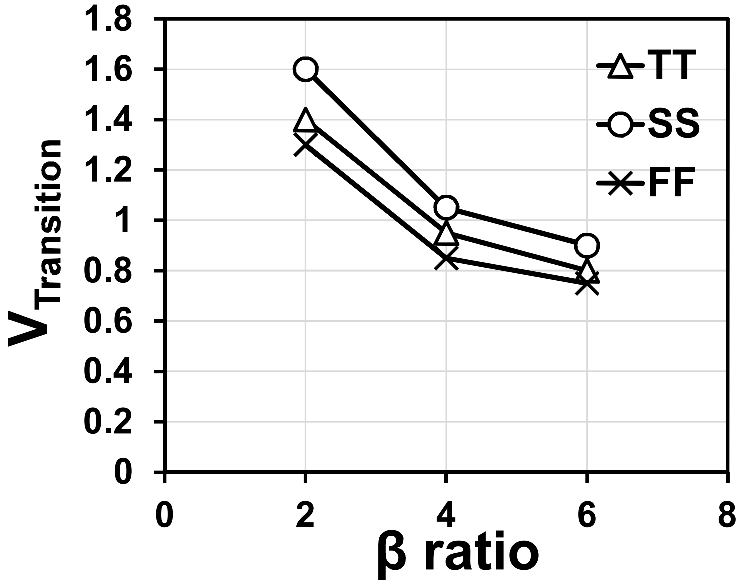


Figure 3.6: Simulated transition voltages where a PUF cell changes from the mono-stable to the bi-stable state with respect to the β ratio at three process conditions.

Table 3.1: Transistor size of the EE SRAM PUF cell.

Transistor	Width	Length (μm)
Load (LL, LR)	1×	0.52
Driver (DL, DR)	1×	0.13
Access (AL, AR)	1×	0.80

This feature is important to the stable evaluation of EE SRAM PUF. Figure 3.5 shows the butterfly curves of an EE SRAM PUF cell during power-up evaluation. Since the DC transfer curves are more linear, the bitcell stays in the mono-stable state even when V_{DD}

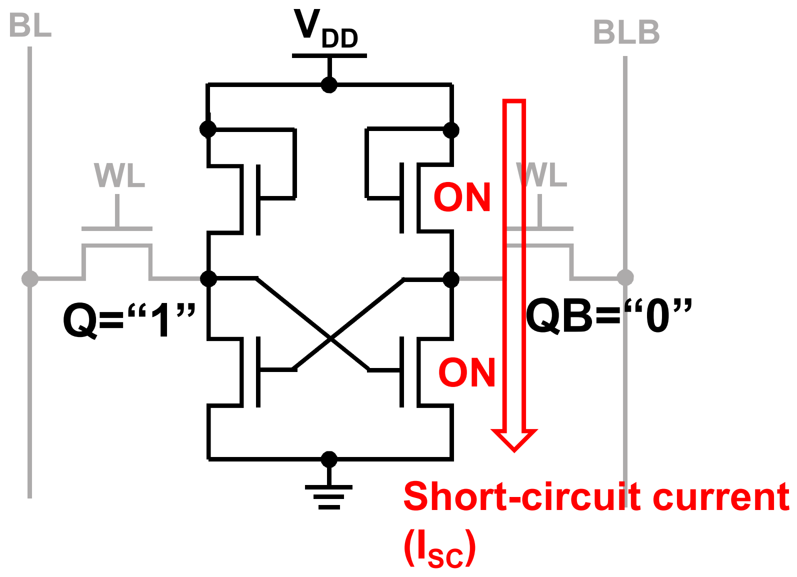


Figure 3.7: Short-circuit currents in EE SRAM PUF.

reaches 0.9 V. Therefore, under low V_{DD}s, it does not suffer from the solution flipping problem like conventional CMOS SRAM PUFs. As the voltage gain increases gradually during power-up, the bitcell smoothly changes from the mono-stable state to the bi-stable state at around 1.0 V. At a higher V_{DD}, the conductivity of the transistors is high enough, or in other words, the currents charging or discharging node Q and node QB are large enough for V_Q and V_{QB} to follow the change of the mono-stable solution point shown as the solid circles in Figure 3.5 as V_{DD} increases to 0.9 V. When V_{DD} becomes 1.0 V, and V_Q and V_{QB} separate in the bi-stable state, the solution smoothly goes to the solid rectangle shown in Figure 3.5. The probability of going to the dotted rectangle is very low.

Because the β ratio relates to the gain and the transition voltage from the mono-stable state to the bi-stable state, it should be carefully chosen. Figure 3.6 shows the simulated transition voltages under different β ratios. A larger β ratio results in a lower transition voltage so that the minimum V_{DD} could be lowered to reduced energy consumption. In this work, to balance the area and power, a β ratio of four is chosen. The sizes of transistors are listed in Table 3.1. A long length is chosen for the load transistors for energy reduction as

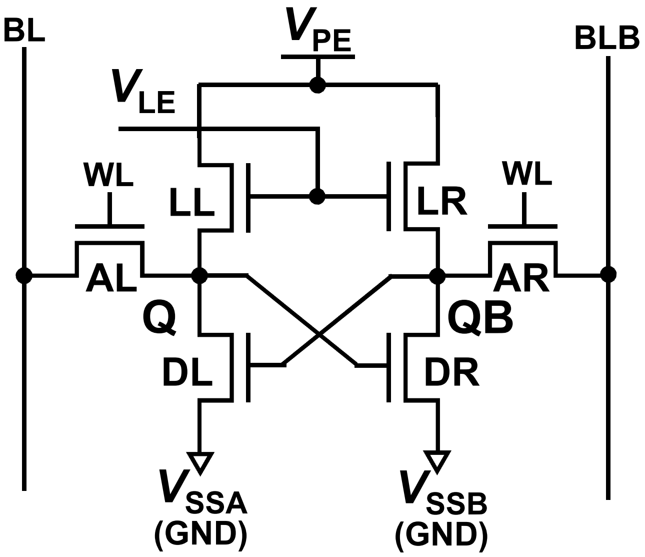


Figure 3.8: Circuit of a revised PUF cell for 2-D power gating.

well as setting the β ratio. Since write-in operation is not needed for a PUF, a long length is used for the access transistors to ensure read stability.

3.3 2-D power gating and array architecture

Although the EE structure could improve the stability of SRAM PUF, it also leads to a short-circuit current problem that increases the energy consumption of the PUF. An example is depicted in Figure 3.7. If Q equals to "1" and QB equals to "0", both the load and the driver transistor on the right-hand side are turned on, and this results in a short-circuit current. For another concern, it might provide a potential attack point for adversaries if the PUF data could be generated by directly powering up V_{DD}, such as the invasive attacks in [34] and the remanence decay side-channel attack in [35]. In order to alleviate both concerns, a 2-D power gating technique is implemented.

The basic concept of 2-D power gating is to power gate bitcells by controlling both vertical and horizontal addresses so that only the cross-point bitcells are powered up. To implement the 2-D power gating, the gate and the drain of the load transistors are separated as V_{LE} and

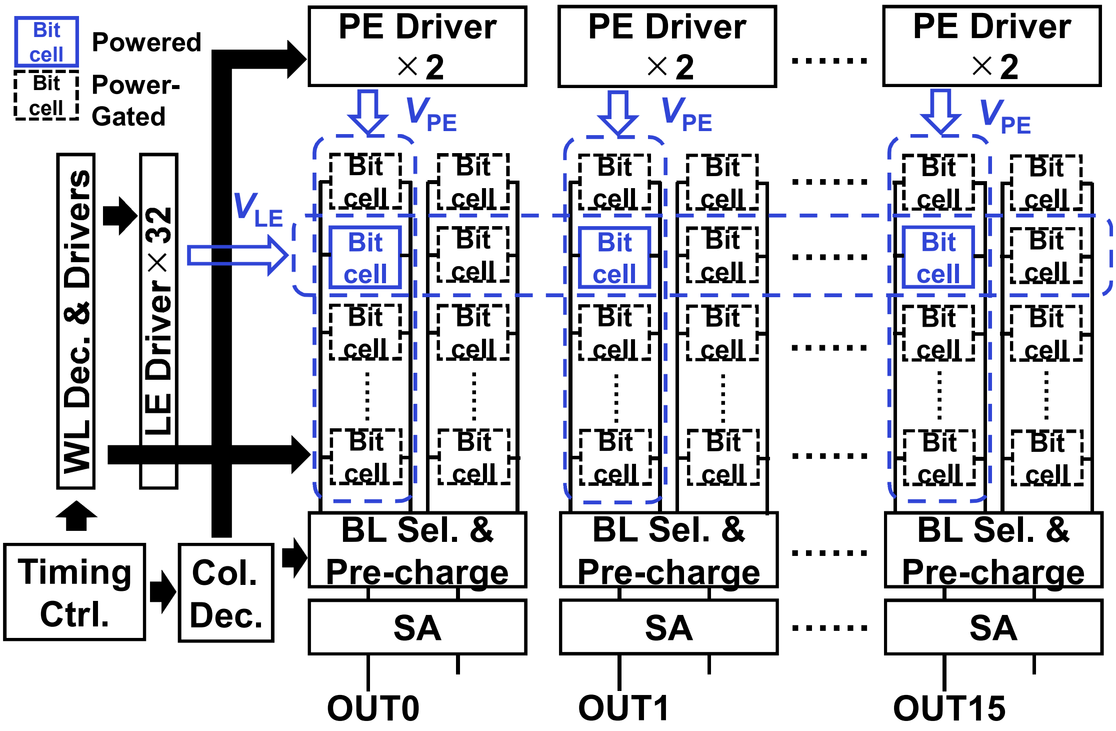


Figure 3.9: Array architecture of EE SRAM PUF.

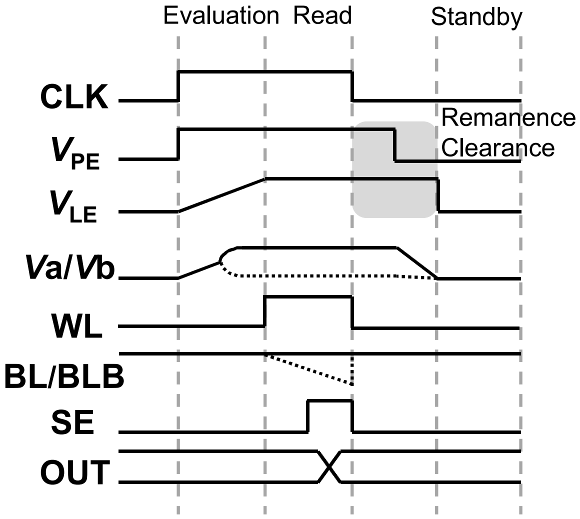


Figure 3.10: Operation waveforms of EE SRAM PUF in one read cycle.

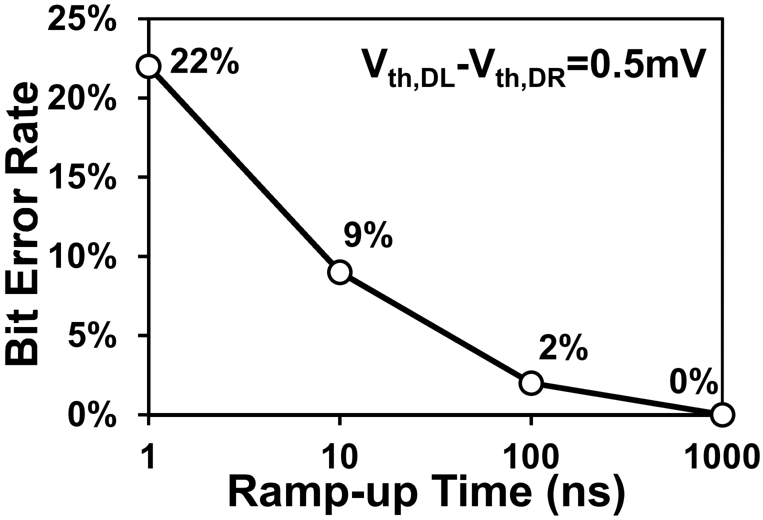


Figure 3.11: Simulation results on the relationship between ramp-up time and BER. A 100-time Monte Carlo simulation is performed on a bitcell with a small mismatch of 0.5 mV.

V_{PE}, respectively, as depicted in Figure 3.8. V_{LE} and V_{PE} are controlled with load-enable (LE) drivers and power-enable (PE) drivers, respectively. Each LE driver controls the V_{LE} of one row and each PE driver controls the V_{PE} of one column. Both the drivers are CMOS inverter-like circuits. The array architecture is shown in Figure 3.9. The array is physically arranged as 32 rows by 32 columns. Logically, it has 64 words by 16 bits. Each parallel output bit comes from a bitcell block. In each bitcell block, there are two columns that share the same column-control logic circuits and SA.

The operation waveforms are shown in Figure 3.10. Normally, when CLK=0, all the bitcells are in OFF state as V_{PE} and V_{LE} are at the GND level. When a read cycle starts, the V_{LE} of one selected row and the V_{PE}s of the selected columns are powered up through the drivers. Therefore, only one selected bitcell in one block is powered up instead of powering up all the bitcells. Since there are 64 bitcells in a block, it could be expected that the energy consumption is reduced by ~64× thanks to the 2-D power gating. After the targeted bitcell is powered up, the word-line (WL) of the selected row is activated and the differential BLs

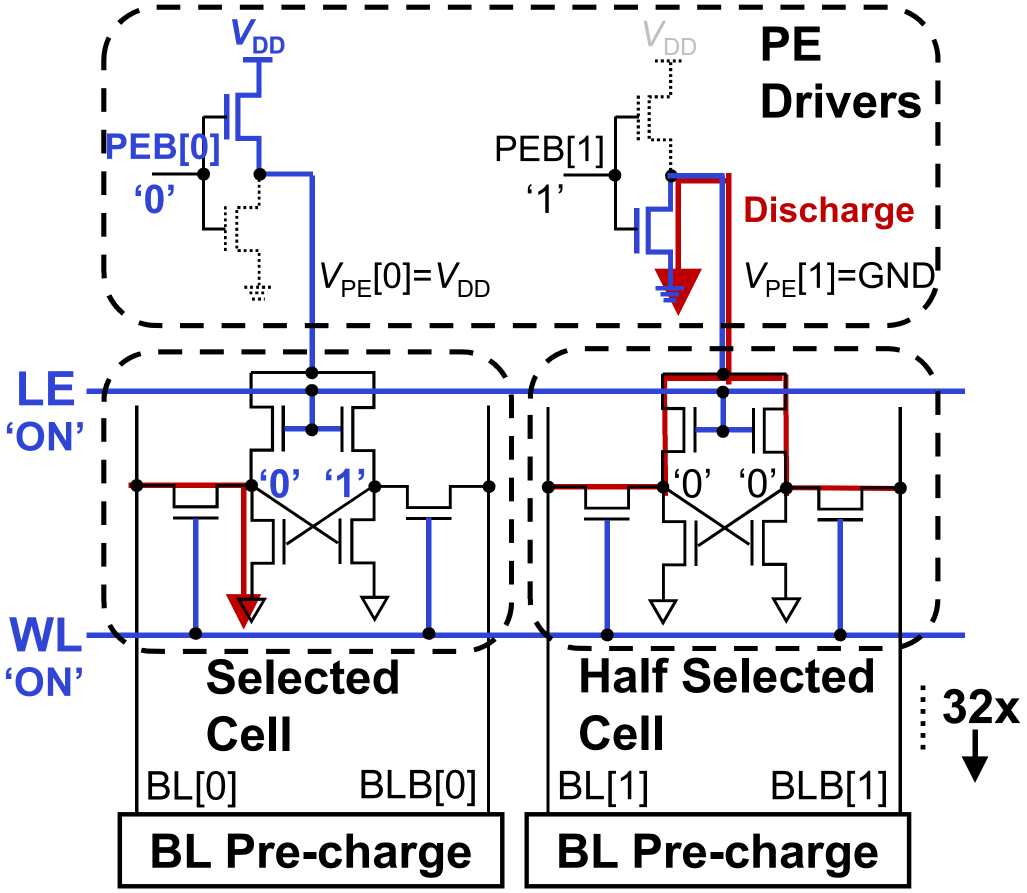


Figure 3.12: Wasted energy issue due to half-selected cells.

or BLBs are discharged depending on the PUF data. The difference between BLs and BLBs is then sensed by the SAs so that the PUF data are read out. After the read-out operation is finished, V_{LE} and V_{PE} are cut off again. Therefore, all the bitcells are turned off in the standby phase and the standby power is very small.

As mentioned in Section 3.2, a slow state transition is beneficial to the stability. The Monte-Carlo simulation results in Figure 3.11 indicate that a longer ramp-up time leads to a lower BER. The speed of state transition is related to the ramp-up rate of V_{LE}. Therefore, in this work, long-channel transistors are used for the LE drivers so that their ramp-up could be slower. The ramp-up time is set to ~200 ns in simulation. Although V_{LE} is set with a slower

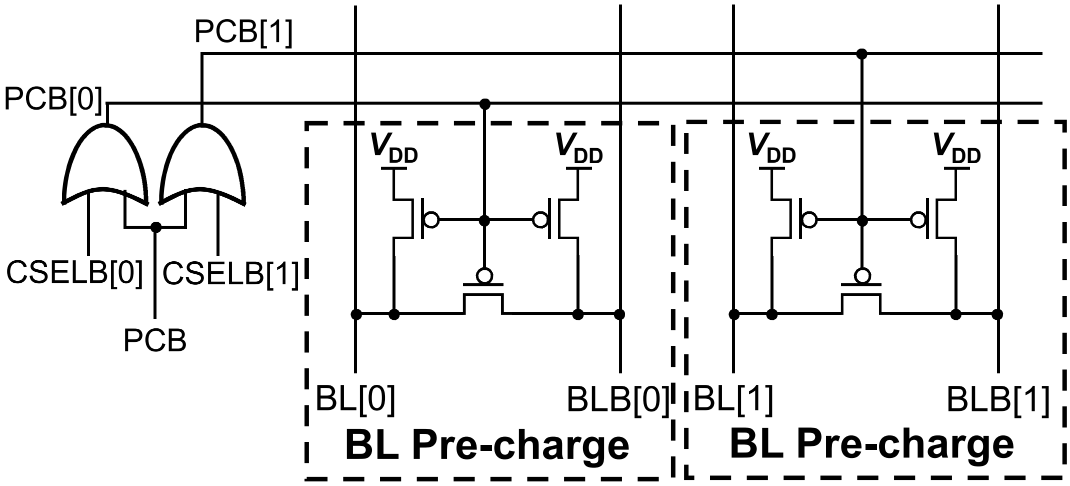


Figure 3.13: The dual PCB signals scheme to solve the wasted energy issue.

ramp-up, this time is still much faster than the 10 μs to 500 ms ramp-up for conventional SRAM PUFs [91].

Despite the contributions on energy reduction, there exists a wasted energy issue due to half-selected cells. This issue is illustrated with Figure 3.12. When the selected cell is being read and the WL is activated, the charges in the BLs of unselected columns flow through the access transistors, the load transistors, and the pull-down nMOS transistors of the PE drivers to GND. As the bitcells are read row-by-row, the repeated charge and discharge of unselected BLs cause wasted energy. To solve this problem, a dual pre-charge-enable signals (PCB) scheme is applied. The control circuits are shown in Figure 3.13. By this design, the bit-line pre-charge circuits of the two columns in a block are separately controlled. When the selected column is being read, the BLs of the unselected column would not be pre-charged. Therefore, in this design, PUF data are read out row by row so that the wasted energy is reduced by the presented scheme.

In addition to energy reduction, the architecture in this work strengthens the security of the EE SRAM PUF in three aspects. First, all the bitcells are normally turned off, and only one bitcell in one block is powered up in one read-out cycle. This limits the access to PUF data and reduces the risk of PUF key leakage. Second, a remanence clearance scheme is

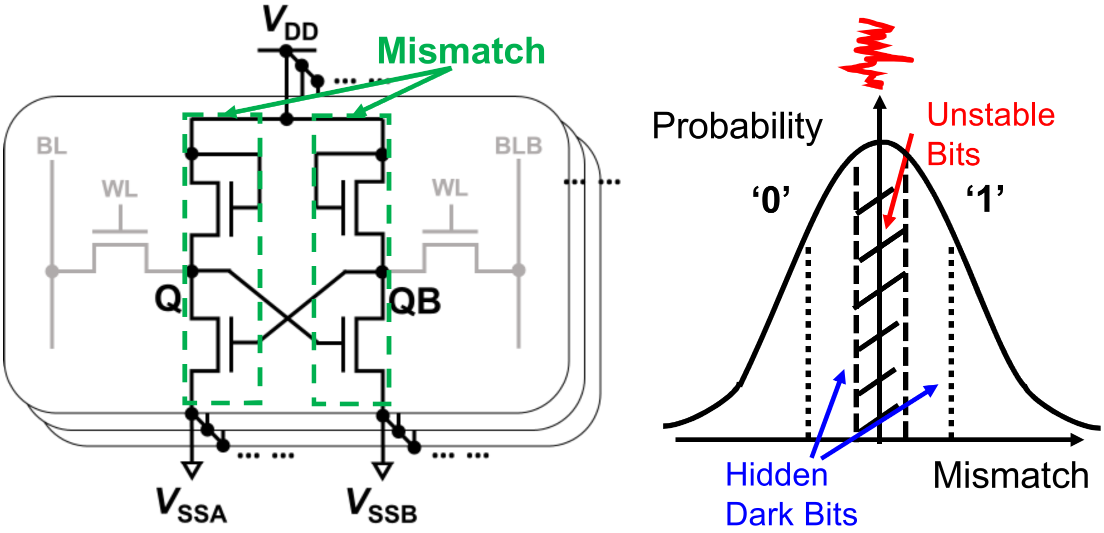


Figure 3.14: Concept of hidden (potential) dark bits.

designed to nullify a reported side-channel attack based on the remanence decay of SRAM cells [35]. As shown in Figure 3.10, after a read-out cycle finishes, V_{PE} is cut off before V_{LE}. By this scheme, the remanence at node Q and QB flow through the load transistors and the pull-down transistors of PE drivers to the GND, and therefore, they are cleared in a short time. In simulation, the charges are cleared within 10 ns even under a very low temperature of −40 °C. Third, the complementary sensing scheme reduces the risk of power analysis attacks since the energies of reading "1" and "0" are the same [21].

