X-ray and Proton Radiation Effects on 40 nm CMOS Physically Unclonable Function Devices

By
Pengfei Wang

Thesis
Submitted to the Faculty of the Graduate School of Vanderbilt University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE in

Electrical Engineering

December 16, 2017
Nashville, Tennessee

Approved:

Robert A. Reed, Ph.D.

Daniel M. Fleetwood, Ph.D.
I would like to acknowledge and express my sincere gratitude to all the professors at Vanderbilt University I have had the good fortune to encounter over the course of my studies and this work. First and foremost, I would like to thank my advisor, Dr. Robert A. Reed, for his guidance in my academic growth and research pursuits. I would also like to thank Dr. Daniel M. Fleetwood for teaching me radiation effects and all the insight suggestions, Dr. Ronald D. Schrimpf for teaching me the device physics and all the helpful discussions. I consider myself especially indebted to Dr. Enxia Zhang. She is patient and always there to help. She provides boundless support to my work. Thank you to Michael W. McCurdy for his help during long days down in the FEL.

I would like to thank my parents, Qi Wang and Xuemei Yan, for their unconditional love and support. I thank my fellow graduate students in the Radiation Effects and Reliability group, in particular Kai Ni, Huiqi Gong, Wenjun Liao, Pan Wang, Charles N. Arutt and Ryan Keller for their advice and friendship. Thank you to collaborators from imec, Kai Hsin Chuang, Dr. Erik Bury and Dr. Dimitri Linten. I would also like to thank AFRL and AFOSR through the Hi-REV program and the Defense Threat Reduction Agency through its Basic Research program for supporting this project.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................................................. ii

LIST OF FIGURES .......................................................................................................................... v

LIST OF TABLES ............................................................................................................................ viii

Chapter

I. Introduction of Radiation Effects ............................................................................................... 1

   1. Overview of Space Radiation Environments ........................................................................ 1
   2. Total Ionizing Dose Effects .............................................................................................. 3
   3. Displacement Damage ......................................................................................................... 6

II. Physically Unclonable Function (PUF) ...................................................................................... 8

   1. Information Security ......................................................................................................... 8
   2. PUF Concept ..................................................................................................................... 12
   3. Types of PUFs ................................................................................................................. 13
      a. Strong PUF ................................................................................................................ 14
      b. Weak PUF .................................................................................................................. 15

III. Experimental Details of BD-PUFs ........................................................................................... 22

   1. Device Information .......................................................................................................... 22
   2. DC Characteristics .......................................................................................................... 23
   3. X-ray Irradiation ........................................................................................................... 26
   4. Proton Irradiation ............................................................................................................ 26

IV: Radiation Effects on BD-PUFs ................................................................................................ 28

   1. X-ray Irradiation Response ............................................................................................. 28
   2. 1.8 MeV Proton Irradiation Response .......................................................................... 30
      a. Grounded Condition .................................................................................................... 30
      b. Read-out Condition .................................................................................................... 34
      c. Standby Condition ...................................................................................................... 35
   3. Selector-transistor and Leakage Effects ........................................................................... 38

V: Conclusions ............................................................................................................................... 46
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 1-1. Artist’s drawing of the earth’s Van Allen radiation belts. Drawing does not Show the SAA</td>
<td>3</td>
</tr>
<tr>
<td>Fig. 1-2. Schematic of n-channel MOSFET illustrating radiation-induced charging of the gate oxide: (a) normal operation and (b) post-irradiation</td>
<td>4</td>
</tr>
<tr>
<td>Fig. 1-3. Schematic energy band diagram for MOS structure, indicating major physical processes underlying radiation response</td>
<td>6</td>
</tr>
<tr>
<td>Fig. 2-1. Relationships between information security, cryptography, physical security and physical roots of trust</td>
<td>11</td>
</tr>
<tr>
<td>Fig. 2-2. Challenge-response system</td>
<td>14</td>
</tr>
<tr>
<td>Fig. 2-3. “Fingerprint” chips</td>
<td>16</td>
</tr>
<tr>
<td>Fig. 2-4. SRAM cell CMOS circuit</td>
<td>17</td>
</tr>
<tr>
<td>Fig. 2-5. SRAM cell logic circuit</td>
<td>17</td>
</tr>
<tr>
<td>Fig. 2-6. SRAM cell voltage transfer curves</td>
<td>18</td>
</tr>
<tr>
<td>Fig. 2-7. SRAM cell power-up transient analysis</td>
<td>20</td>
</tr>
<tr>
<td>Fig. 3-1. (a) Unit cell of BD-PUF, consisting of two minimum sized nFETs ($W \times L = 120 \times 40$ nm$^2$) and a pFET selector. Breakdown can be generated on one of the nFETs randomly. (b) The array with 60 cells, which has been fabricated on a commercial CMOS process</td>
<td>23</td>
</tr>
<tr>
<td>Fig. 3-2. HP4156 semiconductor parameter analyzer</td>
<td>24</td>
</tr>
</tbody>
</table>
Fig. 3. Forming step of a BD-PUF. The $\overline{BL}$ nFET experiences breakdown at $\sim 0.25$ s ..........25

Fig. 3-4. Input/output currents vs. input voltage.................................................................25

Fig. 3-5. 10-keV ARACOR 4100 X-ray irradiator .................................................................26

Fig. 3-6. Pelletron accelerator at Vanderbilt University .........................................................27

Fig. 4-1. Input/output currents vs. input voltage for (a) 20 cycles, demonstrating little cycle-to-

cycle variation, and (b) a BD-PUF at different TID levels for 10-keV X-ray irradiation

.............................................................................................................................................29

Fig. 4-2. All the terminals of the BD-PUF were grounded during proton irradiation ..........31