3.4 V_{SS} bias based dark-bit detection

Dark-bit (i.e., unstable or potentially unstable bits) masking is an efficient post-processing method to reduce bit errors. However, one big challenge is that many dark bits do not appear unstable at the nominal VT condition, but when V_{DD} or temperature changes, they become unstable. It is difficult to find out these hidden dark bits at the room temperature. The conventional method to detect these dark bits is to sweep the test temperature [67].

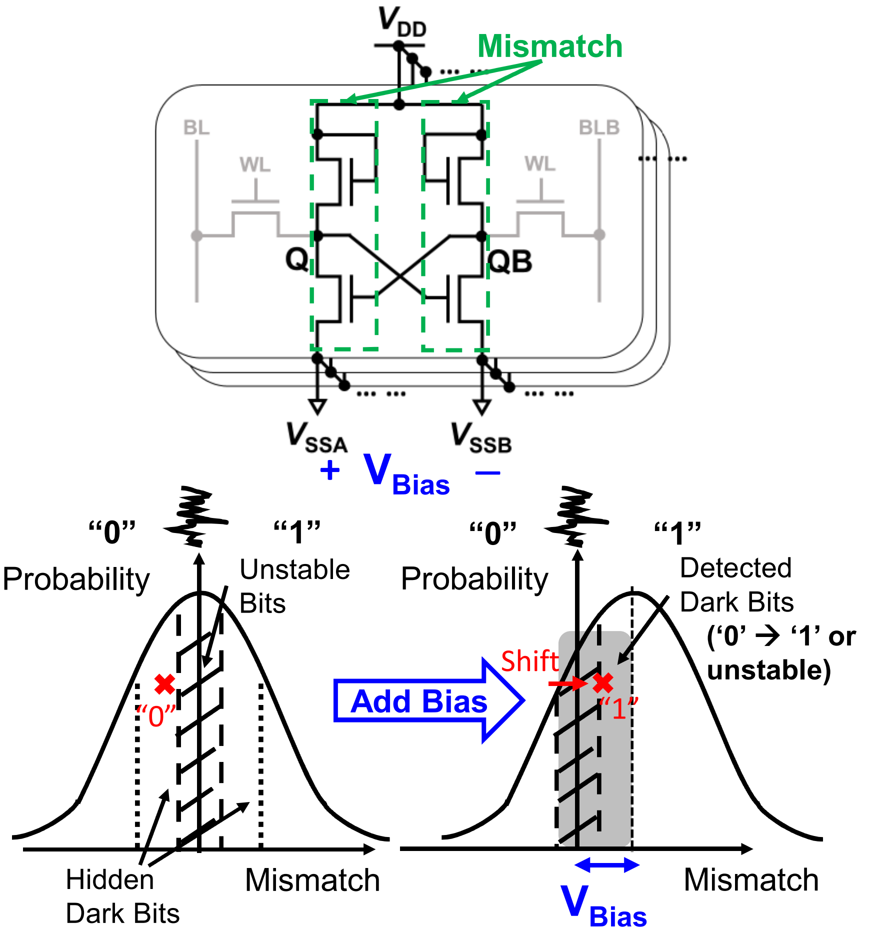


Figure 3.15: Concept of dark-bit detection based on V_{SS} biasing.

However, temperature sweep largely increases the test cost. Therefore, a technique to detect these dark bits under the room temperature is desired.

Hereby, I present a dark-bit detection technique based on V_{SS} biasing.

3.4.1 Concept of V_{SS} bias based dark-bit detection

Figure 3.14 shows the concept of hidden or potential dark bits. The stability of EE SRAM PUFs depends on the mismatch between the cross-coupled EE inverters. The mismatch follows normal distribution [92, 93]. For those bitcells which have a very small mismatch, they are easily influenced by random noise and appear unstable at the nominal VT

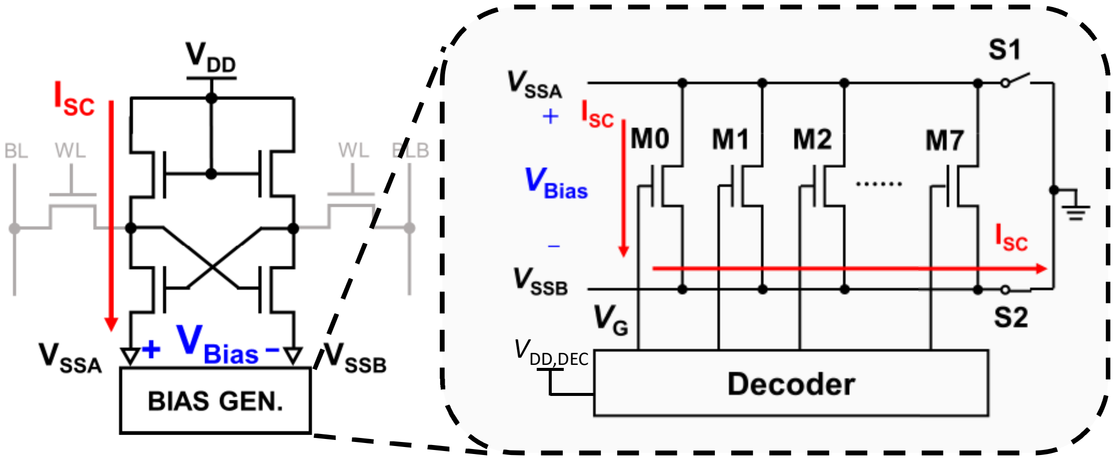


Figure 3.16: Integrated bias generator for bias voltage generation.

condition. For those bitcells which have a very large mismatch, they are always stable, even at the VT corners. However, for those bitcells which have a mismatch that is neither very large nor very small, they appear to be stable at the nominal VT condition, but they may become unstable at a different VT condition. These bitcells are the hidden or potentially unstable bits.

In order to find out these dark bits, the proposed method is to apply a bias voltage to shift the distribution to the right-hand side or to the left hand-side [92]. In the case of EE SRAM PUFs, a bias voltage between the V_{SSA} and V_{SSB} port could be applied. As shown in Figure 3.15, when the distribution is shifted by a proper bias voltage, a hidden dark bit, whose value is “0” before being biasing, could be shifted to the “1” region, and its value is flipped or becomes unstable. By observing the value, this dark bit can be found out. In this example, the distribution is shifted to the right-hand side by a bias voltage on the V_{SSA} port. It can also be shifted to the left-hand side by a bias voltage applied on the V_{SSB} port.

3.4.2 Circuit implementation for bias generation

In order to apply bias voltages on chip, an integrated bias generator is designed. It generates a bias voltage based on the short-circuit currents of EE SRAM cells and the IR drop caused by the I_{SC}. Therefore, the generator is lightweight. The circuit is shown in Figure 3.16. The bias generator is composed of an eight-transistor ladder, two switches, and a decoder which selects one of the eight transistors. These transistors have different sizes so that they have different on-resistances (R_{ON}s). One side of the ladder is connected to the V_{SSA} and V_{SSB} ports, and the other side is connected to the GND. One of the switches is closed and the other is open, in order to control the polarity of the bias voltage.

In the example of Figure 3.16, S1 is open and S2 is closed. In this case, the V_{SSB} node in the bias generator is connected to the GND and the bias voltage is applied on the V_{SSA} node. During the operation, the I_{SC} flow from the EE SRAM PUF to the V_{SSA} port of the bias generator. In this design, only one of M0—M7 is turned on. The current then flows through the selected transistor and generates an IR drop. This IR drop is used as the bias voltage. In the example of Figure 3.16, M0 is selected and the voltage across M0 is used as the bias voltage. This voltage depends on the I_{SC} of EE SRAM and the R_{ON} of the selected transistor M0. We can change the bias voltage by tuning the R_{ON}. R_{ON} could be changed not only by selecting the transistors with different width (W) and length (L), but also by changing the gate voltage (V_G) through the V_{DD} of the decoder ($V_{DD,\ DEC}$). The equation for R_{ON} and the bias voltage (V_{Bias}) are shown as (3.9) and (3.10), respectively:

$$R_{ON} = \frac{1}{\mu C_{OX}(W/L)(V_{DD,DEC} - V_{th})} \tag{3.9}$$

$$V_{Bias} = I_{SC} \times R_{ON} \tag{3.10}$$

3.5 Experimental results

3.5.1 Prototype chip tape-out

To verify the effectiveness, prototype chips are designed in a standard 130-nm CMOS process and tested. The micrograph of the chip and the layout of the bitcell are shown in Figure 3.17. In each chip, there are 1K bits. Due to the absence of p-n boundary, the bitcell has a compact footprint. The actual area is 6.30 μm^2 and the corresponding feature size is 373 F^2.

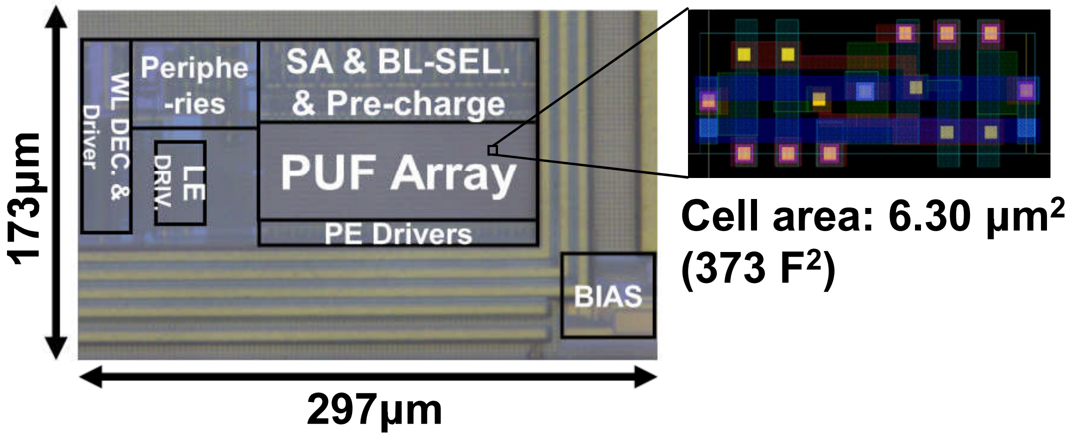


Figure 3.17: Micrograph of the prototype chip and the layout of the bitcell.

3.5.2 Stability at nominal condition

The stability at the nominal condition indicates a PUF's anti-noise ability. It is usually evaluated with BER and the percentage of unstable bitcells (instability). To evaluate the stability at the nominal VT condition, 20 chips (20K bits) are measured at 0.8 V and 23 °C.

The measurement results of the average BER and instability as well as that of the worst chip are shown in Figure 3.18. For accurate results, up to 2000-time evaluations are performed. First, when we look at the instability, it is found that it increases with respect to the number of evaluations. With 500 evaluations, the average instability is 1.82%. With 2000 evaluations, the average instability is 2.14%. The instability of the worst chip is 2.54% with 2000 evaluations. It is also found that BERs do not have an obvious relationship to the number of evaluations. With 2000 evaluations, the average and the worst BER are 0.21% and 0.34%, respectively.

3.5.3 Stability across VT variations

VT variations can inverse the mismatch of PUF cells and lead to majority flipping as well as great BERs. In this work, the measured V_{DD} range is 0.8—1.4 V and the measured

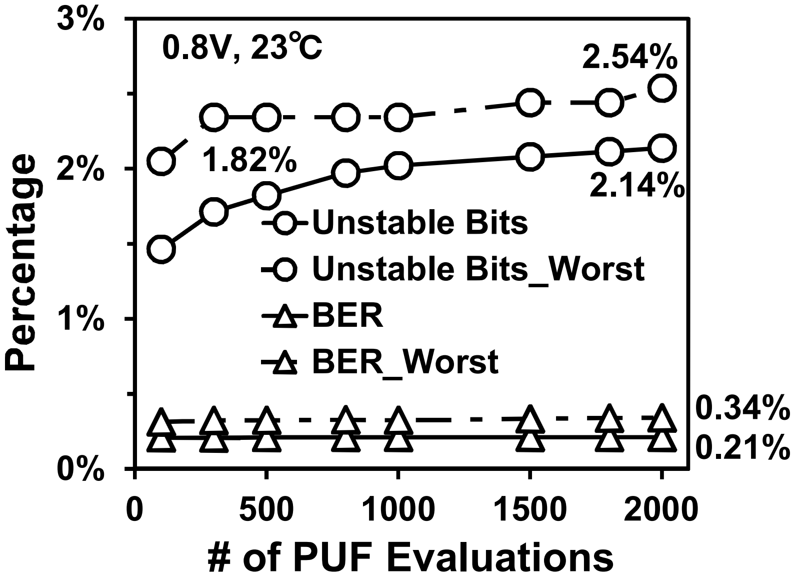
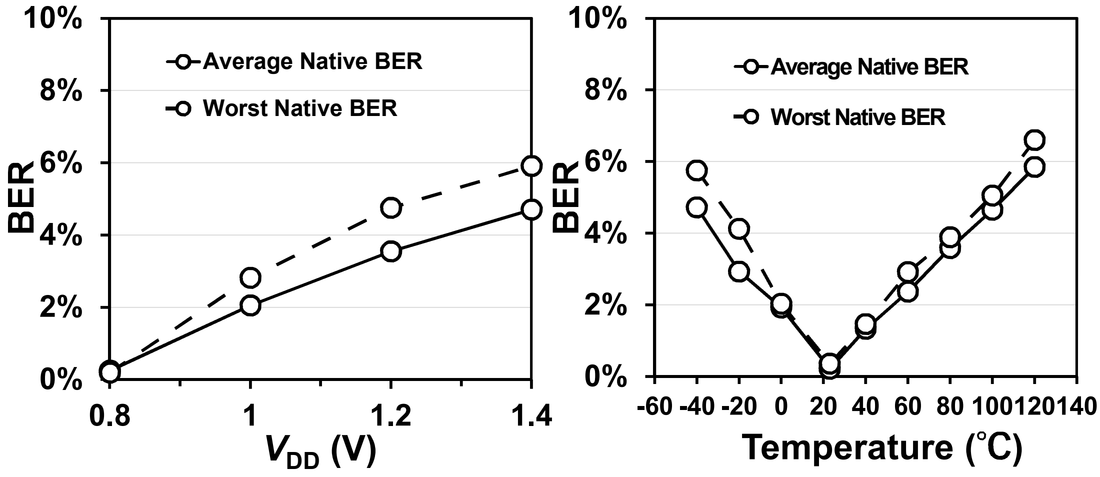


Figure 3.18: Measured BER and instability of 20 chips at 0.8 V and 23 °C.

temperature range is −40—120 °C. Ten chips (10K bits) are measured for 500 evaluations at each VT condition. For BER calculation, the data measured at a different VT condition are compared with the golden data, which are the majority values at the nominal condition of 0.8V and 23 °C.

The measured average native BERs with respect to V_{DD}s are shown with the solid line with circles in Figure 3.19(a). At 1.4 V, the average native BER is 4.70% (0.783%/0.1 V). The BERs of the worst chip are also shown with the dotted line with circles. The worst BER is 5.91% at 1.4 V.

The measured average native BERs with respect to temperature are shown with the solid line with circles in Figure 3.19(b). At −40 °C, the average native BER is 4.69%. At 120 °C, the average native BER is 5.85%. The BER at 120 °C is higher because the temperature difference from the room temperature is larger. When we look at the BER per 10 °C data, the value in the −40 to 23 °C range is 0.75%/10 °C, and the value in the 23 °C to 120 °C range is 0.60%/10 °C. BERs of the worst chip are also shown with the dotted line with circles. At −40 °C, the worst BER is 5.75%, and at 120 °C, the worst BER is 6.60%.



(a)

(b)

Figure 3.19: Measured average and the worst native BERs with respect to (a) V_{DD} variations and (b) temperature variations.

3.5.4 Dark-bit detection at nominal VT condition

The proposed dark-bit detection technique is applied at the room temperature to efficiently find out the hidden dark bits and contribute to the reduction of bit errors caused by VT variations. Before detection, the R_{ON}s of the transistors in the bias generator are measured. The measurement procedure is that, first, the V_{SSB} is set to be connected to the GND by closing the switch at the V_{SSB} side. Then, the targeted transistor is selected through the decoder. Afterward, a 30-mV voltage is applied to node V_{SSA} and the current is monitored. Bitcells are turned off to avoid the interference of bitcell currents. R_{ON} is then calculated by dividing the 30 mV by the measured current value. The measured R_{ON}s are listed in Table 3.2.

For the dark-bit detection measurement, the same 10 chips (10K bits) used for the VT variation measurements are tested. Each test chip is evaluated for 100 times under several bias conditions to find out hidden dark bits. Those bitcells which never change their data

Table 3.2: Measured R_{ON}s of the transistors in the bias generator.

Transistor	Width (µm)	Length (µm)	$V_{DD,DEC}$ (V)	Measured I (µA)	Measured R_{ON} (Ω)
M0	45.0	0.13	0.80	256	117
M1	30.0	0.13	0.80	191	157
M2	21.0	0.13	0.80	149	201
M3	14.0	0.13	0.80	109	276
M4	10.0	0.13	0.80	80	377
M5	7.5	0.13	0.80	60	498
M6	5.2	0.13	0.80	44	688
M7	3.8	0.13	0.80	30	1003
			0.71	17	1804
			0.66	9	3366
			0.64	7	4328
			0.62	5	5687

are regarded as stable bits, and those bitcells that change their values under a bias voltage are regarded as dark bits.

Figure 3.20(a) and Figure 3.20(b) show the BERs at different V_{DD} conditions after masking the detected dark bits and the corresponding BER improvement, respectively. The BER improvement is calculated as follows:

$$BER\ Improvement = \frac{BER_{native} - BER_{masked}}{BER_{native}} \tag{3.11}$$

The lines with triangles, rectangles, and rhombuses represent the BERs and BER improvement after 12.7%, 28.8%, and 60.5% masking, respectively. The corresponding R_{ON}s are 276 Ω, 688 Ω, and 3366 Ω, respectively. In the range of 0.8 V to 1.0 V, a 12.7% masking reduces the BER from 2.05% to 0.21%. In a wider range of 0.8 V to 1.4 V, a

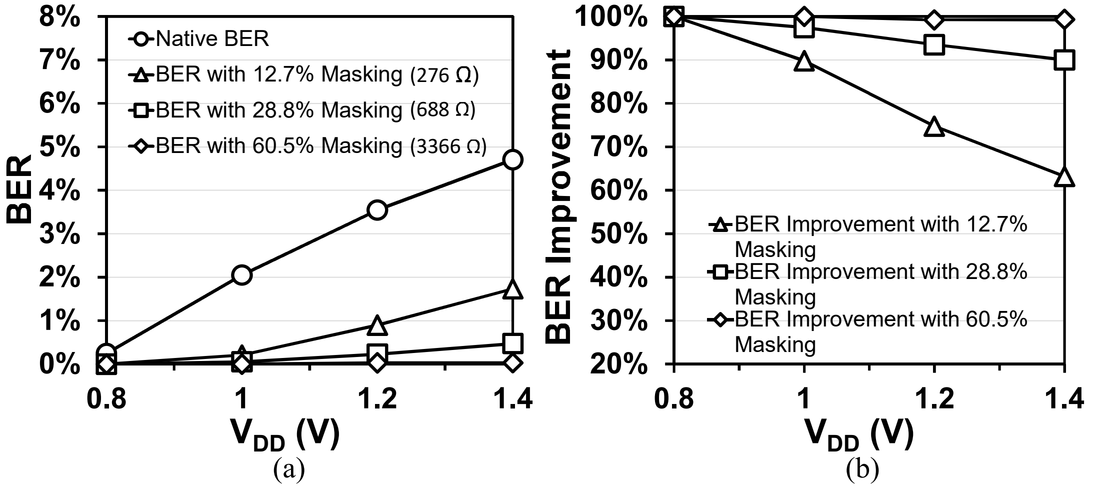


Figure 3.20: Measured (a) BERs and (b) BER improvement with respect to V_{DD} variations.

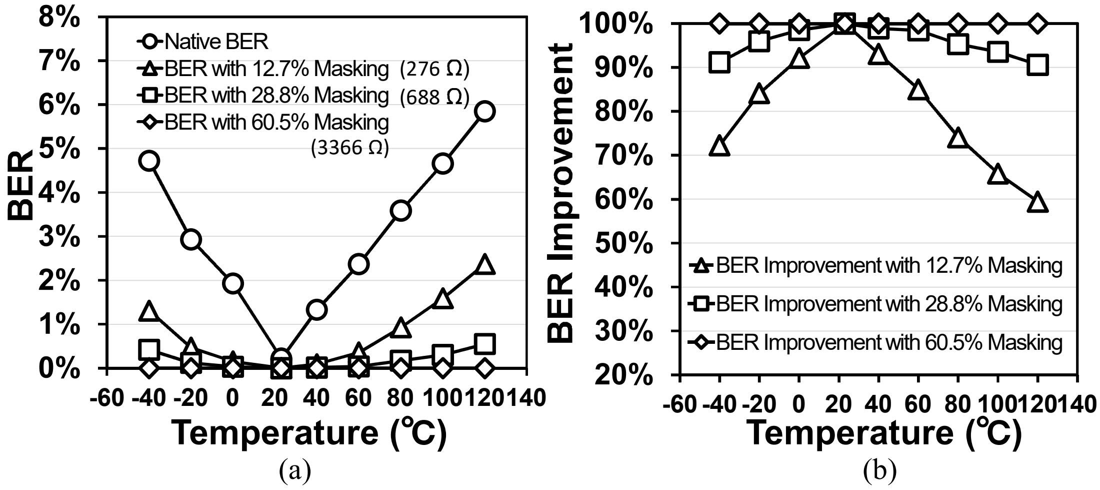


Figure 3.21: Measured (a) BERs and (b) BER improvement with respect to temperature variations.

28.8% masking reduces the BER from 4.70% to 0.47% and achieves 90% improvement. After the 60.5% masking, the BER is reduced to 0.03%.

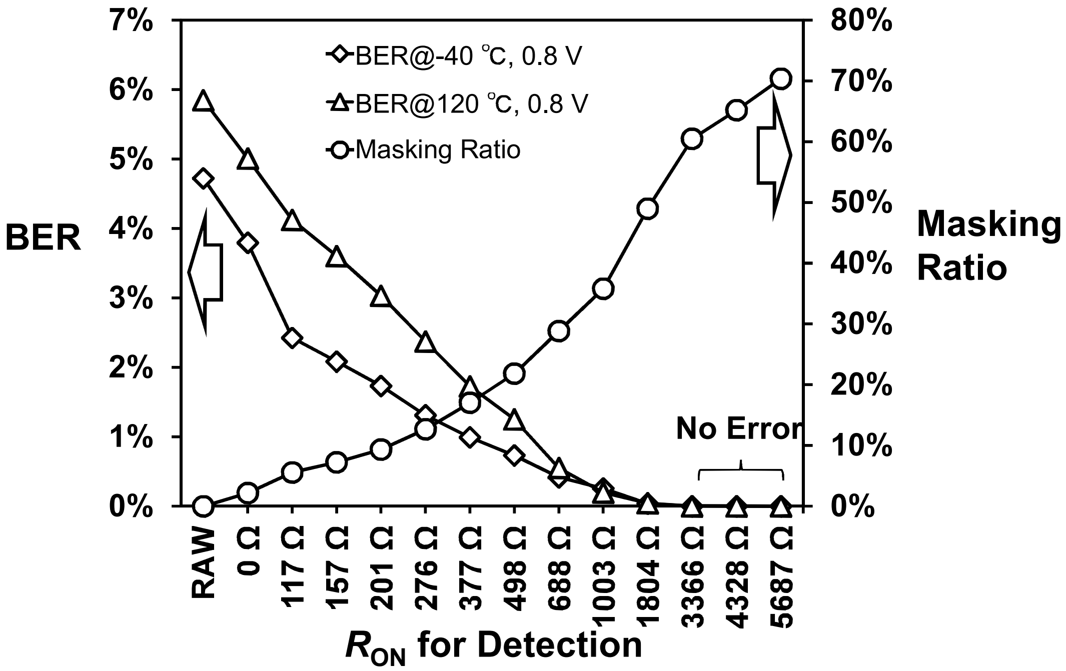


Figure 3.22: The average masking ratio as well as the BERs at two temperature corners with respect to the R_{ON}s used for detection.

Masking the same dark bits also effectively reduces the bit errors caused by temperature variation. The results are shown in Figure 3.21(a) and Figure 3.21(b). In the range of −20 °C to 60 °C, BER is reduced to less than 0.46% after 12.7% masking. As for a more extreme range of −40 °C to 120 °C, a 28.8% masking achieves more than 90.6% improvement, reducing the maximum average BER from 5.85% to 0.55%. After 60.5% masking, no bit errors appear in the measurement.

Figure 3.22 shows the relation between the R_{ON}s used for dark-bit detection and the masking ratios. The corresponding BERs at two temperature corners −40 °C and 120 °C are also depicted. It indicates that as R_{ON} increases, the masking ratio firmly increases, and the corresponding BERs decrease. When the masking ratio exceeds 60.5%, no errors are found at both the two temperature corners. The corresponding measured R_{ON} is 3366 Ω.

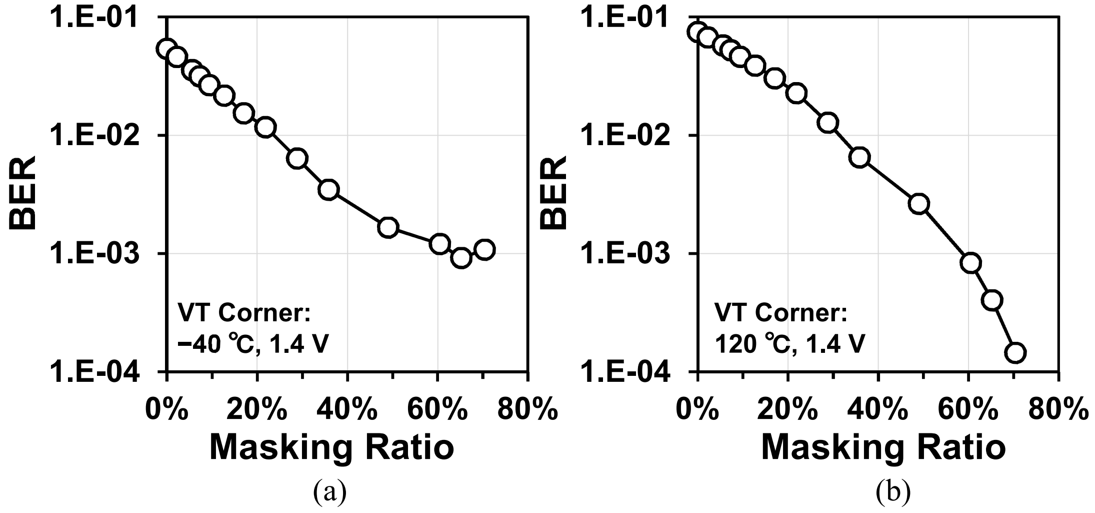
Table 3.3: Comparison of dark-bit masking effectiveness.

Room Temperature as Reference		Temperature			V_{DD}[4]
		80 °C	120 °C		1.4 V
This Work[1] V_{SS} Bias	Masking Ratio	21.8%	12.7%	60.5%	60.5%
	BER Improvement	87.9%	59.5%	100%[3]	99.3%
ISSCC'16 [67][2] Temperature Sweep	Masking Ratio	18.5%	–	–	–
	BER Improvement	90.6% (@85 °C)	–	–	–
ISSCC'17 [69][1] Body Bias	Masking Ratio	–	9.0%	–	–
	BER Improvement	–	60.0%	–	–

1. Dark-bit detection at the room temperature.
2. Requires temperature sweep (i.e., not dark-bit detection)
3. No error in 4048 bits, 500 evaluations.
4. Reference is the golden data at 0.8 V.

Table 3.3 shows the comparison of the proposed dark-bit detection with the previous dark-bit solutions. Compared with the conventional temperature sweep [67], this work achieves a comparable efficiency without requiring high test costs. Compared with the previous dark-bit detection using body bias [69], this work has a much wider detection range until BER approaches zero.

To further study the effectiveness of the dark-bit detection, BERs before and after masking at the two VT corners of −40 °C/1.4 V and 120 °C/ 1.4 V are also measured and analysed. Figure 3.23(a) shows the BERs at the −40 °C/1.4 V corner after masking the dark bits detected at the nominal VT condition. When the masking ratio reaches 70.4%, the BER remains at 0.11%. Figure 3.23(b) shows the BERs at the 120 °C/ 1.4 V corner. Similarly, after masking 70.4% bitcells, the BER fails to reach "zero" and stays at 0.01%. It reveals



(a) (b)

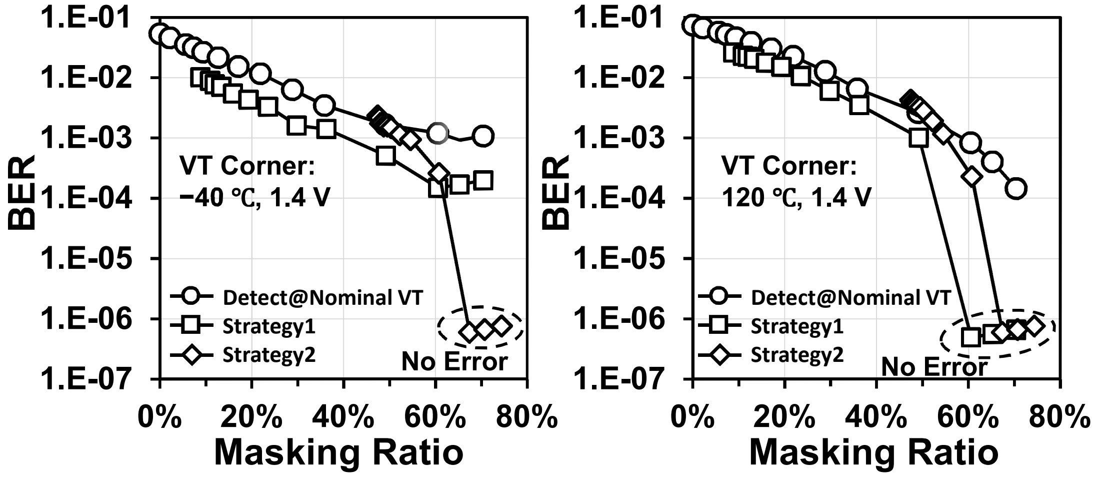
Figure 3.23: BERs at the VT corners of (a) −40 °C/1.4 V and (b) 120 °C/1.4 V. Dark bits are detected at the nominal VT condition of 23 °C/0.8 V.

that under the double effects of V_{DD} and temperature variations, even a bitcell with very large mismatch at the nominal condition could become unstable at the VT corners. To further reduce bit errors at the VT corners, new strategies are needed.

3.5.5 Dark-bit detection at nominal T condition

In order to efficiently mitigate bit errors at the VT corners, the detection strategy has been improved. The new method is to find out the bitcells sensitive to V_{DD} variations by performing dark-bit detection at an elevated V_{DD} condition. Compared with temperature sweep, V_{DD} sweep is easier in testing and does not lead to significantly higher cost in test equipment and test time. To efficiently find out V_{DD}-sensitive bitcells, the elevated V_{DD} is set as 1.6 V, which is 0.2 V higher than the corner V_{DD} of 1.4 V. Hereby, two strategies that combine the dark-bit detection at 0.8 V and the dark-bit detection at 1.6 V are proposed.

In the strategy 1, in addition to masking the bitcells detected by V_{SS} biasing at the nominal VT condition of 23 °C/0.8 V, those bitcells that flip at 23 °C/1.6 V are also masked. The results are shown with the lines with rectangles in Figure 3.24(a) and Figure 3.24(b). At



(a) (b)

Figure 3.24: BERs at the VT corners of (a) −40 °C/1.4 V and (b) 120 °C/1.4 V. Two new detection strategies are included.

both VT corners, BERs become lower with the same masking ratios, compared with the detection at the nominal VT. At the −40 °C/1.4 V corner, only one dark bit out of 10 K bits remains undetected, and the minimum BER is reduced from 0.09% to 0.01%. At the 120 °C/1.4 V corner, bit error is reduced to zero in the measurement.

In the strategy 2, the V_{SS} bias based dark-bit masking is also performed at 1.6 V using an R_{ON} of 276 Ω. Then, the bitcells detected at 1.6 V are masked together with the bitcells detected at 0.8 V. The results are shown with the lines with rhombuses in Figure 3.24(a) and Figure 3.24(b). Since the I_{SC} at 1.6 V is much larger than the I_{SC} at 0.8 V, the generated bias voltage becomes larger and the initial masking ratio reaches 47.4%. With this strategy, no error is observed at both corners after masking 67.4% bitcells. Based on the pessimistic assumption that one error will happen in the next evaluation, the corresponding BER is $<5.99\times10^{-7}$ (assuming one-bit error in 3339 bits × 501 evaluations).

Although 67.4% bitcells need to be masked, the effective bitcells in the PUF are still larger than 300. Compared with a typical key length of 128 or 256, this number is acceptable. In addition, thanks to the small bitcell area, the wasted area is limited. After 67.4% masking,

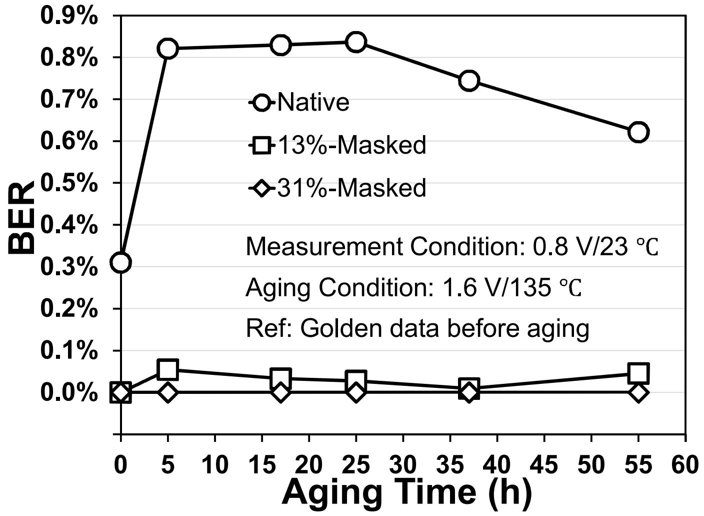


Figure 3.25: BERs with respect to the accelerated aging time.

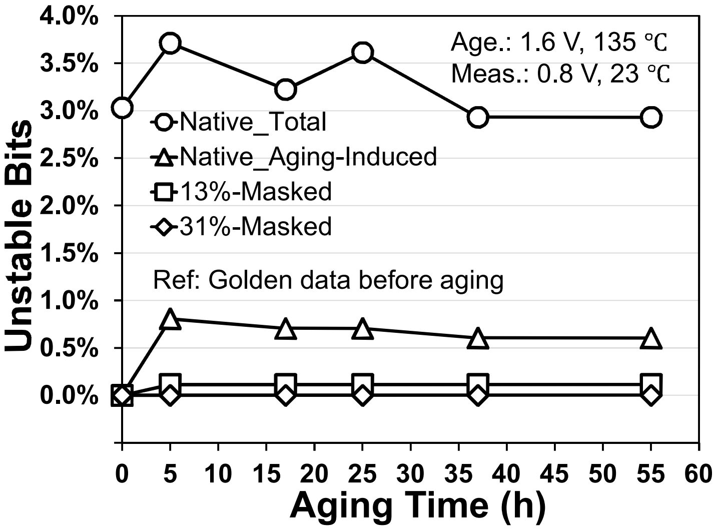


Figure 3.26: Percentage of unstable bits with respect to the accelerated aging time.

the area per bit is 1119 F^2, which is smaller than many of the most advanced PUFs, such as [66] and [73].

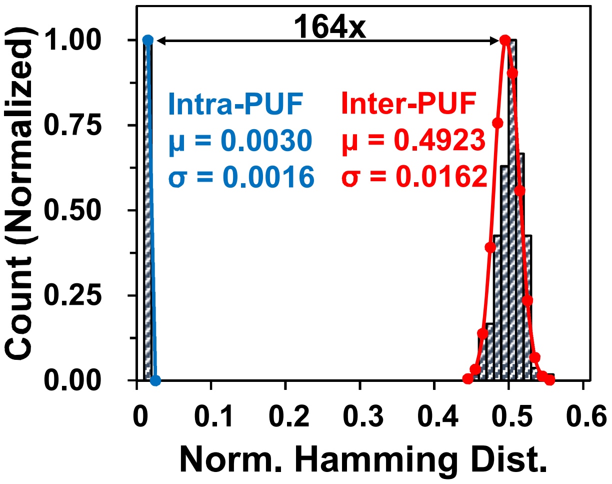


Figure 3.27: Measured inter-PUF and intra-PUF hamming distance of 20 chips.

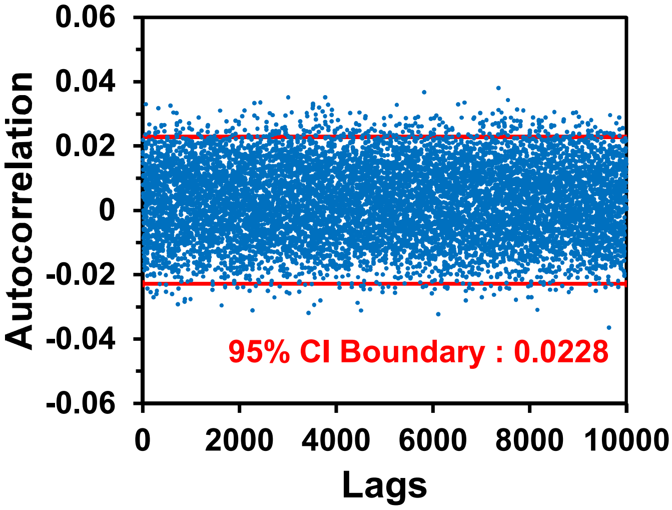


Figure 3.28: Measured autocorrelation among 20K bits.

3.5.6 Long-term reliability

To test the long-term reliability, an accelerated aging test is applied to one chip (1K bits). The test chip is baked in the temperature chamber, and it is kept evaluating and reading under a high temperature and high voltage condition for acceleration. The aging condition

Table 3.4: NIST SP 800-22 randomness tests.

Test Name	Masking Ratio	Stream Length	Runs. #	Avrg. P-Value	Pass?
Frequency	0%	1024	20	0.1868	Yes
	31.2%	704	10	0.3407	Yes
	47.9%	491	10	0.3302	Yes
	75.0%	256	10	0.5665	Yes
Block Frequency	0%	1024	20	0.4148	Yes
	31.2%	704	10	0.3109	Yes
	47.9%	491	10	0.5067	Yes
	75.0%	256	10	0.5944	Yes
Runs	0%	1024	20	0.4880	Yes
	31.2%	704	10	0.5401	Yes
	47.9%	491	10	0.4398	Yes
	75.0%	256	10	0.4259	Yes
Longest Runs of Ones	0%	1024	20	0.3269	Yes
	31.2%	704	10	0.5212	Yes
	47.9%	491	10	0.7648	Yes
	75.0%	256	10	0.5928	Yes
Cumulative Sums	0%	1024	20	0.2104	Yes
	31.2%	704	10	0.3480	Yes
	47.9%	491	10	0.4184	Yes
	75.0%	256	10	0.5647	Yes
FFT	0%	1024	20	0.6134	Yes
	31.2%	704	10	0.3914	Yes
	47.9%	491	10	0.6402	Yes
	75.0%	256	10	0.3751	Yes
Non-Overlapping Template Matching	0%	1024	20	0.5075	Yes
	31.2%	704	10	0.5084	Yes
	47.9%	491	10	0.4610	Yes
	75.0%	256	10	0.6154	Yes
Serial	0%	1024	20	0.4306	Yes
	31.2%	704	10	0.4334	Yes
	47.9%	491	10	0.3624	Yes
	75.0%	256	10	0.4990	Yes
Approximate Entropy	0%	1024	20	0.4224	Yes
	31.2%	704	10	0.5173	Yes
	47.9%	491	10	0.4169	Yes
	75.0%	256	10	0.6539	Yes

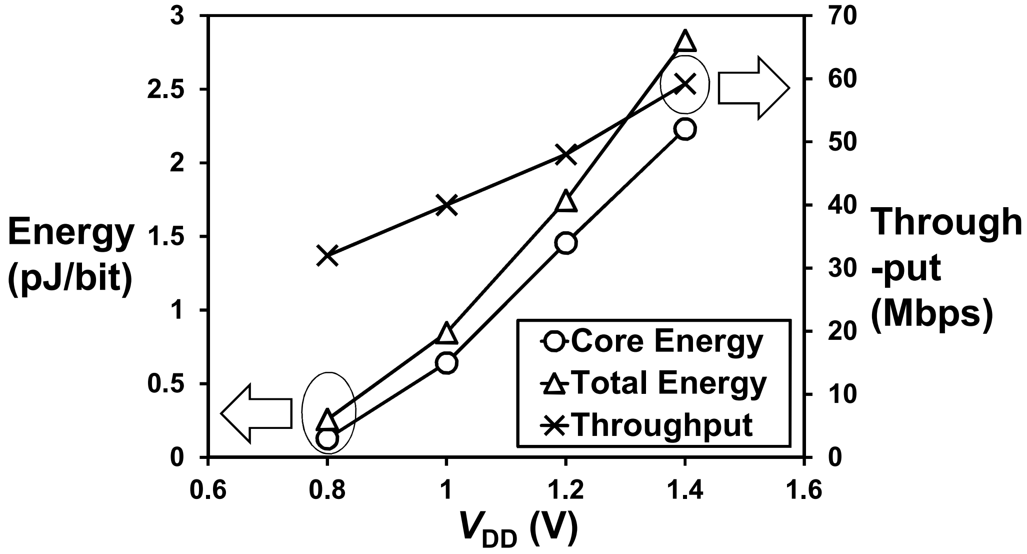


Figure 3.29: Measured energy and throughput in 0.8—1.4 V.

is 1.6 V and 135 °C, and the aging time is up to 55 hours. According to the VAF and TAF equations (2.3) and (2.4) in Chapter 2, the accelerated aging is equivalent to ~11 years of operation. The parameters are referred to [32], where E_a is 0.5 eV, γ is 2.6. The aging-induced BERs and instability are shown in Figure 3.25 and Figure 3.26, respectively. After aging, the BER is increased from 0.31% to 0.84% and the instability is up from 3.03% to 3.71%. However, after masking 13% bitcells which were detected before aging, BER and instability are reduced to 0.04% and 0.11%, respectively. After masking 31% bitcells, all aging induced bit errors are eliminated. Thus, dark-bit detection is effective on detecting the bitcells sensitive to aging.