Fig. 4-3. (a) Current read-out of BD-PUF, and (b) $\overline{M}$ current before and after 1.8 MeV proton

irradiation..................................................................................................................................31

Fig. 4-4. (a) The change of input voltage when the BL current is $-0.2 \ \mu$A, and percentage

decrease of BL current at 1.2 V, and (b) memory ratio as a function of fluence and

annealing time............................................................................................................................33

Fig. 4-5. The voltage applied in the readout condition of the BD-PUF .....................................34

Fig. 4-6. Current read-out of BD-PUF .....................................................................................35

Fig. 4-7. The voltage applied in the standby condition of BD-PUFs. .......................................36

Fig. 4-8. (a) Current read-out of BD-PUF, and (b) BL current before and after 1.8 MeV proton

irradiation..................................................................................................................................36

Fig. 4-9. Memory ratio as a function of fluence and annealing time ......................................37

Fig. 4-10. (a) Semi-log plot of $I_D - V_G$ curve as a function of fluence; (b) threshold voltage shifts

as a function of fluence and annealing time. Error bars here show the full range of

variation observed......................................................................................................................38
Fig. 4-11. Semi-log plots of $I_D - V_G$ curves for (a) the broken $n$FET, and (b) the unbroken $n$FET as a function of fluence. .................................................................40

Fig. 4-12. $I_G - V_G$ curves for (a) the broken $n$FET, and (b) the unbroken $n$FET as a function of fluence .............................................................................................................41

Fig. 4-13. Current in the BD-PUF after its forming step and proton irradiation. Current from the Input flows through the $p$FET selector, and then through the parallel combination of the broken and unbroken $n$FETs. The resistance of the broken $n$FET is much lower than that of the unbroken $n$FET, so before and after proton irradiation, nearly all of the current flows through the broken $n$FET. While the majority of current still flows through the broken $n$FET after proton irradiation, an increasing amount of leakage current flows through the unbroken $n$FET, which shares a common body junction with the broken $n$FET .............................................................................................................43

Fig. 4-14. The $I_D-V_D$ curves of the $p$FET selector and the $n$FETs load line as a function of fluence ........................................................................................................................................44
LIST OF TABLES

Table 1.1 Maximum energies of particles...............................................................................................................1
CHAPTER I

INTRODUCTION OF RADIATION EFFECTS

1. **Overview of Space Radiation Environments**

   Since radiation exists throughout the universe originating from many sources and with varying intensities, topics in radiation environment research cover a wide subject matter. The space radiation environment can be divided into two groups: the particles trapped by planetary magnetospheres in “bells,” including protons, electrons, and heavier ions and transient particles which include protons and heavy ions of all of the elements of the periodic table [1]. The transient radiation includes galactic cosmic ray (GCR) particles and particles from solar events, such as coronal mass ejections and flares [1]-[6]. Table I shows the maximum energy of space radiation particles. Much of the environment is high energy, and shielding is not effective for many radiation source environments.

   Table 1.1 Maximum energies of particles. (After [1].)

<table>
<thead>
<tr>
<th>Particle Type</th>
<th>Maximum Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trapped Electrons</td>
<td>10s of MeV</td>
</tr>
<tr>
<td>Trapped Protons &amp; Heavy Ions</td>
<td>100s of MeV</td>
</tr>
<tr>
<td>Solar Protons</td>
<td>GeV</td>
</tr>
<tr>
<td>Solar Heavy Ions</td>
<td>GeV</td>
</tr>
<tr>
<td>Galactic Cosmic Rays</td>
<td>TeV</td>
</tr>
</tbody>
</table>
Van Allen is credited with discovering the trapped proton and electron regions around the earth. An artist’s drawing of the Earth’s Van Allen radiation belts is shown in Fig. 1-1 [1]. The tilt of the Earth’s magnetic pole form the geographic pole and the displacement of the magnetic field from the center cause a dip in the field over the South Atlantic Ocean, causing a bulge in the underside of the inner belt. This region is called the South Atlantic Anomaly (SAA). SAA plagues spacecraft, and the flux levels are much lower than those at higher altitudes. For most shielded spacecraft systems, ions at energies high enough to penetrate spacecraft materials are too low to be a dominant factor in single event effects rates [1]-[3]. The trapped particles’ levels and locations are highly dependent on particle energy, altitude, inclination, and the activity level of the sun and are highly dynamic. The levels of GCR are modulated by the 11-year solar cycle with the peak GCR populations occurring near solar minimum. Galactic and solar particles have free access to spacecraft outside of the magnetosphere. The transient particles penetrate the Earth’s magnetosphere, and they can reach near-Earth orbiting spacecraft and are particularly hazardous to satellites in polar, highly elliptical, and geostationary (GEO) orbits [1]-[6].
2. *Total Ionizing Dose Effects*

Energetic particles incident on a solid will lose their energy to both ionizing and non-ionizing processes as they travel through a material. The result of this energy loss is the production of electron/hole pairs (ionizing) and displaced atoms (displacement damage). Fig. 1-2 shows one of the basic radiation problems in a MOS transistor. The normal operation of a MOSFET is shown in Fig. 1-2(a) [7]. The device is turned on when a conducting channel has formed between the source and drain after an appropriate gate voltage has been applied to the gate. The effect of ionizing radiation is explained in Fig. 1-2(b) [7]. Radiation induced trapped charge builds up in the gate oxide, and this causes a threshold voltage shift. The device cannot be turned off if the shift is large enough, even at 0 volts, and the device is failed by going depletion mode [7].
Fig. 1-2 Schematic of n-channel MOSFET illustrating radiation-induced charging of the gate oxide: (a) normal operation and (b) post-irradiation. (After [7].)
Fig. 1-3 shows the schematic energy band diagram of a MOS structure, where positive bias is applied to the gate, and electrons flow toward the gate and holes move to the Si substrate. The oxide insulators are the most sensitive parts of a MOS system to radiation. Electron/hole pairs are created by the deposited energy when a gate oxide exposed by radiation. The electrons are more mobile than the holes [7]-[8], and they are swept out of the oxide in a picosecond or less. However, some fraction of the electrons and holes will recombine in this picosecond. That fraction will highly depend on the energy and type of the incident particle. The holes that survive the initial recombination are immobile and remain near their point of generation, so a negative threshold voltage shifts in a MOS transistor [7]. The second process is the transport of the holes to the Si/SiO$_2$ interface, and this causes the short term recovery of the threshold voltage. This process is normally done in less than 1 s at room temperature, but it can take many orders of magnitude slower at low temperature. The third process is that some fraction of the transporting holes fall into relatively deep long-lived trap states. These trapped holes can cause a remnant negative voltage shift. The last process of MOS radiation response is the radiation-induced buildup of interface traps right at the Si/SiO$_2$ interface. Interface traps highly depend on oxide processing, applied field and temperature [7].
Displacement Damage

Displacement damage is the result of a non-ionizing process where an atom is displaced. Vacancies and interstitials are the main lattice defects. A vacancy is the absence of an atom from its normal lattice position. An interstitial is that displaced atom moves into a non-lattice position. The divacancy is the defect formed by two adjacent vacancies. There may be larger grouping of vacancies in irradiated silicon. Additional types of defects can form when vacancies and interstitials are adjacent to impurity atoms [9]. In general, incident energetic particles can produce a mixture of isolated and clustered defects. The defects formed by incident radiation will reorder to form a more stable configuration. The effectiveness of defects in altering the properties of bulk semiconductor material depends on the nature of the specific defects and on the time after defect creation. Defect reordering is also called annealing. The amount of damage
and its effectiveness are reduced. Damage effectiveness are depending on many factors, including particle type, particle energy, irradiation temperature, measurement temperature, time after irradiation, thermal history after irradiation, injection level, material type, and impurity type and concentration[9]. Displacement damage in a radiation environment can cause materials and devices to degrade, and the basic phenomena are: incident particles displace atoms; the resulting defects give rise to new energy levels; and those levels alter material and device electrical and optical properties [9].
CHAPTER II

Physically Unclonable Function (PUF)

A PUF is an essential building block for hardware security, which is typically used in commercial silicon technology, and enables chip level cryptographic functions such as secret key generation, entity authentication and IP protection [10]-[13]. In this chapter, the relationship between information security, cryptography, physical security and physical roots of trust are introduced. The details of the PUF concept and two types of PUFs are explained. SRAM PUFs of weak PUFs are analyzed in detail.

1. Information Security

In our life, we constantly put our trust in many things. For example, when we send someone a letter, we trust the courier of the post office to deliver the letter in a timely manner to the correct person, and to keep the letter closed so that no one else can read its content. Such trust based interactions work out in the right way most of time, because the parties we interact with are trustworthy. However, we do not live in an ideal world. We need systems that induce, guarantee or even enforce trustworthiness of parties in our non-ideal world. This is called security that enables trust [10]-[13]. Information security deals with securing interactions involving communication and information. Countless quantities of private and sensitive information are stored and communicated over the internet and other digital networks every second around us. Our society has become a flow of digital information in many ways, and reliable information security techniques are essential to enable trust in our digital world [10].
Here are several goals for information security techniques: data confidentiality relates to keeping information secret from unauthorized parties; entity authentication deals with obtaining proof of the identity and the presence of the entity one is interacting with; data integrity and authentication is aimed at preventing and detecting unauthorized alteration of data and ensuing the origin of the data [10] -[13].

Cryptography and cryptanalysis are the subfield of cryptology. Cryptography deals with the construction of protocols and algorithms to achieve information security goals, typically on a mathematical basis. Cryptanalysis analyzes the security of cryptographic constructions by attempting to break their anticipated security [10], [14]. A basic design principle for cryptographic constructions is to reduce the security goal they attempt to achieve to the security of a single parameter in the construction, which is called the key. The obtained level of security is the required effort to break it without knowing the key. Cryptographic primitives can be grouped based on the nature of their key [10], [14]. Here are three types: unkeyed primitives are constructions which do not require a key; symmetric-key primitives are based on a single key which is only known to authorized parties and secret to anyone else; public-key primitives are based on a key pair, one of which is public and the other is kept private [14].

For nearly all keyed cryptographic primitives, it needs three parts. The first is the secure key generation. A secure key, which is random, unique and unpredictable, can be generated for every instantiation of the primitive. The second is the secure key storage. The key can be assigned to, stored and retrieved by the instantiation without being revealed. The last is the secure execution [10], [14]. The instantiation can execute the cryptographic algorithm without revealing any partial information about the key or about internal results, and without an outsider being able to influence the internal execution in any possible way. None of these three parts can be achieved
through information security techniques, but require physical security measures. Cryptographic primitives reduce information security objectives into physical security requirements [10], [14].

Secure key generation is difficult, because there is a significant shortage in randomness in a large collected set of actually used public keys from a public key signature scheme, likely caused by badly implemented key generators. For some of the keys in the analyzed collection, this leads to an immediate loss of security [15]. Storing secret keys in a highly secure manner partially contracts the fact that they still need to be in some permanent digital format to be usable in an algorithm. For typical digital implements, this means that the key bits reside somewhere in a non-volatile digital memory on a silicon chip. Even with extensive countermeasures in place, it is very difficult to stop a well-equipped and/or determined adversary from gaining physical access to key memories [16]-[17]. There are many ways an adversary can break the secure execution assumption, both on the software and on the hardware level. Modern cryptographic implementations can no longer ignore side-channel attacks, which take advantage of the reality that all actions on a digital platform leak information about their execution through side channels, e.g., through their execution time [18], their power consumption [19], their electro-magnetic radiation [20], etc.