3.5.7 **Uniqueness and randomness**

Uniqueness and randomness are basic requirements for a PUF. Uniqueness is evaluated with the inter-PUF hamming distance (HD_{Inter}). The results of hamming distance of twenty tested PUFs (i.e., 190 combinations) are shown in Figure 3.27. The average HD_{Inter} value is 0.4923, which is close to the ideal value of 0.5 and shows high uniqueness. Considering

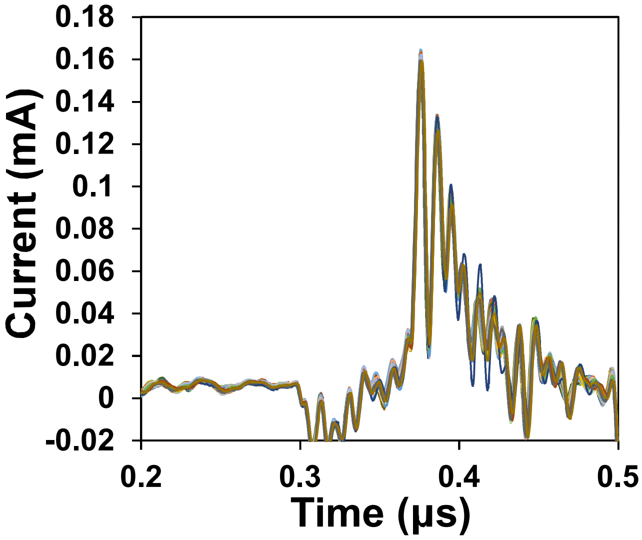


Figure 3.30: Current waveforms of the EE SRAM PUF. 64 waveforms corresponding to 64 words are shown. Each waveform is the average of 4K-time traces.

bit errors, the separation between HD_{Inter} and native HD_{Intra} of 164× shows that the PUF is highly identifiable even with bit errors.

Randomness is evaluated with the autocorrelation among bitcells calculated with equation (2.6) in Chapter 2. The result is shown in Figure 3.28. The 95% confidence interval (CI) boundary of 0.0228 is close to the ideal value of zero, showing that the correlation among between bitcells is very weak and the PUF has high randomness. Randomness is also verified with NIST SP 800-22 randomness tests. In this work, not only the data before masking are tested, but also the data after masking with the three typical masking ratios of 31.2%, 47.9%, and 75.0%. The results are shown in Table 3.4, and all the applicable tests are passed.

3.5.8 Energy and power

Measured energy and throughput in 0.8—1.4 V are shown in Figure 3.29. The core energy (i.e., the energy consumption of the bitcell array) is 0.128 pJ/bit at 0.8 V with a throughput

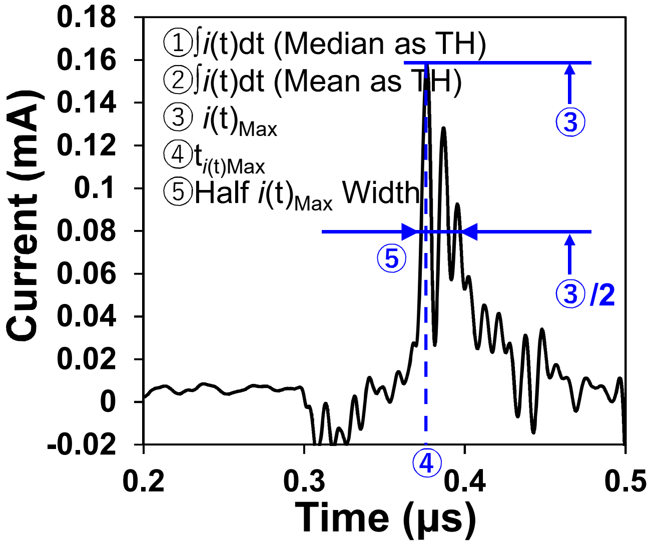


Figure 3.31: Indexes for SPA analysis.

of 32 Mbps. When the energy of peripheral circuits is counted, the total energy at 0.8 V is 0.258 pJ/bit. At the V_{DD} corner of 1.4 V, the core energy and the total energy are 2.23 pJ/bit and 2.83 pJ/bit with a throughput of 59.2 Mbps, respectively.

3.5.9 Side-channel attack analysis

To analyse the resilience to side-channel attacks, a simple power analysis (SPA) [73] is applied. To implement SPA, a 1-kΩ resistor is connected in series with the V_{DD}. Afterward, the PUF is kept evaluating and reading, and the power traces are monitored using a differential probe (Tektronix P6247, 1 GHz) and an oscilloscope (Tektronix MDO3102, 1 GHz, 5 GS/s) to measure the voltage drop across the resistor.

The EE SRAM PUF in this work has 64 words by 16 bits. For each word, it operates by 4K times. The average waveform of the 4K operation is used for analysis. By this mean, 64 waveforms are gotten, as shown in Figure 3.30. Analysing these 64 waveforms, 64 bits of binary number based on a threshold (TH) on several indexes are gotten. The 64-bit bitstream is then compared with the 64-bit binary numbers extracted based on the actual hamming weights of the 64-word PUF data. The correlation coefficients between them are

Table 3.5: Correlation between the SPA-extracted bitstreams and the actual bitstream based on the HW of PUF data.

Index Type	Binarized Bitstream	Correlation Coefficient
Actual HWs	01111111000111110010110000010101	
	01011100101101011101010001010000	
①Energy (Median as TH)	00111001001101011111111110101100	0.156
	10010100001100111100101000010000	
②Energy (Mean as TH)	00111001001101011111111110101100	0.156
	10010100001100111100101000010000	
③i_{Max} (Mean as TH)	11111011110110110011010000001000	−0.092
	10100110010100101000001000010011	
④t_{imax} (Mean as TH)	01000101100010100011111110111011	−0.040
	10110111111001111001001111110101	
⑤$W_{imax/2}$ (Mean as TH)	11101111011101011111111110111111	0.092
	11111110111011110110110101111000	

calculated following (3.12) to evaluate whether the SPA is successful or not, where A is the actual HW and B is the predicted HW using SPA. Figure 3.31 shows the indexes used for SPA.

$$Correlation\ Coefficient = \frac{\sum_{i=1}^{64}(A_i - \mu_A)(B_i - \mu_B)}{N \times \sigma_A \times \sigma_B} \qquad (3.12)$$

The results are listed in Table 3.5. It indicates that with all the indexes, the correlation coefficients are all close to the ideal value of zero. Also, these numbers are within the 95% CI boundary of random guess of 0.245. Therefore, the SPA cannot successfully predict the data of EE SRAM PUFs. It shows that the differential structure of EE SRAMs has a good resilience to power analysis attacks.

Table 3.6: Comparison table.

	This Work	ISSCC’18 [72]	JSSC’18 [66]	ISSCC’17 [69]	JSSC’17 [62]	ISSCC’16 [67]	JSSC’08 [56]
Technology	130nm	55nm	40nm	180nm	14nm	45nm	130nm
Mismatch Based?	YES	NO	YES	YES	YES	YES	YES
Type	SRAM	Anti-Fuse	RCCM	Amp. Chain	SRAM	NAND Chain	SRAM
Cell Area (F^2)	373	218	3644	553	9388	2613	1092
Native BER	0.21%	–[1]	3.2%[2]	0.13%	5.76%[2]	0.10%[3]	3.04%
Native Instability (Evaluations#)	2.14% (2000)	–[1]	2.55% (500)	1.67% (2000)	26.37% (5000)	–	–
Stabilizing Technique	V_{SS} Bias Dark-Bit Detection	Oxide Breakdown	Body Bias Compensation	Body Bias Dark-Bit Detection	TMV & NBTI & DB Masking	Glitch Detection & DB Masking	–
Measured V_{DD} (V)	0.8—1.4	0.75—1.35	0.8—1.0	0.8—1.8	0.55—0.75	–	0.9—1.2
Measured Temperature (°C)	−40—120	−40—150	−40—125	−40—120	25—110	−25—85	0—80
Stabilized BER at the Worst VT Corner	~0%[4]	~0%	2.31%	1.28%[5]	1.46%[6]	0.21%[7]	–
BER Improvement at the Worst VT Corner	~100%[4]	–[1]	6%	60%[5]	74.6%[6]	90%[7]	–
Core Energy (fJ/bit)	128	5200	1.2	11.3	4	–	930

[1] No PUF data before oxide breakdown.

[2] Including VT variations.

[3] After glitch detection, i.e., not native BER.

[4] No error in 3339 bits × 500 evaluations ($< 5.99\times10^{-7}$ BER). Masking ratio is 67.4%.

[5] V_{DD} variation is not included (i.e., not VT corner). Masking ratio is 9%.

[6] Including NBTI burn-in and TMV. Masking ratio is 20%.

[7] V_{DD} variation is not included. Temperature sweep is required. Masking ratio is 18.5%.

3.5.10 **Comparison**

Comparison with prior art is shown in Table 3.6. Compared with conventional SRAM PUFs [56, 62], this work achieves 14× lower BER and is comparable to the state-of-the-art mono-stable PUFs [69]. Among the mismatch-based PUFs, this work is the only one that achieves “zero” error at the VT corners after stabilization. Also, this work is the only design that achieves “zero” error with pure circuit techniques. In addition, the EE SRAM PUF has a compact bitcell that is only larger than a non-mismatch-based PUF using anti-fuses [72].

3.6 Conclusion

In this work, an EE SRAM PUF with a 2-D power gating technique and a V_{SS} bias-based dark-bit detection technique is designed in order to realize a “zero”-error, small-area, and low-energy SRAM PUF solution with pure circuit techniques.

First, high native stability is realized with an EE-structure bitcell. With a native BER of 0.21% and instability of 2.14%, it is the first SRAM PUF that achieves a low native BER that is comparable to the state-of-the-art mono-stable PUFs. Compared with conventional SRAM PUFs, its native BER is 14× lower. In addition, this nMOS-only bitcell has a small bitcell area of 373 F^2, thanks to the absence of p-n boundary.

Second, bit errors across −40—120 °C and 0.8—1.4 V are further lowered to “zero” by the post-processing technique of V_{SS} bias-based dark-bit detection and masking. This technique succeeds to find out all the dark bits across the measured VT range by simply applying V_{SS} bias voltages to the bitcells using a lightweight integrated resistance-based bias generator. This eliminates the need for costly temperature sweep in conventional testing. After masking the detected dark bits, the worst BER is reduced to be smaller than 5.99×10^{-7}. This technique is also the first solution to realize “zero” error based on only circuit techniques.

Third, the attack resilience is improved. A remanent charge clearance scheme clears the charges in the bitcells in every evaluation cycle, and therefore, eliminates the potential risk of side-channel attacks targeting the remanence decay. Also, the differential structure of SRAM bitcells is more resilient to power analysis attacks. Measurement results show that SPA fails to predict the PUF data.

Last but not least, the 2-D power gating technique powers up only one bitcell in a bitcell block and largely reduces the energy consumption to 128 fJ/bit.

Chapter 4
Hybrid SRAM PUF Using Hot Carrier Injection (HCI) Burn-in for Stability Reinforcement

4.1 Introduction

In Chapter 3, a non-ECC solution using a high-stability EE SRAM PUF and an efficient V_{SS} dark-bit detection was introduced. Although it achieves 5.99×10^{-7} BER with only circuit techniques, there are still a few demerits that could be improved. As for the PUF circuit, the energy consumption of 128 fJ/bit is larger than some of the most recent PUFs, which reach 1 fJ/bit to 10 fJ/bit. As for the dark-bit detection, it results in up to 67.4% bitcell loss. In addition, memories are required to store the helper data, resulting in increased costs.

In this chapter, I present a work with two main techniques to improve the above-mentioned weak points. First, a hybrid SRAM PUF is designed. This PUF lowers the minimum V_{DD} from the 0.8 V of EE SRAM PUF to 0.5 V, and therefore, reducing the energy by 62×, while keeping the high native stability feature. Second, the hybrid PUF is compatible with hot carrier injection (HCI) burn-in. Although HCI burn-in is not a pure circuit technique, it can solidly reduce the BER in a short time, and it neither requires helper data nor causes bitcell loss. Compared with NVMs and oxide breakdown, HCI does not cause visible damages or require special processes, and hence, it has a better security and lower cost.

The circuit design of the hybrid SRAM PUF is first introduced. Then, the implementation of HCI burn-in is shown. After that, the array architecture and operation are exhibited, followed by the experimental results and the analysis on the mismatch shift due to HCI. Finally, a conclusion is drawn.

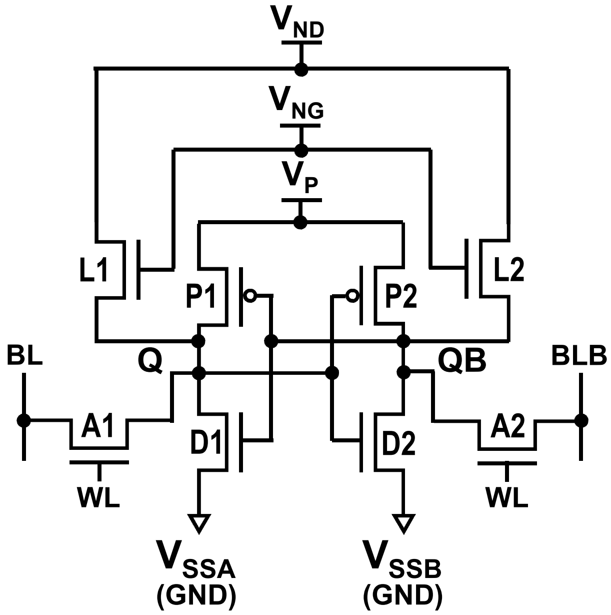


Figure 4.1: Circuit of the hybrid SRAM PUF cell.

4.2 Hybrid SRAM PUF

The circuit of the hybrid SRAM PUF cell is shown in Figure 4.1. It is composed of a 6T SRAM and two nMOS load transistors (i.e., L1 and L2) used in the EE SRAM PUF. It has two operation modes. One is the EE SRAM mode and the other is the CMOS SRAM mode. The operation mode is selected by the three voltages: V_{ND}, V_{NG}, and V_P. These three voltages are controlled through inverter-like drivers, which will be introduced in Section 4.4. The bitcells in different operation modes are shown in Figure 4.2. For high-stability evaluation, the hybrid PUF first works in the EE SRAM mode by setting V_P to the GND level and powering up V_{ND} and V_{NG}. In the EE SRAM mode, the data depend on the mismatch among the two load transistors and the two driver transistors. The two pMOS transistors have little influence on the evaluation, because their V_{GS}s are close to zero until V_Q and V_{QB} separate a lot. After that, V_P is turned on and the PUF works in an EE+CMOS intermediate state. Finally, V_{NG} and V_{ND} are turned off and the PUF works as a CMOS SRAM for low-power read/write operations.

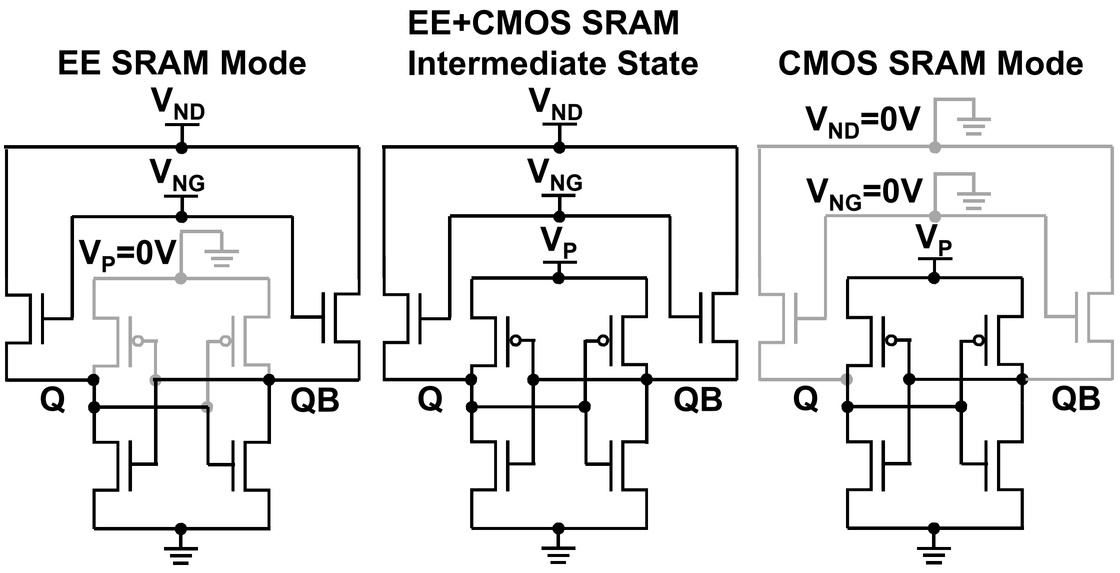


Figure 4.2: Hybrid PUF cells in different operation modes. The two access transistors are omitted.

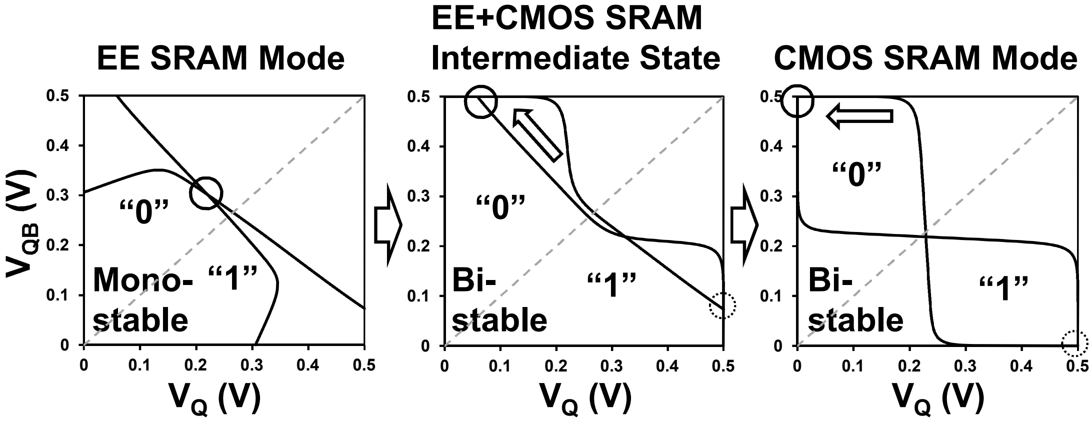


Figure 4.3: Butterfly curves of a hybrid SRAM cell in different operation modes at 0.5 V, simulated with a typical mismatch of 20 mV.

The butterfly curves of the hybrid PUF in these three states are shown in Figure 4.3. In the EE SRAM mode, the hybrid PUF stays in the mono-stable state just like an EE SRAM PUF. The small mismatch of the circuit is amplified, and the mono-stable solution moves

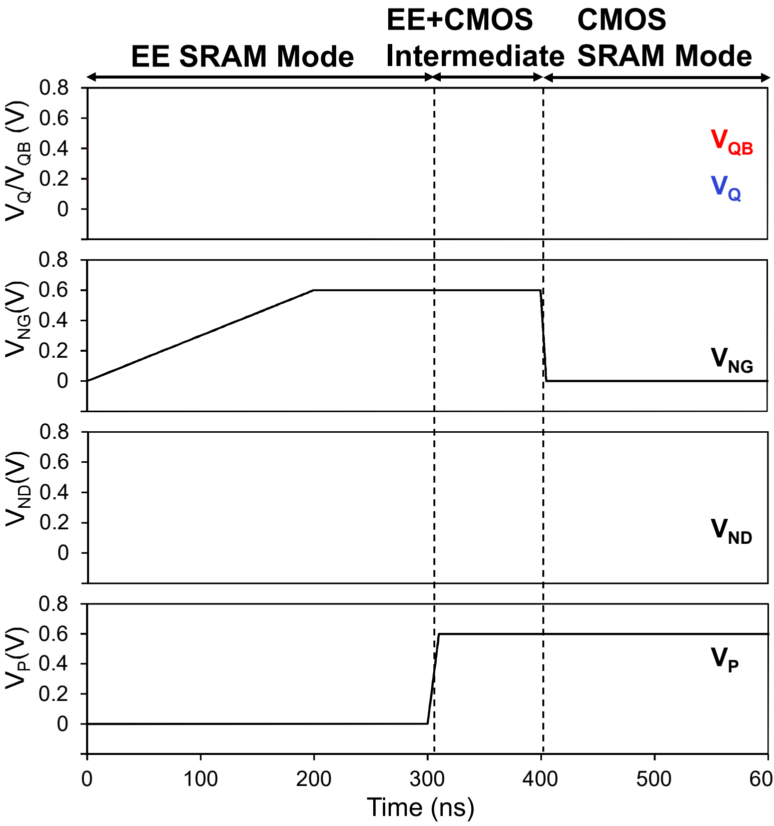


Figure 4.4: Transient simulation waveforms of V_Q, V_{QB}, V_{NG}, V_{ND}, and V_P during power-up evaluation.

away from the V_Q=V_{QB} line. This mono-stable state brings high stability. After powering up to a proper V_{DD}, e.g., 0.5 V, instead of continuing powering up the circuit, V_P turns on and the PUF goes into the intermediate state. In the intermediate state, the CMOS latch starts to function, biasing the PUF into the bi-stable state, and further amplifying the difference between V_Q and V_{QB}. Since the V_{DD} and the bitcell current are large enough under 0.5 V, the circuit is able to follow the solution change stably. Finally, the PUF works in the CMOS SRAM mode and a full-swing voltage is gotten. In the CMOS SRAM mode, there is no I_{SC} so that the energy consumption is reduced.

In the previous EE SRAM PUF, in order to separate V_Q and V_{QB}, a higher V_{DD} around 0.8 V is required. In contrast, in the hybrid SRAM PUF, the data latching scheme can separate

Table 4.1: Transistor size of the hybrid SRAM PUF cell.

Transistor	Width (μm)	Length (μm)
D1 & D2	4×W_{min}	0.13
L1 & L2	W_{min}	0.13
P1 & P2	1.00	0.13
A1 & A2	0.50	0.13

V_Q and V_{QB} even at 0.5 V. This reduces the minimum V_{DD} so that the energy consumption can be further reduced.

Figure 4.4 shows the waveforms of the transient simulation on the power-up evaluation. In the EE SRAM mode, V_Q and V_{QB} separates a little due to the mismatch. When the PUF goes into the EE+CMOS intermediate state, the separation between V_Q and V_{QB} are amplified. In the CMOS SRAM mode, they become a full-swing differential voltage.

The sizes of the transistors are listed in Table 4.1. The large width of the driver transistors D1 and D2 is for a higher gain in the EE SRAM mode as well as a better read-stability [94]. The large width of P1 and P2 is for the stable HCI burn-in, which will be introduced in Section 4.3. The access transistors A1 and A2 also have a large width to ensure the write stability.

4.3 Hot carrier injection burn-in stabilization

As introduced in Section 2.4.5, the reverse-direction HCI effect can efficiently increase the V_{th} of a transistor and is a good candidate for PUF stability reinforcement. In this work, the hybrid PUF is designed to be compatible with HCI burn-in. To verify the difference between forward-direction and reverse-direction HCI effect, a burn-in test on single transistors are performed. The results are shown in Figure 4.5. With a 10-min HCI stress, the V_{th} shift of the reverse direction is more than 2.7× compared with the forward direction. V_{th}s are measured by the extrapolation method using the I_{DS}-V_{GS} curves of the transistors. For a better burn-in efficiency, the reverse-direction HCI effect is chosen for this work.

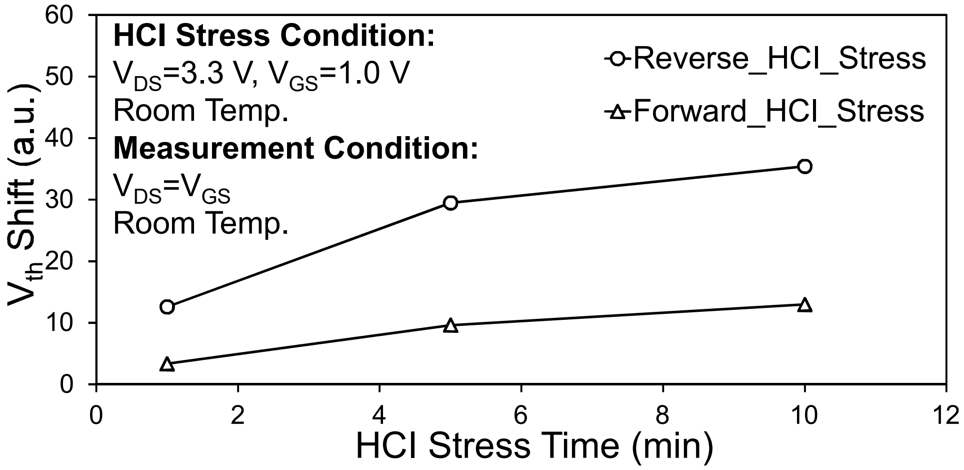


Figure 4.5: Measured V_{th} shifts caused by HCI effect. Both the forward- and reverse-direction stresses were tested.

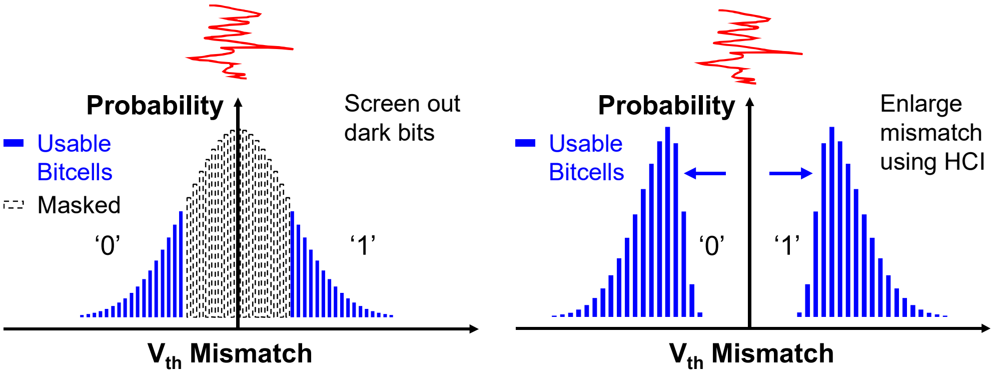


Figure 4.6: The concept of dark-bit masking and HCI burn-in.

Figure 4.6 shows the difference between dark-bit masking and HCI burn-in. Since the V_{th} mismatch follows the normal distribution, most bitcells have a relatively small V_{th} mismatch. Therefore, to ensure stability, a large portion of bitcells need to be masked, resulting in high bitcell loss. In contrast, HCI burn-in stabilization enlarges the V_{th} mismatch of every bitcells and shifts them into the large mismatch regions. In this case, all the bitcells could be stable and no bitcell need to be masked.

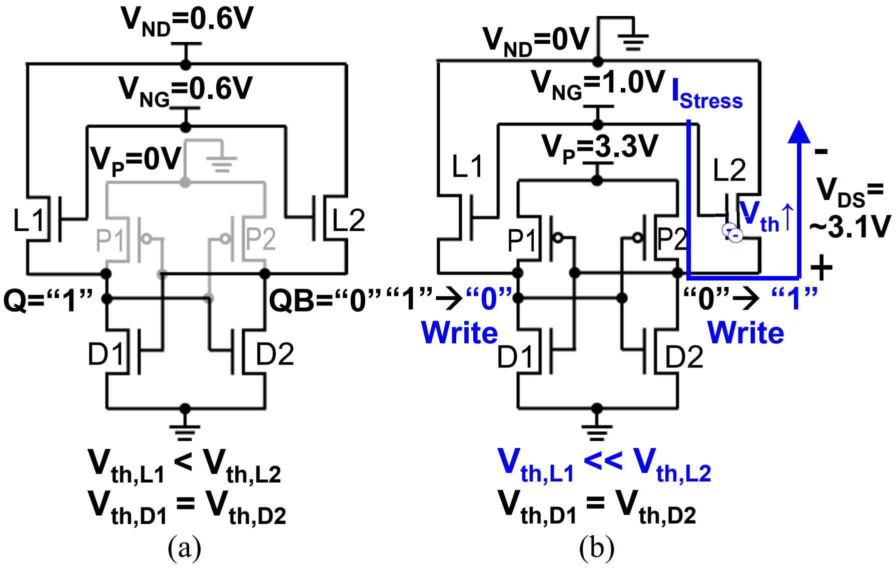
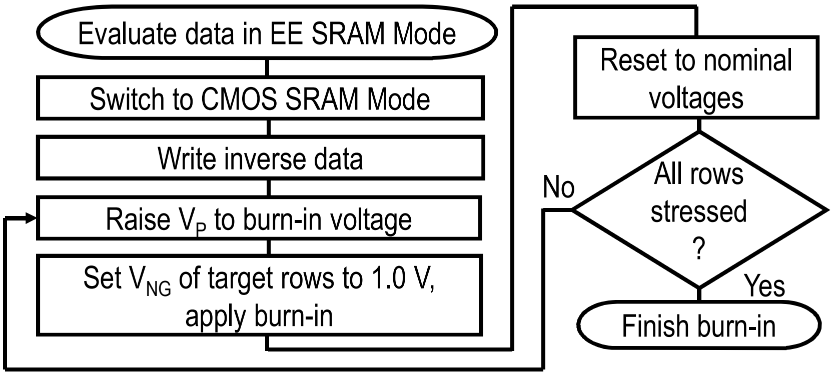


Figure 4.7: An example of HCI burn-in on a hybrid PUF cell. The access transistors are omitted. (a) Data development in EE SRAM mode; (b) HCI burn-in mode with inverse data.



Evaluate data in EE SRAM Mode
Switch to CMOS SRAM Mode
Write inverse data
Raise VP to burn-in voltage
Set VNG of target rows to 1.0 V, apply burn-in
Reset to nominal voltages
All rows stressed ?
No
Yes
Finish burn-in

Figure 4.8: Flowchart of HCI burn-in stabilization.

The implementation of HCI could be explained using the example in Figure 4.7(a) and Figure 4.7(b). Assuming that $V_{th,L1}$ is smaller than $V_{th,L2}$, and $V_{th,D1}$ nearly equals to $V_{th,D2}$, when the PUF evaluates in the EE SRAM mode, its data depend on the mismatch between $V_{th,L1}$ and $V_{th,L2}$ and should be Q equaling to "1" and QB equaling to "0". In this case, to reinforce the stability, the target is to increase $V_{th,L2}$ to further enlarge the V_{th} mismatch. It should be noted that this assumption is an ideal case, which makes the example easier to understand. In a practical case, $V_{th,D1}$ might not equal to $V_{th,D2}$. However, since $W_D L_D$ is four times larger than $W_L L_L$ in this design (see Table 4.1), the standard deviation of load transistor V_{th} ($\delta_{Vth,L}$) is two times larger than $\delta_{Vth,D}$ according to the Pelgrom's law [18]. Thus, this assumption is reasonable.

The flowchart of HCI burn-in is shown in Figure 4.8. First, the PUF is evaluated in the EE SRAM mode at the nominal VT condition. After that, it is switched to the CMOS SRAM mode, and the inverse data are written. V_P is then raised to a high voltage of 3.3 V, and the V_{NG}s of the target rows are set to 1.0 V. In the example in Figure 4.7, the target transistor L2 now has a large drain voltage at node QB, and it operates in the saturation region. In this case, L2 suffers from HCI stress, and the hot carriers are injected into its oxide near node QB. After that, the voltages are reset to the nominal, and these procedures are repeated until all the rows are stressed. After burn-in, when the PUF evaluates in the EE SRAM mode, the direction of channel currents of L2 is reversed from the burn-in direction. Therefore, the injected hot carriers of L2 now distribute at the source, and they largely increase $V_{th,L2}$. By this means, the mismatch is enlarged, and the stability is reinforced. The pMOS transistors act as the high voltage supplier to node Q and QB so that no additional transistors are needed in the bitcell.

In addition to HCI effect, NBTI and PBTI effects also exist during burn-in. In the example in Figure 4.7, P2 is stressed with NBTI and D1 is stressed with PBTI. Since the pMOS transistors have little influence on evaluation, the NBTI effect on P2 does not degrade stability. As for the PBTI on D1, it increases the $V_{th,D1}$, resulting in a stronger "1" at node Q, and reinforcing the stability instead of degrading it.

4.4 Array architecture and operation

Array architecture of the 1K-bit hybrid SRAM PUF is shown in Figure 4.9. The bit/word organization is 64 words by 16 bits and the bitcells are physically arranged with 32 rows and 32 columns. In overall, it is like the architecture of a conventional SRAM, but V_{ND}, V_{NG}, V_P drivers, and level shifters are added. The three drivers are used to control the

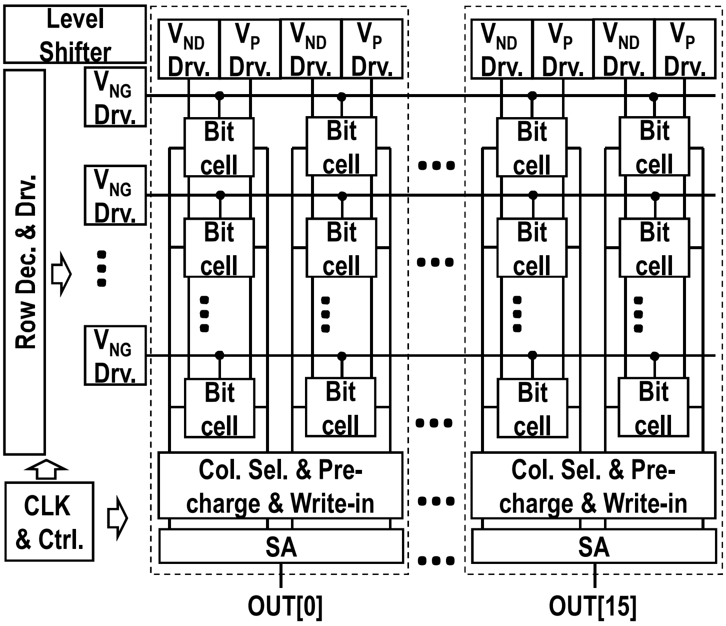


Figure 4.9: Array architecture of the hybrid SRAM PUF.

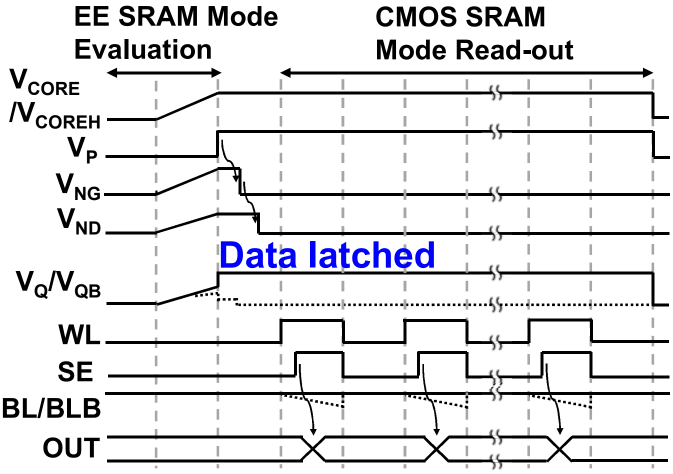


Figure 4.10: Operation waveforms of the hybrid SRAM PUF.

operation mode of the hybrid PUF. In order to apply high voltages for burn-in, only the V_P drivers and the level shifters are implemented with thick-oxide transistors, which are usually used for I/O circuits and are universally available in standard libraries.

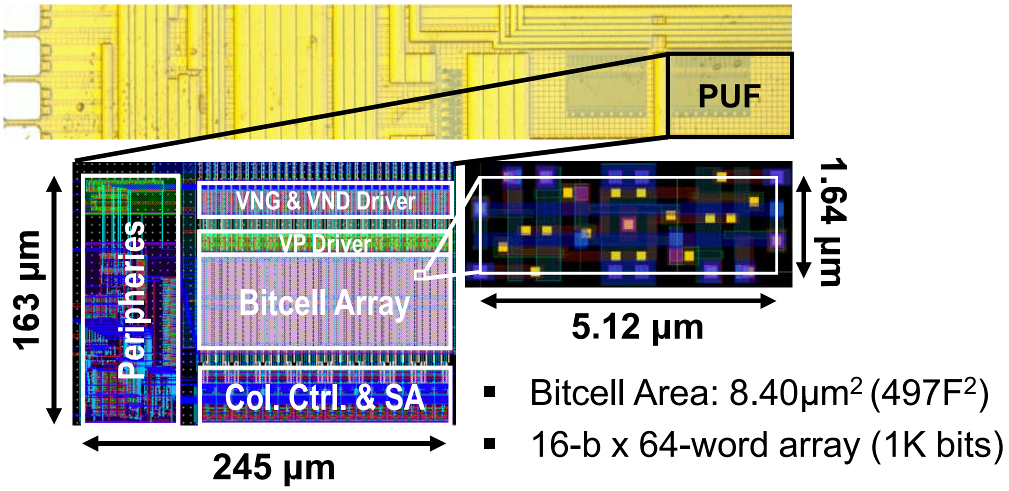


Figure 4.11: Micrograph of the prototype chip, the layout of the PUF array, and the layout of a bitcell.

The operation waveforms of one evaluation cycle are shown in Figure 4.10. The complementary sensing scheme is used for data read-out. It is considered to be more resilient to power analysis attacks as described in Chapter 3. After the data are latched in the CMOS SRAM mode, WL of the selected row is activated. According to the PUF data, either BL or BLB is discharged, and the difference between them is sensed by a StrongARM latch sense amplifier [95].

4.5 Experimental results

To verify the effectiveness of the design, test chips are fabricated in a 130-nm standard CMOS process. The micrograph of the chip, the layout of the PUF array, and the layout of a bitcell are shown in Figure 4.11. Since the hybrid PUF cell does not require additional transistors in the bitcell to apply HCI burn-in, it keeps a small area of 8.40 μm^2 and 497 F^2 in feature size. Due to the two additional pMOS transistors, the area overhead compared with an EE SRAM PUF bitcell is about 33%.

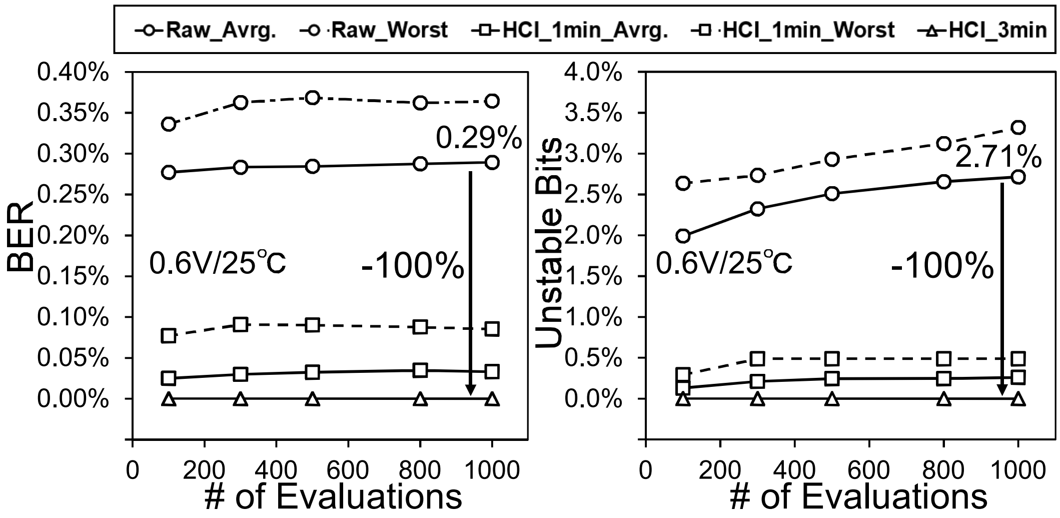


Figure 4.12: BER and instability before and after HCI burn-in.

4.5.1 **Stability at nominal condition**

The hybrid PUF has high native stability thanks to the evaluation in EE SRAM mode. Native stability (raw chips) and stability after HCI burn-in at the nominal condition of 0.6 V/25 °C are shown in Figure 4.12. Ten chips (10K bits) are measured for the native stability and five chips (5K bits) are measured with burn-in. With 1000 evaluations, the average native BER is 0.29% and the average native instability is 2.71%, which are similar to the EE SRAM PUF. As for the errors of the worst chip, its BER is 0.36% and its instability is 3.32% with 1000 evaluations.

After 1-min HCI stress, the average BER is reduced to 0.03% (89.6% reduction) and the average instability is reduced to 0.26% (90.4% reduction). The BER and instability of the worst chip are also reduced to 0.08% (77.8% reduction) and 0.49% (85.2% reduction), respectively. After 3-min HCI, both the BER and the instability at the nominal condition are reduced to zero in the measurement.

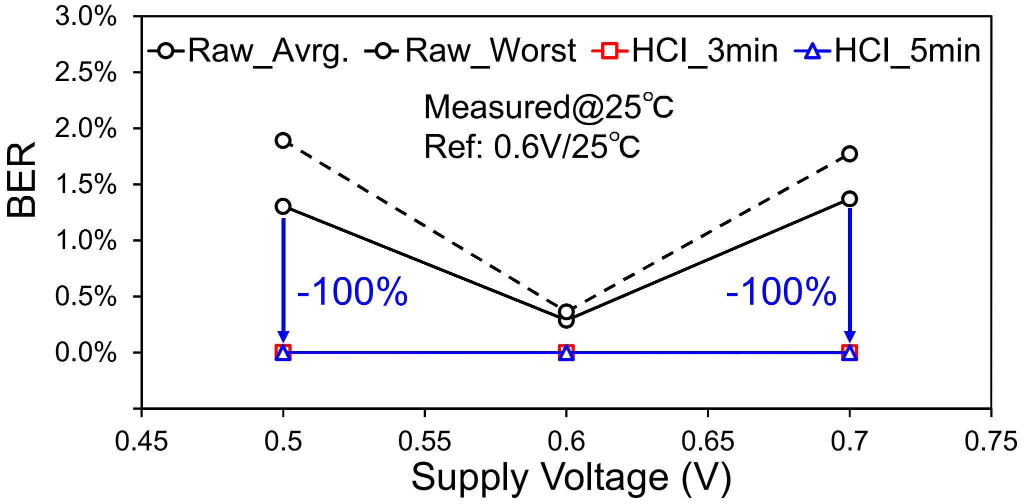


Figure 4.13: BERs across V_{DD} variations before and after HCI burn-in.

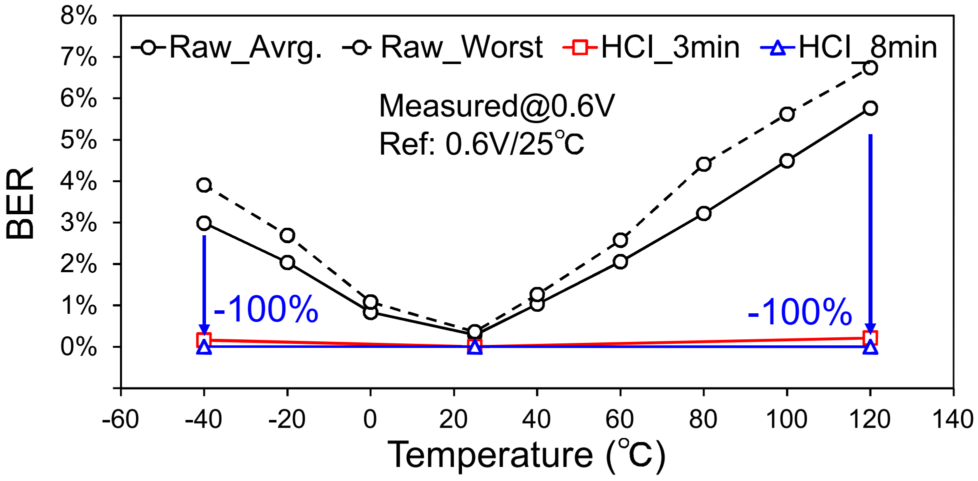


Figure 4.14: BERs across temperature variations before and after HCI burn-in.