We cannot depend on mathematical reductions any more to provide physical security objectives. We need to develop physical techniques and primitives based on physical reasoning, which can be trusted to withstand certain physical attacks and can hence provide certain physical security objectives. These primitives are called physical roots of trusts. How information security objectives can be achieved from physical security and eventually from physical roots of trust is shown in Fig. 2-1 [10]. Possible candidates of physical roots of trust are: true random number generators or TRNGs [21]-[22] harvest random numbers from truly physical sources of
randomness and can therefore be trusted to produce highly random keys for cryptographic purposes; design styles for digital silicon circuits have been developed which minimize and ideally eliminate certain physical side channels [23]; physically unclonable functions produce unpredictable and instance-specific values and can be used to provide physically secure key generation and storage.

Fig. 2-1 Relationships between information security, cryptography, physical security and physical roots of trust. (After [10].)
2. **PUF Concept**

The current best practice for providing a secure memory or authentication source in a mobile system is to place a secret key in a nonvolatile electrically erasable programmable read-only memory or battery-backed static random-access memory and use hardware cryptographic operations such as digital signatures or encryption. This approach is expensive both in terms of design area and power consumption. Moreover, a nonvolatile memory is often vulnerable to invasive attack mechanisms. Protection against such attacks requires the use of active tamper detection/prevention circuitry which must be continually powered \[10\]. Physical unclonable functions are a promising innovative primitive that are used for authentication and secret key storage without the requirement of secure electrically erasable programmable read-only memories and other expensive hardware \[24\]-\[25\]. In an attempt to express the concept of a Physical unclonable function in a single phrase, one of the best possible descriptions would be: “a PUF is an object’s fingerprint”. Here are some reasons why PUFs are similar to fingerprints. First, a human fingerprint is a feature which strongly expresses individualism. In a more inanimate sense, a PUF is an identifying feature of a specific instance of a class of objects, or for short is an instance-specific feature. Second, as an individualizing feature, a fingerprint is inherent. In the same way, PUFs are inherently present in an object from its creation, as a result of unique variations during its creation process. Finally, fingerprints are unclonable \[10\]. In fact, unclonability is one of the core properties of a PUF. Here is the definition of the PUF concept given in \[10\]: “a PUF is an expression of an inherent and unclonable instance-specific feature of a physical object”.


3. Types of PUFs

There are two primary applications of PUFs: low-cost authentication and secure key generation [10]. These two applications have resulted from the fact that PUFs designed during the past decade have fallen into two groups. These categories are described as “strong PUFs” and “weak PUFs”. Strong PUFs are typically used for authentication, while weak PUFs are used for key storage. Each PUF can be modeled as a black-box challenge-response system like that in Fig. 2-2. A PUF is passed an input challenge c, and returns a response \( r = f(c) \), where \( f(\cdot) \) describes the input/output relations of the PUF. The black-box is appropriate here, because the internal parameters of \( f(\cdot) \) are hidden from the user since they represent the internal manufacturing variability that the PUF uses to generate a unique challenge-response set [10], [26]. PUF security relies on the difficulties of measurement or estimation of these parameters and the difficulties of manufacturing two chips with the same set of parameters. A PUF is based on the idea that even though the mask and manufacturing process is the same among different ICs, each IC is actually slightly different due to normal manufacturing variability [10]. PUFs leverage this variability to derive secret information that is unique to the chip. Due to the manufacturing variability that defines the secret, one cannot manufacture two identical chips, even with full knowledge of the chip’s design. PUF architectures exploit manufacturing variability in multiple ways, like gate delay, power-on state of SRAM, threshold voltage, and many other physical characteristics [10], [26].
The fundamental difference between weak and strong PUFs is the domain of \( f(\cdot) \), or informally, the number of unique challenges \( c \) that the PUF can process [26]. A weak PUF can only support a small number of challenges. In some cases, a weak PUF only has a single challenge. A strong PUF can support a large number of challenges such that complete determination/measurement of all challenge-response pairs (CRPs) within a limited time is not feasible [10], [26].

\( \text{a. Strong PUF} \)

A strong PUF can be authenticated directly without using any cryptographic hardware, because it can support a large number of CRPs [26]. Here are some requirements for a strong PUF. First, a strong PUF needs to have large enough challenge-response space such that an adversary cannot enumerate all CRPs within a fixed time. Ideally, it is exponential in the number of challenge bits. Second, its response is stable to the environment during multiple readings. Third, an adversary with a polynomial-sized sample of adaptively chosen CRPs cannot predict the response to a new, randomly chosen challenges. Fourth, it is not feasible to manufacture two PUFs with the same responses. Finally, the readout only reveals the response \( r = f(c) \), and no other data about the internal functionality of the PUF [10], [26].
A weak PUF can provide authentication capabilities if the weak PUF is paired with crypto hardware [26]. The security models for weak and strong PUFs are different. The output of a weak PUF must be kept private, while a strong PUF’s responses do not have the same restriction. When \( r = f(c) \) is revealed, the strong PUF has the additional requirement of readout access restriction due to this difference in security models. The readout time of the PUF in conjunction with the number of CRPs must be considered to prevent total enumeration of the strong PUF. A faster PUF response allows for faster enumeration of all PUF CRPs. The surrounding digital cryptographic hardware is responsible for limiting access to the weak PUF output, because a weak PUF provides a secret key [10], [26]. However, the strong PUF does not need the use of additional crypto hardware to provide authentication services [26].