4.5.2 **Stability across VT variations**

For stability across VT variations, five chips (5K bits) are measured with 500 evaluations at each VT condition. The test V_{DD} range is 0.5—0.7V, and the test temperature range is −40—120 °C.

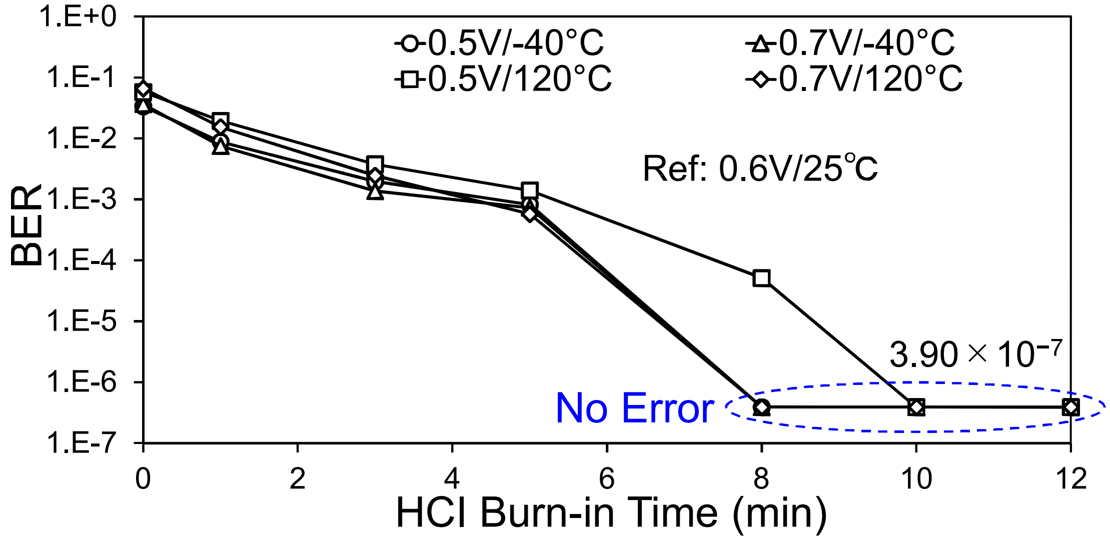


Figure 4.15: BERs at the four VT corners with respect to HCI burn-in time.

V_{DD} dependency across 0.5—0.7 V is shown in Figure 4.13. The BERs are calculated by comparing with the golden data gotten at the nominal condition of 0.6 V/25 ℃. As for native stability, at 0.5 V, the average BER is 1.30%, and the BER of the worst chip is 1.89%. At 0.7 V, the average BER is 1.37%, and the BER of the worst chip is 1.77%. After 3-min HCI, the average BER at 0.5 V is reduced to $4.88{\times}10^{-5}$ (> 99.99% reduction), and no bit errors are observed at 0.7 V. After 5-min HCI, no errors are observed at both the V_{DD} corners.

Temperature dependency across −40—120 ℃ is shown in Figure 4.14. Due to the wide test range, the maximum tested BER caused by temperature is higher than V_{DD}. At −40 ℃, the average BER is 2.99%, and the BER of the worst chip is 3.91%. At 120 ℃, the average BER of 5 chips is 5.76%, and the BER of the worst chip is 6.74%. After 3-min HCI, the BER at −40 ℃ is reduced to 0.16% (94.6% reduction), and the BER at 120 ℃ is reduced to 0.21% (96.4% reduction). After 8-min HCI, the BERs at both the temperature corners are reduced to zero in the measurement.

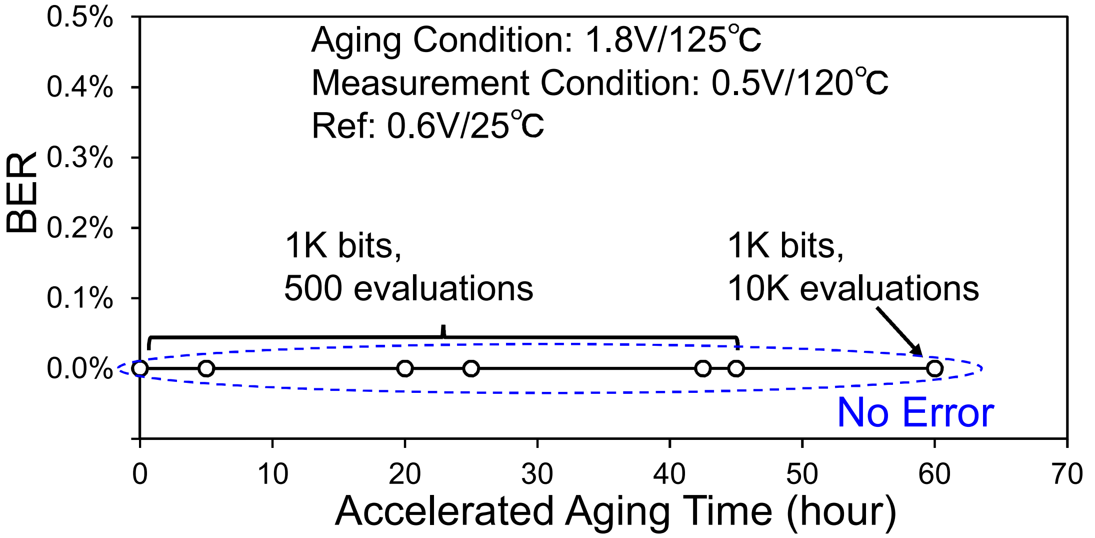


Figure 4.16: BERs with respect to accelerated aging time.

BERs at the four extreme VT corners of 0.5 V/−40 °C, 0.5 V/120 °C, 0.7 V/−40 °C, and 0.7 V/ 120 °C are also measured. Before HCI burn-in, the BER is up to 6.60% under the double effects of V_{DD} and temperature variations. But HCI burn-in solidly reduces the BERs. At the three corners of 0.5 V/−40 °C, 0.7 V/−40 °C, and 0.7 V/ 120 °C, bit errors are reduced to zero after 8-min HCI burn-in. The corresponding BER is 3.90×10^{-7}, based on the pessimistic assumption that one-bit error occurs in 5120 bits × 501 evaluations. Even at the worst corner of 0.5 V/120 °C, bit errors are reduced to zero after 10-min HCI.

4.5.3 Long-term reliability

To evaluate the long-term reliability of the hybrid SRAM PUF, an accelerated aging test is applied. One test chip with 1K bits is baked under 1.8 V and 125 °C for up to 60 hours. This is equivalent to approximately 21 years of operation, according to the assumption model in Section 2.1.3. During the aging test, the PUF is kept evaluating and reading. At the 60-h point, the 1K bitcells are tested with 10K evaluations. At other test points, they are tested with 500 evaluations. The measurement condition is the worst VT corner of 0.5 V and 120 °C.

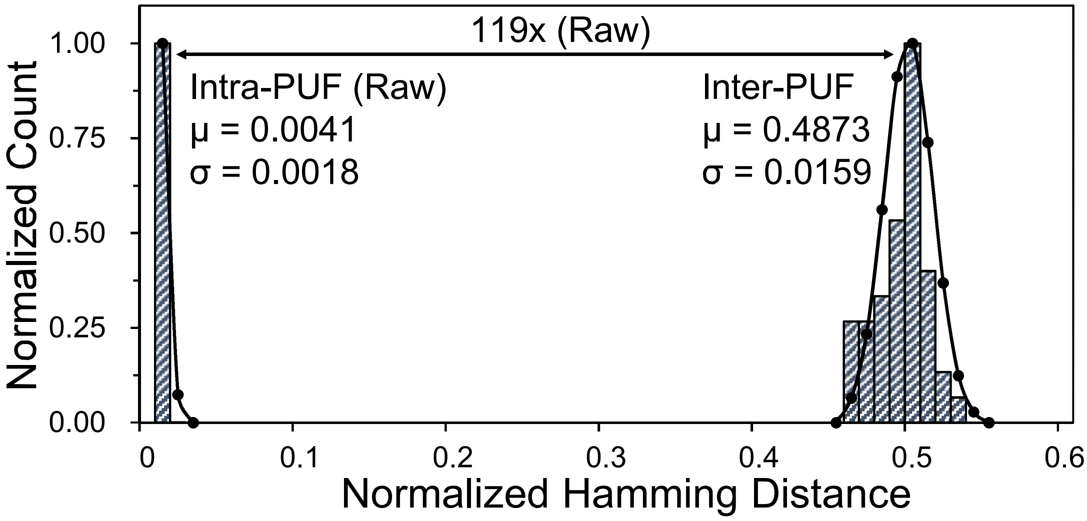


Figure 4.17: Measured native hamming distances.

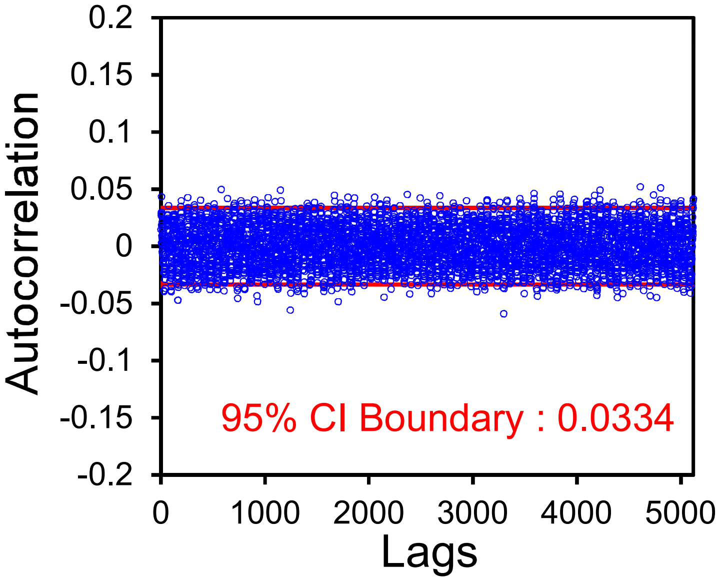


Figure 4.18: Measured autocorrelation.

The measurement results are shown in Figure 4.16. No errors are observed throughout the whole aging test. At the 60-h point, the corresponding BER is 9.76×10^{-8} under the

Table 4.2: NIST SP 800-22 randomness tests on the hybrid SRAM PUF.

Test Name	Stream Length	Runs. #	Pass Rate	Avrg. P-Value	Pass?
Frequency	1024	10	10/10	0.509	Yes
Block Frequency	1024	10	9/10	0.491	Yes
Runs	1024	10	10/10	0.428	Yes
Longest Run of Ones	1024	10	10/10	0.514	Yes
Cumulative Sums	1024	10	10/10	0.527	Yes
FFT	1024	10	10/10	0.449	Yes
Non-Overlapping Template Matching	1024 (m=5)	10	10/10	0.497	Yes
Serial	1024 (m=4)	10	10/10	0.489	Yes
Approximate Entropy	1024 (m=4)	10	9/10	0.344	Yes

pessimistic assumption that one error occurs at the next evaluation.

4.5.4 Uniqueness, identifiability, and randomness

Uniqueness, identifiability, and randomness are fundamental requirements for a PUF. The measured native HDs of ten chips (1K bits × 10) are shown in Figure 4.17. The average value of 0.4873 of HD_{Inter} shows that the hybrid PUF has good uniqueness. The 119× native separation between HD_{Inter} and HD_{Intra} indicates good identifiability considering bit errors.

Randomness is evaluated with the autocorrelation of bitcells. The measured autocorrelation is shown in Figure 4.18. The 95% CI boundary of 0.0334 is close to the ideal value of zero, showing high randomness. Randomness is also verified with NIST SP 800-22 randomness tests. The results are shown in Table 4.2. The hybrid PUF passes all the applicable tests, showing high randomness.

4.5.5 Energy efficiency and throughput

Thanks to the data latching scheme, the hybrid SRAM PUF can operate down to 0.5 V and realize low energy. For the energy measurement, a 10-Ω resistor is connected in series with

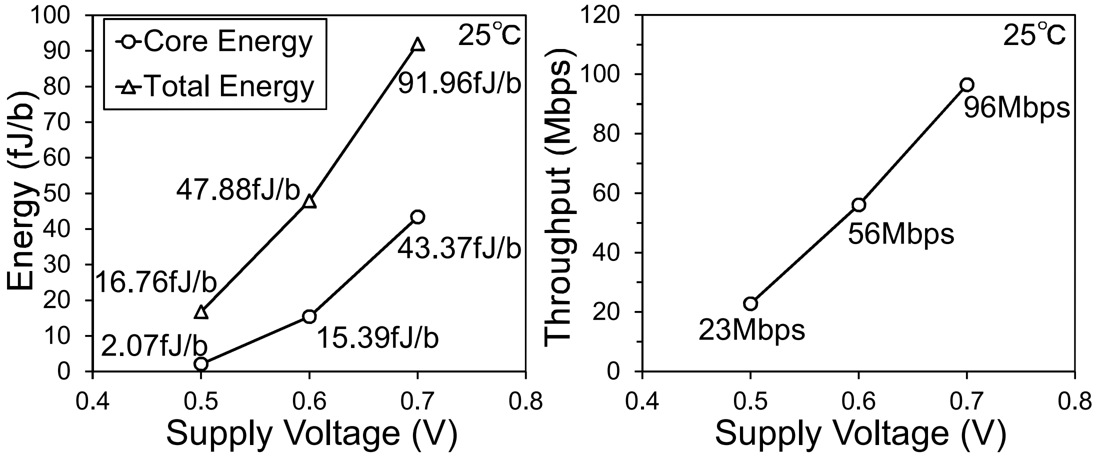


Figure 4.19: Energy and throughput in 0.5—0.7 V.

V_{DD}, and an oscilloscope (Tektronix MDO3102, 1 GHz, 5 GS/s) with a differential probe (Tektronix P6247, 1 GHz) is used to monitor the transient currents during evaluation.

The measured core energy (i.e., the energy of the bitcell array) and the total energy (i.e., the energy of the whole circuit including peripheries) are shown in Figure 4.19. The corresponding throughput is also depicted in Figure 4.19. At 0.5 V, the PUF achieves the best energy efficiency, with 2.07 fJ/bit core energy and 23 Mbps throughput. At the nominal V_{DD} of 0.6 V, the core energy is 15.39 fJ/bit with a throughput of 56 Mbps. The total energy at 0.5 V and 0.6 V are 16.76 fJ/bit and 47.88 fJ/bit, respectively.

4.5.6 Comparison with prior art

The comparison table is shown in Table 4.3. The hybrid PUF has high native stability that is comparable with mono-stable PUFs [71] and EE SRAM PUF [87] while achieving more than 61× lower energy than the EE SRAM. It also has a compact footprint of 497 F^2. The HCI burn-in is the only “zero” error solution that does not require helper data compared with other mismatch-based PUFs [68, 71, 87]. Compared with the oxide-breakdown [72, 73] or NVM [74] based designs, this work neither require visible oxide damages that might degrade security nor require additional fabrication processes that lead to higher cost.

Table 4.3: Comparison table.

	This Work	ISSCC'20 [68]	JSSC'20 [87]	JSSC'20 [71]	JSSC'19 [73]	ISSCC'19 [74]	ISSCC'18 [72]	CHES'13 [85]
Technology	130nm	28 nm	130 nm	65 nm	40 nm	130 nm	55 nm	65 nm
Mismatch Based?	YES	YES	YES	YES	NO	NO	NO	YES
Type	SRAM	NAND Chain	SRAM	Inverter Chain	Oxide Breakdown	NVM	Oxide Breakdown	SA
Stabilizing Technique	HCI Burn-in	F/F Fault Detection & DB Masking	Dark-Bit Detection & Masking	Dark-Bit Detection & Reconfig.	Soft Oxide Breakdown	ReRAM Storage	Hard Oxide Breakdown	HCI Burn-in
Without Visible Oxide Damages/Filaments?	YES	YES	YES	YES	NO	NO	NO	YES
Without Helper Data?	YES	NO	NO	NO	YES	YES	YES	YES
Cell Area (F^2)	497	3699	373[2]	562	1875	169	218	4926
Native BER	0.29%	0.089%[3]	0.21%	0.30%	–[1]	–[1]	–[1]	2.60%
Native Instability (Evaluations#)	2.71% (1000)	~25%[4]	2.14% (2000)	2.95% (2000)	–[1]	–[1]	–[1]	–
Measured V_{DD} (V)	0.8—1.4	0.81—0.99	0.8—1.4	0.7—1.4	0.9—1.5	1.4—2.2	0.75—1.35	0.8—1.2
Measured Temperature (°C)	−40—120	−40—150	−40—120	−55—125	−20—120	25—150	−40—150	−20—85
Stabilized BER at the Worst VT Corner	~0%	0.97%	~0%	2.3%[5]	~0%	~0%	~0%	~0%
BER Improvement at the Worst VT Corner	~100%	90.8%	~100%	45.6%	–[1]	–[1]	–[1]	~100%
Core Energy (fJ/bit)	2.07 (0.5 V) 15.39 (0.6 V)	2969	128	0.076	51.8	3028	5200	–

1. No PUF data before oxide breakdown or programming.
2. The area per cell is up to 1144 F^2, when the redundancy cost for dark-bit masking is considered.
3. After the flip-flop fault detection and dark-bit masking (i.e., not native).
4. Including 10% V_{DD} variations.
5. V_{DD} variation is not included (i.e., not VT corner).

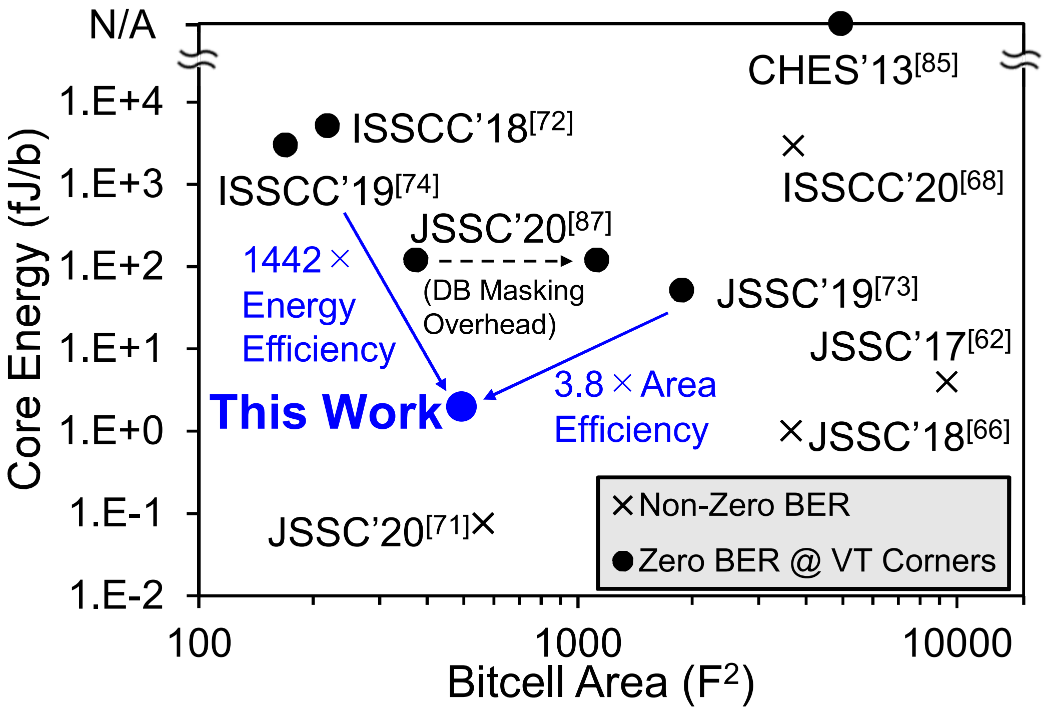


Figure 4.20: Comparison with prior art with respect to core energy and bitcell area.

Figure 4.20 shows the comparison with previous works with respect to core energy and bitcell area. Among all the "zero" error PUFs, the hybrid PUF has the lowest energy consumption. As for the area, it is only larger than two NVM-based solutions. Furthermore, HCI burn-in is the only solution that can achieve "zero" error without any helper data, visible damages, or additional fabrication processes.

4.6 Mismatch analysis

In this section, the mismatch shift caused by HCI burn-in is analyzed. The distributions of mismatch before and after HCI burn-in are measured at the nominal condition and the four VT corners to analyze the mismatch shift. One chip (1K bits) is measured and applied with HCI burn-in. To measure the mismatch, one of V_{SSA} and V_{SSB} is applied with a bias and the other is connected to the GND. A 10-Ω resistor is put between the GND and the bias voltage node to stabilize the bias voltage. The measurement setup is shown in Figure 4.21.

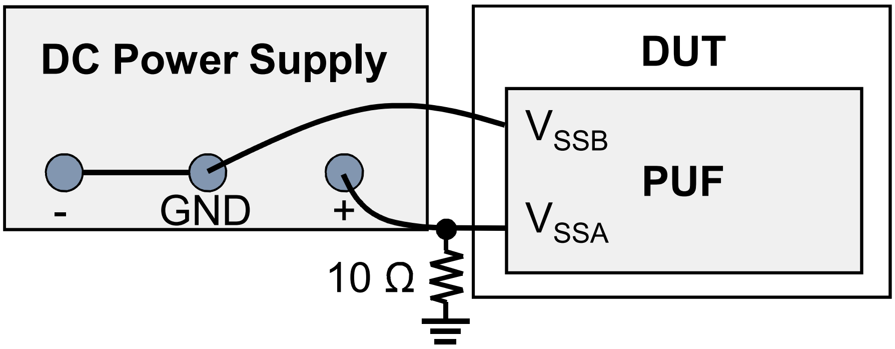


Figure 4.21: Measurement setup for mismatch shift analysis.

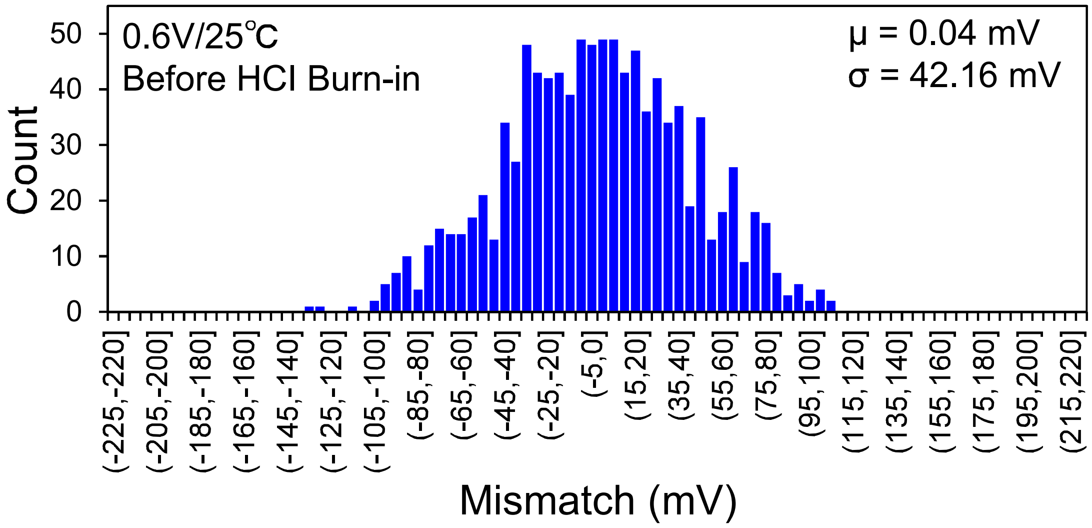


Figure 4.22: Measured mismatch distribution at the nominal condition of 0.6 V and 25 °C before burn-in.

The bias added to V_{SSA} is regarded as positive, and the bias added to V_{SSB} is regarded as negative. The bias voltage is added in a 5-mV step successively until all the bitcells are biased to "0" or "1'. The PUF is evaluated for 100 times at each bias condition, and its majority value is measured. When a bitcell's majority value flips at a bias voltage, its mismatch is regarded to be within the two bias voltages where the bitcell turns from its original value to the flipped value.

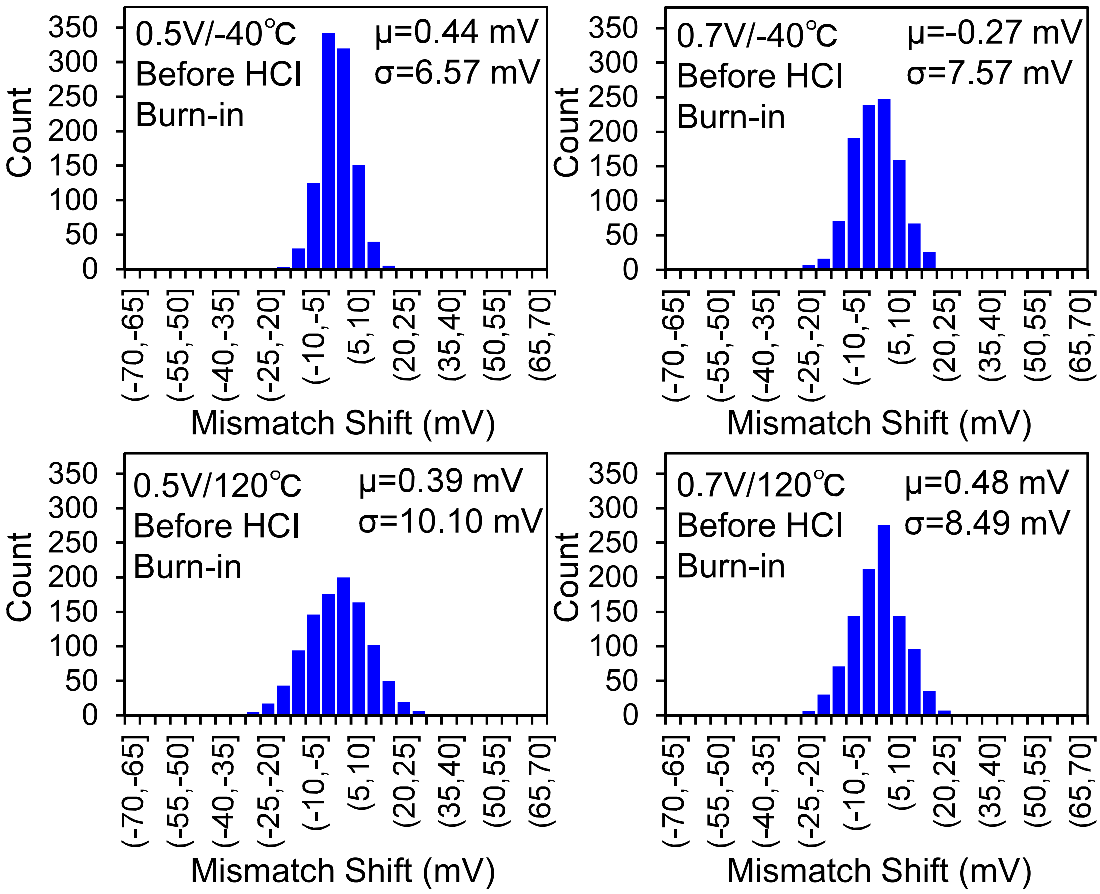


Figure 4.23: Mismatch shift caused by VT variations at the four VT corners.

Figure 4.22 shows the distribution at the nominal condition before burn-in. It shows that the mismatch distribution follows normal distribution, matching the theory. The standard deviation (δ) is 42.16 mV. Among all the bitcells, 9.47% of them fall in the −5- to 5-mV region. This is more than 3 times larger than the 2.71% unstable bits shown in Figure 4.12. Therefore, the bit error threshold should be smaller than 5 mV, and the target for burn-in is to enlarge the mismatch of all the bitcells to be larger than 5 mV at all the VT corners.

Figure 4.23 shows the distribution of mismatch shift caused by VT variations. If the shift is too large, a bitcell would change its mismatch polarity and generate erroneous data. The

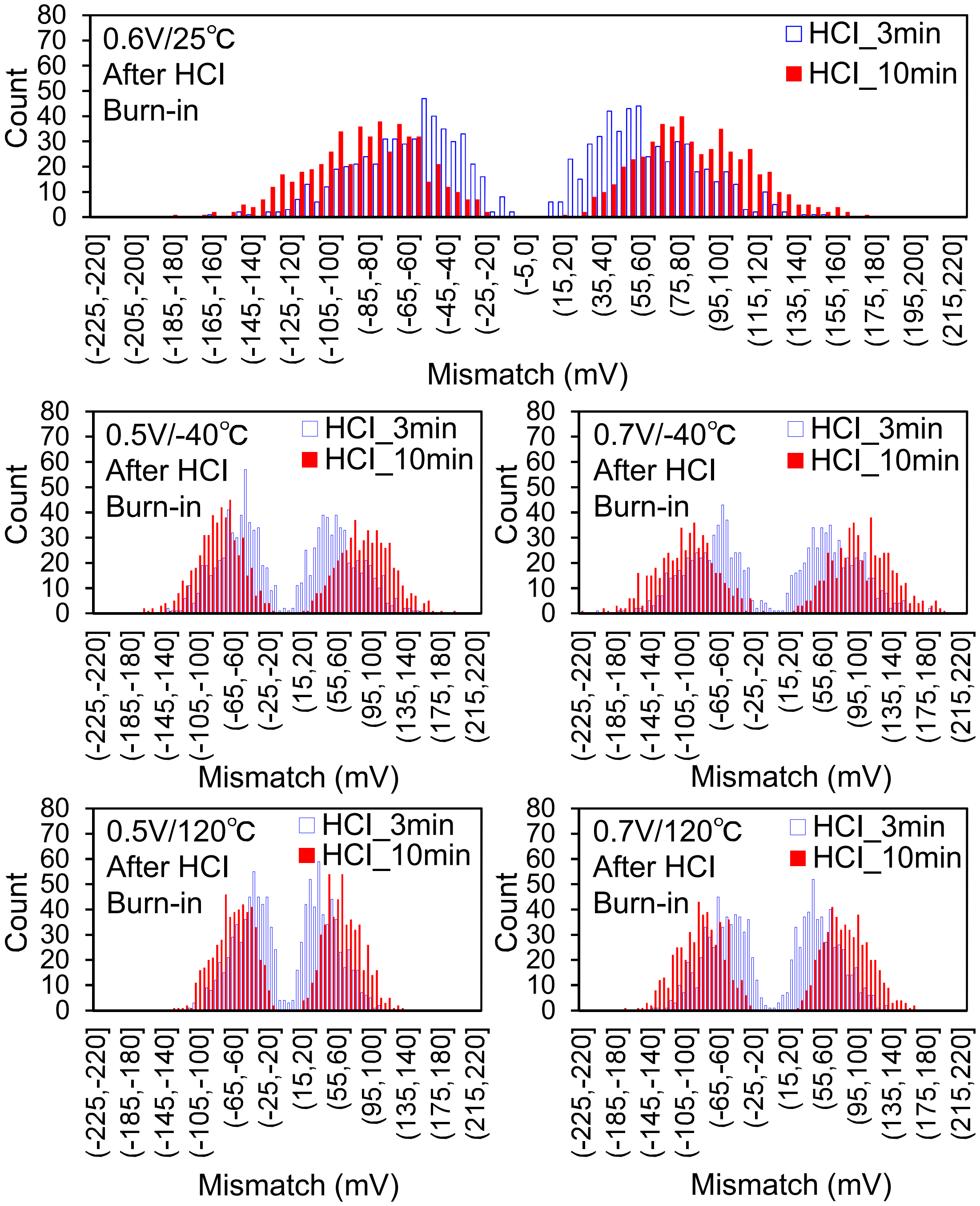


Figure 4.24: Mismatch distribution of hybrid SRAM PUF after 3-min and 10-min HCI burn-in.

δ at the worst VT corner of 0.5 V/120 °C is 10.10 mV, indicating that VT variations have a strong influence on the circuit mismatch.

Figure 4.24 shows the distribution of mismatch after 3-min and 10-min HCI burn-in. From the figure, it is known that HCI burn-in successfully enlarges the mismatch at all the VT conditions. After 3-min burn-in, no bitcells remain in the −5- to 10-mV region at the nominal condition. It complies with the results in Figure 4.12 that a 3-min HCI burn-in eliminates bit errors at the nominal VT. However, at the VT corners, there are still 1.2% of bitcells staying in the −5- to 5-mV region, indicating that 3 minutes are not enough to eliminate bit errors at the VT corners. After 10-min HCI burn-in, there is no bitcell staying in the −10- to 10-mV region even at all the four VT corners. This is two times of the 5-mV target and it shows the effectiveness of the HCI burn-in stabilization.

4.7 Conclusion

In this chapter, a hybrid SRAM PUF that is compatible with HCI burn-in stabilization is presented. Prototype chips in 130-nm CMOS are tested for metrics verification.

The hybrid SRAM PUF can operate in both the EE SRAM mode and the CMOS SRAM mode. During evaluation, it works in the EE SRAM mode to realize high stability. The measured BER and instability are 0.29% and 2.71%, respectively. The stability is comparable with the state-of-the-art mono-stable PUFs. After that, it switches to the CMOS SRAM mode to realize low-energy operation. The mode switching also lowers the operation V_{DD} from the 0.8 V of EE SRAM PUF in Chapter 3 to 0.5 V in this work, further reducing the energy consumption. The measured energy is 2.07 fJ/bit at 0.5 V, which is 61× smaller compared with EE SRAM PUFs. It also has a small footprint of 497 F^2, but this is 33% larger than the EE SRAM PUF in Chapter 3 due to additional pMOS transistors.

HCI burn-in succeeds to realize “zero” error. With 3-min burn-in, bit errors are eliminated a .6 V and 25 °C. After 10-min burn-in, all the bit errors are eliminated at all the VT corners in −40—120 °C and 0.5—0.7 V. With 12-min burn-in, no bit error is found in the accelerated aging test equivalent to ~21 years of operation.

The effectiveness of the burn-in is further studied. By mismatch analysis, it is found that after 10-min burn-in, no bitcell is in the −10—10-mV mismatch region, which is two times of my target.

Chapter 5
Conclusion

In this dissertation, two works using a combination of high-stability low-energy SRAM PUFs and efficient post-processing techniques to realize an error-free PUF solution have been presented. Results are all verified with test chips in 130-nm CMOS.

In Chapter 1, the weakness in conventional IoT authentication solution and the merits of the PUF-based authentication solution were introduced.

In Chapter 2, the metrics to evaluate a PUF, prior PUF circuits, and prior post-processing stabilization techniques were introduced.

In Chapter 3, a PUF solution achieving the "zero" error target with pure circuit techniques was presented.

A stable SRAM PUF, named EE SRAM PUF, is designed. The EE structure raises the transition V_{DD} from the mono-stable to bi-stable state and ensure a smooth switching between them. It is the first high-stability SRAM PUF, achieving 0.21% native BER.

A 2-D power gating technique reduces the energy consumption by about 64× to 128 fJ/bit by power gating the PUF cells in two dimensions. This also improves the security since PUF data only appear in the selected PUF cells, and the PUF data of other bitcells remain unknown.

A dark-bit detection technique using the V_{SS} bias is proposed. A lightweight bias generator that reuses the I_{SC} of EE SRAM cells is designed. It achieves "zero" error across all the VT conditions by masking 67.4% bitcells and "zero" error across the temperature conditions by masking 60.5% bitcells. This is the first error-free solution achieved with pure circuit techniques.

In Chapter 4, the hybrid SRAM PUF is designed to achieve a lower energy than the EE SRAM PUF in Chapter 3, and it takes an approach of device characteristics modification to achieve the "zero" error target.

The low energy feature is realized by a hybrid SRAM cell, which works not only in the EE SRAM mode for high stability, but also in the CMOS SRAM mode for low energy. Thanks to the mode switching, low-V_{DD} operation is achieved. Compared with the 0.8-V $V_{DD,min}$ for the EE SRAM PUF, the hybrid SRAM PUF lowers the operation V_{DD} down to 0.5 V. This again reduces the energy consumption. Measurement results show that the hybrid SRAM PUF has a minimum energy consumption of 2.07 fJ/bit, which is 61× lower than the EE SRAM PUF. Compared with previous "zero" error designs, the hybrid SRAM PUF has the lowest energy.

As for stability, the hybrid SRAM PUF has only 0.29% native BER, thanks to the EE SRAM mode operation. HCI burn-in is used to stabilize the PUF as post-processing. The hybrid SRAM PUF is compatible with HCI burn-in without additional transistors in the bitcell. After 10-min HCI burn-in, bit errors in all the VT corners in the range of −40—120 °C and 0.5—0.7 V are eliminated. After 12-min burn-in, no bit errors occur in the accelerated aging test that is equivalent to about 21 years of operation, even operating at the worst VT corner. HCI burn-in has two main advantages. First, it does not result in bitcell loss. Second, it does not require memories to store helper data. Also, HCI does not cause visible damages, which might lead to security risks. Compared with other post-processing techniques that realize "zero" error, HCI is the only solution that does not require any of helper data, additional fabrication processes or visible damages.

BIBLIOGRAPHY

[1] H. Tankovsa, "Internet of Things - active connections worldwide 2015-2025," Statista Research Department, Sep. 2020. [Online]. Available: https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/. [Accessed Dec. 2020].

[2] A. Mosenia and N. K. Jha, "A comprehensive study of security of Internet-of-Things," IEEE Trans. Emerging Topics Comput., vol. 5, no. 4, pp. 586-602, Oct.-Dec. 2016.

[3] I. Verbauwhede, "Security adds an extra dimension to IC design: Future IC design must focus on security in addition to low power and energy," IEEE Solid-State Circuits Mag., vol. 9, no. 4, pp. 41-45, Nov. 2017.

[4] M. Alioto, "Trends in hardware security: From basics to ASICs," IEEE Solid-State Circuits Mag., vol. 11, no. 3, pp. 56-74, Aug. 2019.

[5] J. Delvaux, R. Peeters, D. Gu and I. Verbauwhede, "A survey on lightweight entity authentication with Strong PUFs," ACM Comput. Surveys, vol. 48, no. 2, pp. 26:1-26:42, Oct. 2015.

[6] National Institute of Standards and Technologies (NIST), "Advanced Encryption Standard (AES)," Nov. 2001. [Online]. Available: https://www.nist.gov/publications/advanced-encryption-standard-aes. [Accessed Dec. 2020].

[7] J. Borghoff, A. Canteaut, T. Güneysu, E. B. Kavun, M. Knezevic, L. R. Knudsen, G. Leander, V. Nikov, C. Paar, C. Rechberger, P. Rombouts, S. S. Thomsen and T. Yalçın, "PRINCE – A low-latency block cipher for pervasive computing applications," in Proc. IACR Int. Conf. Theory Appl. Cryptogr. Inf. Secur. (ASIACRYPT), Dec. 2012, pp. 208-225.

[8] R. Beaulieu, D. Shors, J. Smith, S. Treatman-Clark, B. Weeks and L. Wingers, "The SIMON and SPECK families of lightweight block ciphers," Jun. 2013. [Online]. Available: https://eprint.iacr.org/2013/404.pdf.

[9] T. Suzaki, K. Minematsu, S. Morioka and E. Kobayashi, "TWINE: A lightweight, versatile block cipher," Jan. 2011. [Online]. Available: https://www.nec.com/en/global/rd/tg/code/symenc/pdf/twine_LC11.pdf. [Accessed Jan. 2021].

[10] O. Kömmerling and M. G. Kuhn, "Design principles for tamper-resistant smartcard processors," in Proc. USENIX Workshop Smartcard Technol. (WOST), May 1999, pp. 1-13.

[11] F. Courbon, S. Skorobogatov and C. Woods, "Reverse engineering Flash EEPROM memories using Scanning Electron Microscopy," in Proc. IFIP Int. Conf. Smart Card Res. Adv. Appl. (CARDIS), Nov. 2016, pp. 57-72.

[12] S. Skorobogatov, "Local heating attacks on Flash memory devices," in Proc. IEEE Int. Workshop Hardw.-Oriented Secur. Trust (HOST), Jul. 2009, pp. 1-6.

[13] S. Skorobogatov, "Flash memory 'bumping' attacks," in Proc. IACR Int. Conf. Cryptogr. Hardw. Embeded Syst. (CHES), Aug. 2010, pp. 158-172.

[14] S. Skorobogatov, "How microprobing can attack encrypted memory," in Proc. Euromicro Conf. Digit. Syst. Design (DSD), Aug. 2017, pp. 244-251.

[15] Intrinsic ID, "Protecting the IoT with invisible keys | White Paper," 2019. [Online]. Available: https://www.intrinsic-id.com/resources/white-papers/protecting-iot-invisible-keys-white-paper/. [Accessed Dec. 2020].

[16] C. Böhm and M. Hofer, Physical Unclonable Functions in Theory and Practice, 1st ed., New York, NY, USA: Springer-Verlag, 2013.

[17] R. Maes, Physically Unclonable Functions, 1st ed., Berlin, Germany: Springer-Verlag, 2013.

[18] M. J. M. Pelgrom, A. C. J. Duinmaijer and A. P. G. Welbers, "Matching properties of MOS transistors," IEEE J. Solid-State Circuits, vol. 24, no. 5, pp. 1433-1440, Oct. 1989.

[19] A. Sheikholeslami, "Process variation and Pelgrom's law [Circuit Instuitions]," IEEE Solid-State Circuits Mag., vol. 7, no. 1, pp. 8-9, Feb. 2015.

[20] Y. Chen, "ReRAM: History, status, and future," IEEE Trans. Electron Devices, vol. 67, no. 4, pp. 1420-1433, Apr. 2020.

[21] S.-Y. Chou, Y.-S. Chen, J.-H. Chang, Y.-D. Chih and T.-Y. J. Chang, "11.3 A 10nm 32Kb low-voltage logic-compatible anti-fuse one-time-programmable memory with anti-tampering sensing scheme," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2017, pp. 200-201.

[22] K. Yang, D. Blaauw and D. Sylvester, "Hardware designs for security in ultra-low-power IoT systems: An overview and survey," IEEE Micro, vol. 37, no. 6, pp. 72-89, Nov. 2017.

[23] U. Rührmair, F. Sehnke, J. Sölter, G. Dror, S. Devadas and J. Schmidhuber, "Modeling attacks on physical unclonable functions," in Proc. ACM Conf. Comput. Commun. Secur. (CCS), Oct. 2010, pp. 237-249.

[24] G. T. Becker, "On the pitfalls of using arbiter-PUFs as building blocks," IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 34, no. 8, pp. 1295-1307, Aug. 2015.

[25] Intrinsic ID, "SRAM PUF: The secure silicon fingerprint," 2020. [Online]. Available: https://www.intrinsic-id.com/wp-content/uploads/2020/08/sram-puf-secure-silicon-fingerprint-white-paper.pdf. [Accessed Dec. 2020].

[26] Maxim Integrated, "ChipDNA embeded security PUF technology," 2020. [Online]. Available: https://www.maximintegrated.com/en/design/partners-and-technology/design-technology/chipdna-puf-technology.html. [Accessed Dec. 2020].

[27] T. Lu, R. Kenny and S. Atsatt, "Secure device manager for Intel Stratix 10 devices provides FPGA and SoC security," 2020. [Online]. Available: https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/wp/wp-01252-secure-device-manager-for-fpga-soc-security.pdf. [Accessed Dec. 2020].

[28] Samsung Electronics, "Embedded security keeps mobiles safe," 2020. [Online]. Available: https://www.samsung.com/semiconductor/security/ese/. [Accessed Dec. 2020].