\[ b. \text{ Weak PUF} \]

Weak PUFs can be thought of as PUFs that directly digitize some “fingerprint” of the circuit as shown in Fig. 2-3 [26]-[27]. A weak PUF can only be interrogated by one or a small number of challenges, because the fingerprint signature remains invariant. The \( f(\cdot) \) only has a domain of one or a small number of inputs, and will also have a very small range, as a given challenge should always result in the same response when we ignore noise [10], [26]. Here are some properties of weak PUFs. First, it has a small number of CRPs, which linearly related to the number of components whose behavior depends on manufacturing variation. Second, response is stable and robust to environmental conditions and multiple reading so that a challenge always yields the same response. Third, responses are unpredictable and depend strongly on the innate manufacturing variability of the device. Finally, it is impractical to manufacture two devices with the same physical fingerprint [26].
The static random-access memory (SRAM) is an extreme example of a weak PUF in the sense that it only has one CRP [10], [26]. In a typical CMOS implementation, an individual SRAM cell is built with six transistors, as shown in Fig. 2-4. The logic memory functionality of a cell comes from two cross-coupled invertors at its core, shown in Fig. 2-5, each built from two MOSFETs, one pMOS and one nMOS [10]. From an electronic view, this circuit contains a positive feedback loop which reinforces current state. In a logic sense, this circuit has two stable values, and by residing in one of the two states the cell stores one binary digit. Two additional access MOSFETs are used to read and write the data [10]. Many SRAM cells are arranged in large memory array structures, capable of storing many kilobits or megabits. An SRAM is volatile which means the state will be lost shortly after power-down [10], [26].
Fig. 2-4 SRAM cell CMOS circuit. (After [10].)

Fig. 2-5 SRAM cell logic circuit. (After [10].)
Fig. 2-6, which draws the voltage transfer curves of the two cross-coupled inverters, shows the operation of an SRAM cell [10]. The cross-coupled CMOS inverter structure has three possible operating points of which only two are stable and one is metastable. The stable points are characterized by properties such that deviations from these points are reduced and the stable point condition is stored. This does not hold for the metastable point. Electronic circuits are constantly affected by small deviations due to random noise, an SRAM cell will never stay in metastable state long, and will quickly and randomly end up in one of the two stable states [26].

![Fig. 2-6 SRAM cell voltage transfer curves. (After [10].)]
The operation principle of an SRAM PUF is based on the transient behavior of an SRAM cell when it is powered up. The circuit will evolve to one of operating points, but it is not immediately clear to which one. The preferred initial operating point of an SRAM cell is determined by the difference in strength of the MOSFETs in the cross-coupled inverter circuit [10]. Fig. 2-7 shows the transient behavior of an SRAM at power-up. Typical SRAM cells are designed to have perfectly matched inverter for efficiency and performance reasons. The actual difference in strength between the two invertors called device mismatch, which is the transistor threshold mismatch, is caused by random process variations in the production process. Each cell will have a random preferred initial operating point. When one of the inverters is significantly stronger than the other one, the preferred initial operating point will be a stable state and the preference will be very distinct, i.e. such a cell will always power-up in the same stable state, but which state (‘0’ or ’1’) is randomly determined for every cell [10]. When the mismatch in cell is small, the effect of random circuit noise such as die temperature, power supply fluctuations, and common mode process variations, comes into play [26]. Cells with a small mismatch still have a preferred initial stable state which is determined by the sign of the mismatch, but due to voltage noise there is a non-negligible probability that they power-up in their non-preferred state. Finally, cells have a negligible mismatch between their inverters, will power up in, or very close to the metastable operating point. Their final stable state will be random for every power-up [10].

The magnitude of the impact of process variations on random device mismatch in SRAM cells causes most cells to have a strongly preferred but cell-specific initial state, and only few cells have a weak preference or no preference at all [10], [26]. The power-up state of a typical SRAM cell shows strong PUF behavior. Large arrays of SRAM cells are able to provide thousands to millions of response bits for this SRAM PUF. The address of a specific cell in the
array can be considered the challenge of the SRAM PUF [10]. The first experimental implementation was performed in 2007, where a custom SRAM array based on 0.13 μm technology was shown to generate random values based on threshold mismatches [28]-[29]. Additional work showed that SRAM initialization can produce a unique physical fingerprint for each chip [30]-[31].

![SRAM cell power-up transient analysis](image)

**Fig. 2-7** SRAM cell power-up transient analysis. (After [10].)

Although a SRAM cell is symmetric, manufacturing variability will give each cell a tendency toward a logical “1” or “0” at power-on. This variability is random across the entire SRAM and gives it a unique fingerprint on power-on that can be identified. If the response consists of the entire SRAM state at power-on, the notion of a challenge is not useful, as there is only one possible challenge: powering on the SRAM [10], [26]. The output signature is always
the same when we ignore noise. One can allow for more output bits by increasing the size of the SRAM, but the response space is still linearly related to the number of components subject to manufacturing variation [26]. The weak PUF output may be used as the key in a keyed-hash message authentication code challenge-response sequence, and as a secret key to encrypt/decrypt data on the device [10], [26].
CHAPTER III

Experimental Details of BD-PUFs

The widely discussed SRAM PUF uses process-induced threshold voltage variation to generate random values, but lacks the required stability [31]. Hence a novel PUF using the intrinsic randomness of the oxide breakdown (BD) positions in transistors (BD-PUF) [32] has been evaluated in this work. This chapter introduces the device information, electrical characteristics, and measurement techniques employed in BD-PUFs.

1. Device Information

The test structure for the unit cell of the BD-PUF is shown in Fig. 3-1(a), and the unit cell has been designed into an array configuration as shown in Fig. 3-1(b) with 60 BD-PUF cells in each array. This design is then fabricated in a commercial 40nm CMOS technology.

The unit cell of the BD-PUF examined consists of two minimum sized nFETs, each with shorted source and drain, and a pFET selector [32]. A forming step is used to establish the PUF unit by (random) breakdown of the gate dielectric in one of the two nFETs. In practice, a high voltage is applied to the gates of the nFETs by enabling the pFET compliance transistor. The high voltage applied to the nFET gate generates random defects within the gate oxide until hard failure occurs. As soon as one of the nFETs experiences breakdown, the current through the broken oxide will create a voltage drop on the pFET selector, which now acts as a compliance FET in saturation mode [33], limiting the stress voltage and current. The breakdown path in the broken nFET will further wear-out during this condition, in a current-limited way [34]. The unbroken nFET, however, will not accumulate additional damage in this phase due to the
reduced stress voltage. As a result, a “soft” breakdown path only will have been generated in one

\( nFET \) [32], [35].

![Diagram](image)

Fig. 3-1 (a) Unit cell of BD-PUF, consisting of two minimum sized \( nFETs \) \((W \times L = 120 \times 40 \text{ nm}^2)\) and a \( pFET \) selector. Breakdown can be generated on one of the \( nFETs \) randomly. (b) The array with 60 cells, which has been fabricated on a commercial CMOS process.

2. **DC Characteristics**

The forming process and measurement of the \( I_D-V_G \) characteristics were performed using a HP4156 semiconductor parameter analyzer as shown in Fig. 3-2. Forming currents vs. time are shown in Fig. 3-3. The forming step was accomplished by applying 2 V to the input, 1 V on the \textit{word line} WL, and -2 V on the \textit{bit line} BL, and bulk contacts for 5 seconds. A straightforward
method to recognize the device in which breakdown occurs is to compare the current of $BL$ and $\overline{BL}$. In this example, $BL$ experiences breakdown after 0.25 s of stress. Breakdown in the $BL$ nFET is represented by a logical “0”; conversely, if breakdown occurs on the $\overline{BL}$ nFET, it is represented as a logical “1.” After forming process, the current read-out is done by sweeping the input voltage from 0 V to 1.5 V with all other terminals grounded. The behavior of the PUF is shown in Fig. 3-4.

![HP4156 semiconductor parameter analyzer](image)

**Fig. 3-2** HP4156 semiconductor parameter analyzer.
Fig. 3-3 Forming step of a BD-PUF. The $\overline{BL}$ nFET experiences breakdown at ~ 0.25 s.

Fig. 3-4 Input/output currents vs. input voltage.
3. **X-ray Irradiation**

X-ray irradiation was performed on unlidded packaged parts using an ARACOR 4100 10-keV X-ray system with a dose rate of 31.5 krad(SiO$_2$)/min as shown in Fig. 3-5. All terminals of the device under test (DUT) were grounded during exposure. DC characteristics are measured before and after irradiation up to 2 Mrad(SiO$_2$).

![Fig. 3-5 10-keV ARACOR 4100 X-ray irradiator.](image)

4. **Proton Irradiation**

1.8 MeV proton irradiation experiments were conducted using the Pelletron accelerator at Vanderbilt University, as shown in Fig. 3-6. The beam size was sufficient to irradiate the entire die uniformly. The TID levels for proton fluences of $3 \times 10^{13}$, $5 \times 10^{13}$, $7 \times 10^{13}$ and $1 \times 10^{14}$ cm$^{-2}$ are 58, 96, 134 and 192 Mrad(Si), respectively [36]. The irradiation is performed at room
temperature.

Fig. 3-6 Pelletron accelerator at Vanderbilt University.
CHAPTER IV

Radiation Effects on BD-PUFs

Space systems require integrated circuits to perform operations such as protection of ground-to-spacecraft command and control communications in a reliable, inexpensive, and highly secure way. New research and development in the commercial electronics sector on approaches to secure communication between systems may prove to be useful for these low-cost, space-based applications. The current practice in commercial electronic systems is to place a secret key in non-volatile memory, and use cryptographic primitives such as digital signature and encryption to protect confidential information. While analogous approaches may be useful in larger space systems, such approaches are difficult and expensive to implement in low-cost, small-satellite communication systems.

An example of how PUFs might be used in space applications is to encrypt data transmission between spacecraft or between spacecraft and ground stations. To be able to function in this role in space, the PUF must be resilient to the space radiation environment. In this chapter, the radiation response of BD-PUFs is evaluated using 10-keV X-rays and 1.8-MeV protons, and the mechanism is explained.

1. X-ray Irradiation Response

All terminals of the device under test (DUT) were grounded during exposure. The pre-irradiation behavior of the PUF is stable, as shown in Fig. 4-1(a). The currents associated with breakdown show no significant variation during 20 subsequent sweeps. Fig. 4-1(b) shows $I-V$
read curves of the BD-PUF before and after 10-keV X-ray irradiation up to 2 Mrad(SiO$_2$). Less than 11% change in current ratio at 1.2 V was observed with low-dose X-ray exposure, i.e., the BD-PUF stability is not affected significantly by X-ray irradiation.
2. 1.8 MeV Proton Irradiation Response

One BD-PUF has three conditions in the circuits: grounded, read-out and standby conditions. The proton irradiation effects for three conditions are explored in this section.

a. Grounded Condition

BD-PUFs were irradiated with all terminals grounded as shown in Fig. 4-2. Fig. 4-3(a) plots the measured electrical response of the BD-PUF before and after proton exposure. Below a fluence of $3 \times 10^{13}$ cm$^{-2}$, there is no radiation-induced change of the input and BL currents. However, the input currents and BL currents decrease with fluences above $3 \times 10^{13}$ cm$^{-2}$. In contrast, the $\overline{BL}$ current Fig. 4-3(b) increases significantly and already noticeably at the lowest
radiation dose. Note that the BL current, which is quite small, is plotted on a log scale for visibility.