[29] W. Che, M. Martin, G. Pocklassery, V. K. Kajuluri, F. Saqib and J. Plusquellic, "A privacy-preserving, mutual PUF-based authentication protocol," Cryptography, vol. 1, no. 1, p. 3, 2017.

[30] S. K. Satpathy, S. K. Mathew, R. Kumar, V. Suresh, M. A. Anders, H. Kaul, A. Agarwal, S. Hsu, R. K. Krishnamurthy and V. De, "An all-digital unified physically unclonable function and true random number generator featuring self-calibrating hierarchical von Neumann extraction in 14-nm tri-gate CMOS," IEEE J. Solid-State Circuits, vol. 54, no. 4, pp. 1074-1085, Apr. 2019.

[31] M. Alioto and S. Taneja, "Enabling ubiquitous hardware security via energy-efficient primitives and systems," in Proc. IEEE Custom Integr. Circuits Conf. (CICC), Apr. 2019, pp. 1-8.

[32] R. Maes, V. Rožić, I. Verbauwhede, P. Koeberl, E. van der Sluis and V. van der Leest, "Experimental evaluation of physically unclonable functions in 65 nm CMOS," in Proc. Eur. Solid-State Circuits Conf. (ESSCIRC), Sep. 2012, pp. 486-489.

[33] National Institute of Standards and Technology, "NIST SP 800-22: Download Documents and Softwares," 2018. [Online]. Available: https://csrc.nist.gov/projects/random-bit-generation/documentationandsoftware.

[34] D. Nedospasov, J.-P. Seifert, C. Helfmeier and C. Boit, "Invasive PUF analysis," in Proc. Workshop Fault Diagnosis Tolerance Cryptogr. (FDTC), Aug. 2013, pp. 30-38.

[35] S. Zeitouni, Y. Oren, C. Wachsmann, P. Koeberl and A.-R. Sadeghi, "Remanence decay side-channel: The PUF case," IEEE Trans. Inf. Forensics Secur., vol. 11, no. 6, pp. 1106-1116, Jun. 2016.

[36] P. C. Kocher, J. Jaffe and B. Jun, "Differential power analysis," in Proc. Annu. Int. Cryptology Conf. Advances Cryptology (CRYPTO), Aug. 1999, pp. 388-397.

[37] T. S. Messerges, E. A. Dabbish and R. H. Sloan, "Examining smart-card security under the threat of power analysis attacks," IEEE Trans. Comput., vol. 51, no. 5, pp. 541-552, May 2002.

[38] R. W. Hamming, "Error detecting and error correcting codes," The Bell Syst. Tech. J., vol. 29, no. 2, pp. 147-160, Apr. 1950.

[39] A. Hocquenghem, "Codes correcteurs d'erreurs," Chiffres, vol. 2, no. 2, pp. 147-156, Sep. 1959.

[40] R. C. Bose and D. K. Ray-Chaudhuri, "On a class of error correcting binary group codes," Inf. Control, vol. 3, no. 1, pp. 68-79, Mar. 1960.

[41] D. E. Muller, "Application of Boolean algebra to switching circuit design and to error detection," Trans. IRE Prof. Group Electron. Comput., vol. 3, no. 3, pp. 6-12, Sep. 1954.

[42] I. S. Reed, "A class of multiple-error-correcting codes and the decoding scheme," Trans. IRE Prof. Group Inf. Theory, vol. 4, no. 4, pp. 38-49, Sep. 1954.

[43] M. Hiller and A. G. Önalan, "Hiding secrecy leakage in leaky helper data," in Proc. IACR Int. Conf. Cyptogr. Hardw. Embeded Syst. (CHES), Sep. 2017, pp. 601-619.

[44] J. Dai and L. Wang, "A study of side-channel effects in reliability-enhancing techniques," in Proc. IEEE Int. Symp. Defect Fault Tolerance VLSI Syst. (DFT), Oct. 2009, pp. 236-244.

[45] D. Karakoyunlu and B. Sunar, "Differential template attacks on PUF enabled cryptographic devices," in Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS), Dec. 2010, pp. 1-6.

[46] C.-H. Chang, Y. Zheng and L. Zhang, "A retrospective and a look forward: Fifteen years of physical unclonable function advancement," IEEE Circuits Syst. Mag., vol. 17, no. 3, pp. 32-62, Aug. 2017.

[47] K. Lofstrom, W. R. Daasch and D. Taylor, "IC identification circuit using device mismatch," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2000, pp. 372-373.

[48] B. Gassend, D. Clarke, M. van Dijk and S. Devadas, "Silicon physical random functions," in Proc. ACM Conf. Comput. Commun. Secur. (CCS), Nov. 2002, pp. 148-160.

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

[50] 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 Symp. VLSI Circuits Dig. Tech. Papers, Jun. 2004, pp. 176-179.

[51] K. Yang, Q. Dong, D. Blaauw and D. Sylvester, "14.2 A physically unclonable function with BER <10−8 for robust chip authentication using oscillator collapse in 40nm CMOS," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2015, pp. 254-255.

[52] Z.-Y. Liang, H.-H. Wei and T.-T. Liu, "A wide-range variation-resilient physically unclonable function in 28 nm," IEEE J. Solid-State Circuits, vol. 55, no. 3, pp. 817-825, Mar. 2020.

[53] D. E. Holcomb, W. P. Burleson and K. Fu, "Power-up SRAM state as an identifying fingerprint and source of true random numbers," IEEE Trans. Comput., vol. 58, no. 9, pp. 1198-1210, Sep. 2009.

[54] 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 IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2014, pp. 278-279.

[55] M. Bhargava, C. Cakir and K. Mai, "Attack resistant sense amplifier based PUFs (SA-PUF) with deterministic and controllable reliability of PUF responses," in Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust (HOST), Jun. 2010, pp. 106-111.

[56] Y. Su, J. Holleman and B. P. Otis, "A digital 1.6pJ/bit chip identification circuit using process variations," IEEE J. Solid-State Circuits, vol. 43, no. 1, pp. 69-77, Jan. 2008.

[57] K. Ishibashi and K. Osada, Low Power and Reliable SRAM Memory Cell and Array Design, Berlin, Germany: Springer-Verlag, 2011.

[58] K. Anami, M. Yoshimoto, H. Shinohara, Y. Hirata and T. Nakano, "Design consideration of a static memory cell," IEEE J. Solid-State Circuits, vol. 18, no. 4, pp. 414-418, Aug. 1983.

[59] J. Lohstroh, E. Seevinck and J. De Groot, "Worst-case static noise margin criteria for logic circuits and their mathematical equivalence," IEEE J. Solid-State Circuits, vol. 18, no. 6, pp. 803-807, Dec. 1983.

[60] E. Seevinck, F. J. List and J. Lohstroh, "Static-noise margin analysis of MOS SRAM cells," IEEE J. Solid-State Circuits, vol. 22, no. 5, pp. 748-754, Oct. 1987.

[61] S. Satpathy, S. Mathew, J. Li, P. Koeberl, M. Anders, H. Kaul, G. Chen, A. Agarwal, S. Hsu and R. Krishnamurthy, "13 fJ/bit probing-resilient 250K PUF array with soft dark-bit masking for 1.94% bit-error in 22nm tri-gate CMOS," in Proc. Eur. Solid-State Circuits Conf. (ESSCIRC), Sep. 2014, pp. 239-242.

[62] 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 J. Solid-State Circuits, vol. 52, no. 4, pp. 940-949, Apr. 2017.

[63] A. Alvarez, W. Zhao and M. Alioto, "14.3 15fJ/b static physically unclonable functions for secure chip identification with <2% native bit instability and 140× Inter/Intra PUF hamming distance separation in 65nm," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2015, pp. 256-257.

[64] A. B. Alvarez, W. Zhao and M. Alioto, "Static physically unclonable functions for secure chip identification with 1.9–5.8% native bit instability at 0.6–1 V and 15

fJ/bit in 65 nm," IEEE J. Solid-State Circuits, vol. 51, no. 3, pp. 763-775, Mar. 2016.

[65] S. Taneja, A. Alvarez, G. Sadagopan and M. Alioto, "A fully-synthesizable C-element based PUF featuring temperature variation compensation with native 2.8% BER, 1.02fJ/b at 0.8–1.0V in 40nm," in Proc. IEEE Asian Solid-State Circuits Conf., Nov. 2017, pp. 301-304.

[66] S. Taneja, A. B. Alvarez and M. Alioto, "Fully synthesizable PUF featuring hysteresis and temperature compensation for 3.2% native BER and 1.02 fJ/b in 40 nm," IEEE J. Solid-State Circuits, vol. 53, no. 10, pp. 2828-2839, Oct. 2018.

[67] 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 IEEE Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2016, pp. 158-159.

[68] Y. Choi, B. Karpinskyy, K.-M. Ahn, Y. Kim, S. Kwon, J. Park, Y. Lee and M. Noh, "Physically unclonable function in 28nm FDSOI technology achieving high reliability for AEC-Q100 Grade 1 and ISO26262 ASIL-B," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2020, pp. 426-427.

[69] K. Yang, Q. Dong, D. Blaauw and D. Sylvester, "8.3 A 553F2 2-transistor amplifier-based Physically Unclonable Function (PUF) with 1.67% native instability," in IEEE Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2017, pp. 146-147.

[70] D. Li and K. Yang, "25.1 A 562F2 Physically Unclonable Function with a zero-overhead stabilization scheme," in IEEE Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2019, pp. 400-401.

[71] D. Li and K. Yang, "A self-regulated and reconfigurable CMOS physically unclonable function featuring zero-overhead stabilization," IEEE J. Solid-State Circuits, vol. 55, no. 1, pp. 98-107, Jan. 2020.

[72] M.-Y. Wu, T.-H. Yang, L.-C. Chen, C.-C. Lin, H.-C. Hu, F.-Y. Su, C.-M. Wang, J. P.-H. Huang, H.-M. Chen, C. C.-H. Lu, E. C.-S. Yang and R. S.-J. Shen, "A PUF scheme using competing oxide rupture with bit error rate approaching zero," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2018, pp. 130-131.

[73] K.-H. Chuang, E. Bury, R. Degraeve, B. Kaczer, D. Linten and I. Verbauwhede, "A physically unclonable function using soft oxide breakdown featuring 0% native BER and 51.8 fJ/bit in 40-nm CMOS," IEEE J. Solid-State Circuits, vol. 54, no. 10, pp. 2765-2776, Oct. 2019.

[74] Y. Pang, B. Gao, D. Wu, S. Yi, Q. Liu, W.-H. Chen, T.-W. Chang, W.-E. Lin, X. Sun, S. Yu, H. Qian, M.-F. Chang and H. Wu, "A reconfigurable RRAM physically unclonable function utilizing post-process randomness source with <6×10-6 native bit error rate," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2019, pp. 402-403.

[75] D. Jeon, J. H. Baek, Y.-D. Kim, J. Lee, D. K. Kim and B.-D. Choi, "A physical unclonable function with bit error rate <2.3 × 10−8 based on contact formation probability without error correction code," IEEE J. Solid-State Circuits, vol. 55, no. 3, pp. 805-816, Mar. 2020.

[76] J. Lee, D. Lee, Y. Lee and Y. Lee, "A 445F2 leakage-based physically unclonable function with lossless stabilization through remapping for IoT security," in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2018, pp. 132-133.

[77] M. Agarwal, B. C. Paul, M. Zhang and S. Mitra, "Circuit failure prediction and its application to transistor aging," in Proc. IEEE VLSI Test Symp. (VTS), May 2007, pp. 1-6.

[78] M. Bhargava, C. Cakir and K. Mai, "Reliability enhancement of bi-stable PUFs in 65nm bulk CMOS," in Proc. IEEE Symp. Hardw.-Oriented Secur. Trust (HOST), Jun. 2012, pp. 25-29.

[79] A. Garg and T. T. Kim, "Design of SRAM PUF with improved uniformity and reliability utilizing device aging effect," in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), Jun. 2014, pp. 1941-1944.

[80] S. Rangan, N. Mielke and E. C. Yeh, "Universal recovery behavior of negative bias temperature instability [PMOSFETs]," in Proc. IEEE Int. Electron Devices Meeting (IEDM), Dec. 2003, pp. 14.3.1-14.3.4.

[81] E. Takeda and N. Suzuki, "An empirical model for device degradation due to hot-carrier injection," IEEE Electron Device Lett., vol. 4, no. 4, pp. 111-113, Apr. 1983.

[82] A. Kawasumi, Y. Takeyama, O. Hirabayashi, K. Kushida, Y. Fujimura and T. Yabe, "A low-supply-voltage-operation SRAM with HCI trimmed sense amplifiers," IEEE J. Solid-State Circuits, vol. 45, no. 11, pp. 2341-2347, Nov. 2010.

[83] K. Miyaji, S. Tanakamaru, K. Honda, S. Miyano and K. Takeuchi, "Improvement of read margin and its distribution by VTH mismatch self-repair in 6T-SRAM with asymmetric pass gate transistor formed by post-process local electron injection," IEEE J. Solid-State Circuits, vol. 46, no. 9, pp. 2180-2188, Sep. 2011.

[84] K. Miyaji, T. Suzuki, S. Miyano and K. Takeuchi, "A 6T-SRAM with a post-process electron injection scheme that pinpoints and simultaneously repairs disturb fails for 57% less read delay and 31% less read energy," IEEE J. Solid-State Circuits, vol. 48, no. 9, pp. 2239-2249, Sep. 2013.

[85] M. Bhargava and K. Mai, "A high reliability PUF using hot carrier injection based response reinforcement," in Proc. IACR Int. Conf. Cryptogr. Hardw. Embeded Syst. (CHES), Aug. 2013, pp. 90-106.

[86] K. Liu, Y. Min, X. Yang, H. Sun and H. Shinohara, "A 373 F2 2D power-gated EE SRAM physically unclonable function with dark-bit detection technique," in Proc. IEEE Asian Solid-State Circuits Conf. (A-SSCC), Nov. 2018, pp. 161-164.

[87] K. Liu, Y. Min, X. Yang, H. Sun and H. Shinohara, "A 373-F2 0.21%-native-BER EE SRAM physically unclonable function with 2-D power-gated bit cells and VSS bias-based dark-bit detection," IEEE J. Solid-State Circuits, vol. 55, no. 6, pp. 1719-1732, Jun. 2020.

[88] K. Liu, H. Pu and H. Shinohara, "A 0.5-V 2.07-fJ/b 497-F2 EE/CMOS hybrid SRAM physically unclonable function with < 1E-7 bit error rate achieved through hot carrier injection burn-in," in IEEE Custom Integr. Circuits Conf. (CICC), Mar. 2020, pp. 1-4.

[89] K. Liu, X. Chen, H. Pu and H. Shinohara, "A 0.5-V hybrid SRAM physically unclonable function using hot carrier injection burn-in for stability reinforcement," IEEE J. Solid-State Circuits, pp. 1-12, Accepted, Available on IEEE Xplore Early Access.

[90] B. Razavi, Fundamentals of Microelectronics, 2nd ed., Hoboken, NJ: John Wiley & Sons, Inc., 2013.

[91] M. Cortez, S. Hamdioui, V. van der Leest, R. Maes and G.-J. Schrijen, "Adapting voltage ramp-up time for temperature noise reduction on memory-based PUFs," in Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust (HOST), Jun. 2013, pp. 35-40.

[92] Z. Cui, B. Zheng, Y. Piao, S. Liu, R. Xie and H. Shinohara, "Measurement of mismatch factor and noise of SRAM PUF using small bias voltage," in Proc. Int. Conf. Microelectronic Test Struct. (ICMTS), Mar. 2017, pp. 1-4.

[93] H. Shinohara, B. Zheng, Y. Piao, B. Liu and S. Liu, "Analysis and reduction of SRAM PUF bit error rate," in Proc. Int. Symp. VLSI Design, Autom. Test (VLSI-DAT), Apr. 2017, pp. 1-4.

[94] K. Agarwal and S. Nassif, "Statistical analysis of SRAM cell stability," in Proc. ACM/IEEE Des. Autom. Conf. (DAC), Jul. 2006, pp. 57-62.

[95] B. Razavi, "The StrongARM latch [A circuit for all Seasons]," IEEE Solid-State Circuits Mag., vol. 7, no. 2, pp. 12-17, Jun. 2015.

PUBLICATIONS

6.1 Journal articles

[1] Kunyang Liu, Xinpeng Chen, Hongliang Pu, and Hirofumi Shinohara, “A 0.5-V hybrid SRAM physically unclonable function using hot carrier injection burn-in for stability reinforcement,” *IEEE Journal of Solid-State Circuits*, Accepted, Available on IEEE Xplore Early Access, pp. 1-12.

[2] Kunyang Liu, Yue Min, Xuan Yang, Hanfeng Sun, and Hirofumi Shinohara, “A 373-F^2 0.21%-native-BER EE SRAM physically unclonable function with 2-D power-gated bit cells and VSS bias-based dark-bit detection,” *IEEE Journal of Solid-State Circuits*, vol. 55, no. 6, pp. 1719-1732, Jun. 2020.

6.2 International conference papers

[1] Ruilin Zhang, Xingyu Wang, Luying Wang, Xinpeng Chen, Fan Yang, Kunyang Liu, and Hirofumi Shinohara, “A 0.186-pJ per bit latch-based true random number generator with mismatch compensation and random noise enhancement,” in *Proc. Symp. VLSI Circuits*, Kyoto, Japan, Jun. 2021. (Accepted)

[2] Kunyang Liu, Zihan Fu, Gen Li, Hongliang Pu, Zhibo Guan, Xingyu Wang, Xinpeng Chen, and Hirofumi Shinohara, “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 *IEEE International Solid-State Circuits Conference (ISSCC) Dig. Tech. Papers*, San Francisco, USA, pp. 502-503, Feb. 2021.

[3] Xingyu Wang, Hongjie Liu, Ruilin Zhang, Kunyang Liu, and Hirofumi Shinohara, “An inverter-based true random number generator with 4-bit von-Neumann post-processing circuit,” in *Proc. IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, Springfield, USA, pp. 285-288, Aug. 2020.

[4] Kunyang Liu, Hongliang Pu, and Hirofumi Shinohara, “A 0.5-V 2.07-fJ/b 497-F^2 EE/CMOS hybrid SRAM physically unclonable function with < 1E-7 bit error rate achieved through hot carrier injection burn-in,” in *Proc. IEEE Custom Integrated Circuits Conference (CICC)*, Boston, USA, pp. 1-4, Mar. 2020.

[5] Kunyang Liu, Yue Min, Xuan Yang, Hanfeng Sun, and Hirofumi Shinohara, “A 373 F^2 2D power-gated EE SRAM physically unclonable function with dark-bit detection technique,” in *Proc. IEEE Asian Solid-State Circuits Conference (A-SSCC)*, Tainan, Taiwan, pp. 161-164, Nov. 2018.

6.3 Domestic conference papers

[1] Shiyu Liu, Baikun Zheng, Yanhao Piao, Kunyang Liu, Ronghao Xie, and Hirofumi Shinohara, “Low bit error rate latch PUF with EE structure,” in *Proc. IEICE Symposium on Cryptography and Information Security (SCIS)*, 4 pages, Naha, Japan, Jan. 2017.

6.4 Reports

[1] Kunyang Liu and Hirofumi Shinohara, “A hybrid SRAM PUF with HCI-based stability reinforcement,” in *Proc. International collaboration Symposium on Information, Production and Systems (ISIPS)*, 1 page, Kitakyushu, Japan, Nov. 2020.

[2] Kunyang Liu and Hirofumi Shinohara, “Effective dark-bit detection scheme based on EE SRAM physically unclonable function and VSS biasing technique,” in *Proc. International collaboration Symposium on Information, Production and Systems (ISIPS)*, 1 page, Kitakyushu, Japan, Nov. 2019.

[3] Kunyang Liu and Hirofumi Shinohara, “A 373 F^2 2D-power-gated EE SRAM physically unclonable function with dark-bit detection technique,” in *IEICE Technical Report of ICD on Memory*, 1 page, Tokyo, Japan, Apr. 2019. (Invited)

[4] Kunyang Liu and Hirofumi Shinohara, “A latch-type physically unclonable function (PUF) with <0.32% native bit error rate across 1.0-2.0V,” in *Proc. International*

collaboration Symposium on Information, Production and Systems (ISIPS), 4 pages, Kitakyushu, Japan, Nov. 2017.

ACKNOWLEDGMENTS

Throughout the five years in the Graduate School of IPS, Waseda University, I am fortunate enough to receive so many help and advices in my study and daily life. I would like to thank all the people I met here.

At first, I would like to express my deepest appreciations to my supervisor Prof. Hirofumi Shinohara, who always gives enlightening suggestions to me. Without your instructions and recognitions, I cannot achieve what I have achieved. I feel very lucky and grateful that I could be a member of this laboratory.

Second, I would like to thank Prof. Takashi Ohsawa and Prof. Kiyoto Takahata. Your advices help me a lot in improving my dissertation and presentation. I feel enlightened with your suggestions.

Third, I would like to thank all the other Professors at IPS for your suggestions and reviews on my study. Your advices could help improve my research to the next level.

I would also like to thank Mr. Koichi Takanashi from ROHM Company Ltd. for his support and the technical discussions.

Moreover, I would like to thank all the members of the laboratory, who are indispensable in my research and daily life. Especially, I would like to thank the senior alumni Dr. Jing Wang, other two doctoral students Ruilin Zhang and Xingyu Wang, all the co-authors of the related publications Yue Min, Xuan Yang, Hanfeng Sun, Hongliang Pu, and Xinpeng Chen, and the senior students Shiyu Liu, Yanhao Piao, and Baikun Zheng, who have guided me in this research.

Last but not least, I must thank my families, especially my father and mother, for their support throughout my whole life. Without you, I cannot go this far, especially realizing my wish to study in Japan. Thank you very much.