Fig. 4-2 All the terminals of the BD-PUF were grounded during proton irradiation.
Fig. 4-3 (a) Current read-out of BD-PUF, and (b) $I_{BL}$ current before and after 1.8 MeV proton irradiation.

The BL current ($I_{BL}$) is -0.2 $\mu$A at 1.2 V for the pre-irradiation test, as shown in Fig. 4-3(a). For a fixed value of $I_{BL}$, the corresponding input voltage increases as the fluence becomes larger. This increase is characterized as $\Delta V_{BD}$. Fig. 4-4(a) shows $\Delta V_{BD}$ as a function of fluence and the percentage decrease of the BL-current magnitude. $\Delta V_{BD}$ first increases with fluence and subsequently shows signs of recovery. Moreover, after enhanced recovery by annealing at high temperature (100 °C), $\Delta V_{BD}$ returns near its value prior to irradiation.

Memory ratio ($I_{BL}/I_{BL-bar}$) is the crucial application parameter to distinguish between a “0” or a “1,” is extracted. Changes in memory ratio ($I_{BL}/I_{BL-bar}$) shown in Fig 4-4(b). Similar to what is observed with $\Delta V_{BD}$, the memory ratio between the $I_{BL}$ and $I_{BL-bar}$ at 1.2 V decreases with fluence, then partially recovers at room temperature, and finally recovers back to the original value after annealing.
Fig. 4-4 (a) The change of input voltage when the BL current is -0.2 μA, and percentage decrease of BL current at 1.2 V, and (b) memory ratio as a function of fluence and annealing time.
b. Read-out Condition

For the read-out condition, the input voltage was 1.2 V, and the remaining terminals of the BD-PUF were grounded during the irradiation, as shown in Fig. 4-5. Breakdown happened in the left $n$FET (BL) for the device used here. Fig. 4-6 plots the measured electrical response of the BD-PUF before and after proton exposure. There is no significant change with proton irradiation.

Fig. 4-5 The voltage applied in the readout condition of the BD-PUF.
Fig. 4-6 Current read-out of BD-PUF.

c. **Standby Condition**

For most applications, BD-PUFs are in the standby condition for a substantial fraction of the operational time. The voltage applied in the standby condition is shown in Fig. 4-7. Keeping the word line voltage equal to the input voltage turns the pFET selector off. Breakdown happened in the right nFET ($\overline{BL}$) for the device used here. Fig. 4-8(a) shows the measured electrical response of the BD-PUF before and after several proton exposures up to $3 \times 10^{14}$ cm$^{-2}$ fluence. The breakdown current does not change significantly. The current through the unbroken nFET increases significantly as shown in Fig. 4-8(b). It should be noted that the $BL$ current, which is quite small compared with $\overline{BL}$ current, is plotted on a log scale for visibility. Fig. 4-9 plots the memory ratio between the $\overline{BL}$ current and $BL$ current at 1.2 V decreases with fluence, then partially recovers at room temperature, and finally recovers back to the original value after annealing.
Fig. 4-7 The voltage applied in the standby condition of BD-PUFs.
Fig. 4-8 (a) Current read-out of BD-PUF, and (b) BL current before and after 1.8 MeV proton irradiation.

Fig. 4-9 Memory ratio as a function of fluence and annealing time.
3. **Selector-transistor and Leakage Effects**

Fig. 4-10(a) shows the $I_D$-$V_G$ characteristics for proton tests on a $p$FET selector. Threshold-voltage shifts for different proton fluences and annealing times are shown in Fig. 4-10(b). Two $p$FETs were used for the proton measurements, and all package terminals were grounded for the $p$FET selector during exposure. $\Delta V_{th}$ increases significantly with fluence and partially recovers during annealing. The change of the $p$FET $\Delta V_{th}$ here is consistent with the response of $\Delta V_{BD}$ observed in Fig. 4-4(a).
The $I_D-V_G$ characteristics are shown in Fig. 4-11 as a function of proton fluence for broken and unbroken $n$FETs. All package terminals were grounded for the $n$FETs during exposure. The off-state leakage currents of both the broken and unbroken $n$FET increase as the fluence becomes larger. One obvious reason that off-state leakage currents might increase for either the broken or unbroken $n$FET is increased gate leakage current due to proton-induced defect formation \[37,38\]. However, Fig. 4-12 shows the $I_G-V_G$ characteristics for the (a) broken and (b) unbroken $n$FET, and in neither case does the gate leakage current change significantly with proton irradiation. Hence, the increased leakage current in Figs. 4-11(a) and 4-11(b) must have a different origin.
Fig. 4-11. Semi-log plots of $I_D - V_G$ curves for (a) the broken $n$FET, and (b) the unbroken $n$FET as a function of fluence.
$V_D = V_S = V_B = 0$ V

Fig. 4-12. $I_G - V_G$ curves for (a) the broken nFET, and (b) the unbroken nFET as a function of fluence.
Fig. 4-13 shows a schematic diagram of the current through the PUF after its forming step and proton irradiation. Current from the \textit{Input} flows through the \textit{pFET} selector, and then through the parallel combination of the broken and unbroken \textit{nFETs}. The resistance of the broken \textit{nFET} is much lower than that of the unbroken \textit{nFET}, so before and after proton irradiation, nearly all of the current flows through the broken \textit{nFET}. Fig. 4-10(a) shows that the current through the \textit{pFET} selector decreases with proton fluence, as a result of the buildup of radiation-induced charge and the corresponding negative $V_{th}$ shift [39],[40]. This leads to the overall decrease in read-out current of the BD-PUF in Fig. 4-3(a). To understand the increase in the $\overline{BL}$ current in Fig. 4-3(b), we must consider the parallel combination of the broken and unbroken \textit{nFETs} in Fig. 4-13. While the majority of current still flows through the broken \textit{nFET}, an increasing amount of leakage is observed in Fig. 4-11(b) after proton irradiation through the \textit{unbroken} \textit{nFET}, which shares a common body junction with the \textit{broken} \textit{nFET}. This leakage is independent of gate voltage, and due most likely to proton-induced displacement damage and interface traps at the body to \textit{S/D} junctions [36]. The strong correlation of the $\overline{BL}$ current in Fig. 4-3(b) and the increased off-state leakage of the \textit{unbroken} \textit{nFET} therefore suggests that a small percentage (~0.1 to 1\% in this case) of the total current flows through the body-to-\textit{S/D} contacts of the unbroken \textit{nFET} at the highest observed proton fluence. While this does not significantly affect the measured read-out current of the BD-PUF in Fig. 4-3(a), it does account for the increased $\overline{BL}$ current in Fig. 4-3(b).
Fig. 4-13. Current in the BD-PUF after its forming step and proton irradiation. Current from the Input flows through the pFET selector, and then through the parallel combination of the broken and unbroken nFETs. The resistance of the broken nFET is much lower than that of the unbroken nFET, so before and after proton irradiation, nearly all of the current flows through the broken nFET. While the majority of current still flows through the broken nFET after proton irradiation, an increasing amount of leakage current flows through the unbroken nFET, which shares a common body junction with the broken nFET.

Fig. 4-14 shows the $I_D-V_D$ characteristics of the selector and the $I_G-V_G$ curve (load line) of the broken nFET as a function of proton fluence. The gate voltage of the broken nFET is equal to the drain voltage of the pFET selector in the BD-PUF. The cross points shown in Fig. 4-14 are operating points of the BD-PUF in typical circuit operation. The voltage drop across the pFET selector increases as the fluence becomes larger, and the maximum voltage drop is 0.11 V at a fluence of $10^{14}$ cm$^{-2}$, leading to the observed drop in read-out current of the BD-PUF.
Fig. 4-14. The $I_D-V_D$ curves of the $p$FET selector and the $n$FETs load line as a function of fluence.

For the read-out condition of BD-PUFs, the $p$FET selector is on during proton exposure, and large current flows through the $p$FET selector and $n$FETs. For the bias condition with the $p$FET selector during irradiation, the source/input voltage is 1.2 V, the gate/WL voltage is 0 V, and the drain voltage is around 1.1 V according to the load line in Fig. 4-14. The breakdown current through the broken $n$FET does not change significantly. The leakage current through the unbroken $n$FET does not change significantly due most likely to the large current through the BD-PUF during exposure to limit the defects generated in the device. For the standby condition of BD-PUFs during exposure, the $p$FET selector was biased with the source/input voltage of 1.2 V, the gate/WL voltage of 1.2 V, and the drain voltage of around 0 V. The breakdown current
through the *broken nFET* does not change significantly due most likely to the similar reason as the read-out condition. The leakage current through the *unbroken nFET* increases due most likely to the similar reason as the grounded condition we discussed before.

Because the non-ionizing energy loss of 1.8-MeV protons is much higher than that of the higher-energy protons that typically result in the degradation in space systems [41],[42], the equivalent displacement damage doses in this study are quite high compared with most realistic space environments [36]. Thus, these types of BD-PUFs may well exhibit excellent radiation tolerance in most space environments of interest.
CHAPTER V

CONCLUSIONS

The goal of this work was to evaluate the possibility of using a CMOS-based PUF device, which utilizes the randomness of breakdown positions in transistors, to meet the secure communication needs of space systems. 10-keV X-ray and 1.8 MeV proton radiation response of BD-PUFs is reported in this thesis. The device exhibited excellent radiation tolerance to 10-keV X-ray irradiation up to 2 Mrad(SiO$_2$). Less than 11% change in current ratio at 1.2 V was observed.

When irradiated with 1.8 MeV protons, the read-out window of programmed BD-PUFs, which are biased in the grounded condition during exposure, decreases significantly at high dose proton irradiation, and then recovers back to the original value after annealing. The breakdown current through the broken nFET decreases, and the leakage current through the unbroken nFET increases. The current through the pFET selector decreases after proton irradiation, as a result of the buildup of radiation-induced charge and the corresponding negative $V_{th}$ shift. This leads to the decrease in breakdown current of BD-PUFs. In addition, the voltage drop across the pFET selector increases as the fluence becomes larger, also leading to the observed drop in breakdown current. The radiation response of the broken and unbroken nFETs indicates that a small percentage (~0.1 to 1% in this case) of the total current flows through the body-to-S/D contacts of the unbroken nFET at the highest observed proton fluence. While this does not significantly affect the breakdown current, it does account for the increased leakage current of BD-PUFs. When BD-PUFs are biased in the read-out condition during proton exposure, the read-out
window does not have significant change. When BD-PUFs are biased in the standby condition during exposure, the breakdown current shows no significant change and the leakage current increases with larger fluence.

In summary, we have evaluated the radiation response of BD-PUFs in detail. BD-PUFs likely will perform well in typical low-fluence space environments, but their suitability for high-fluence environments must be evaluated carefully, relative to system requirements.
REFERENCES


