Implementation of Stable PUFs Using Gate Oxide Breakdown

Wei-Che Wang, Yair Yona, Yizhang Wu, Szu-Yao Hung, Suhas Diggavi and Puneet Gupta
Department of Electrical Engineering, University of California, Los Angeles
{weichewang, yairyo99, wuyizhang, syhung, suhasdiggavi}@ucla.edu, puneet@ee.ucla.edu

Abstract—**Instability has been an Achilles heel for physical unclonable functions (PUF) requiring complex error correction or other stability enhancement approaches. This instability originates from parametric nature of variations leveraged as a source of randomness. We develop highly stable PUFs using two random gate oxide breakdown mechanisms: plasma induced damage during semiconductor manufacturing and voltage stressed damage post manufacturing. These gate oxide breakdown PUFs can be easily implemented in commercial silicon processes without extra cost on PUF manufacturing and design, and they are stable and resistant to physical attacks. We fabricated bit generation units for the stable PUFs on 99 testchips with 65nm CMOS bulk technology. Measurement results show that the plasma induced breakdown can generate completely stable responses for all 2871 bits and significant area reduction compared with SRAM PUF can be achieved by eliminating the error correction code (ECC) hardware implementation. For the voltage stressed breakdown, the area cost is further reduced, and its 0.12% bit error rate at a worst case corner can be effectively accommodated by taking the majority vote from multiple measurements without ECC. We show that the responses of gate oxide breakdown PUFs are unique. In addition, we analyze the data of our testchips and show through various statistical distance measures that the bits of our fabricated PUFs are independent.**

I. Introduction

Physical Unclonable Functions (PUFs) [1] have been considered as promising security primitives that enables lightweight hardware implementations of identification [2], Trojan detection [3], device tracking [4], [5], authentication [6], or secret key generation [7] and storage [8], [9]. The randomness of a PUF is extracted from random uncontrollable process variations, and its behavior, or Challenge Response Pair (CRP) [10], is uniquely tied to a given device and is hard to predict or replicate. Since the first physical unclonable identification was fabricated in [11], extensive efforts have been devoted into the area, and different silicon PUF implementations have been proposed, including Arbiter PUF [12], [13], Ring Oscillator (RO) PUF [14], memory-based PUF [15], [16], and many other variations.

The instability of a parametric PUF limits the practical application of a PUF. Since these PUFs are parametric, they are in nature susceptible to environmental variations, and the behavior of a PUF can be altered consistently in two different environments. Sources of the instability could be the measurement noise [17], environmental fluctuations [18], or device aging [19]. For example, for a RO pair of a RO PUF, one RO can have a faster frequency than another RO at 20°C but a slower frequency at 100°C [13]. To make a PUF more stable, extra overhead is required, including hardware or latency cost [20]. Techniques such as error correction code (ECC) or helper data come with the cost of extra hardware implementation or possible security concerns [21].

Recently, a stability-guaranteed Locally Enhanced Defectivity PUF (LEDPUF) proposed in [22] shows completely stable responses by utilizing random hard defect generated from Directed Self Assembly (DSA) process. However, it is difficult to fabricate given that DSA is not well accepted into commercial silicon manufacturing yet. In [23] a reliable RRAM PUF with actual PUF fabrication using Resistive Random Access Memory (RRAM) is presented. However, an off-chip characterization of the split current and offset for the sense amplifiers are required, and the reliability results under voltage variations are not reported, which can dramatically impact the stability. Another reliable PUF using Hot Carrier Injection (HCI) is presented in [24]. However, post calibration steps are still needed and the randomness of the most stable responses was not reported. In this paper we implement and analyze stable source of randomness from silicon gate oxide breakdown. The contributions of the paper are given as follows:

- We propose two mechanisms exploiting gate oxide breakdown as the source of randomness for stable PUFs: the plasma induced oxide breakdown and the voltage stressed oxide breakdown. The oxide breakdown PUFs are resistant to invasive attacks such as imaging attacks.
- Test structures violating antenna rules are fabricated with 65nm CMOS bulk technology. Measured results from 99 testchips show that the responses are highly stable across combinations of voltage (0.8V, 1.0V, 1.2V) and temperature variations (25°C, 100°C). Compared to a practical SRAM PUF, significant area reduction can be achieved by eliminating ECC implementation for the highly stable responses.
- We analyze the data from these testchips and show based on various statistical distance measures that pairs of bits with the same antenna ratio as well as bits that are located next to each other are effectively statistically independent.

II. Stable PUFs Exploiting Gate Oxide Breakdown

In this section, we first introduce the gate oxide breakdown and describe two approaches exploiting the gate oxide breakdown as randomness sources of stable PUFs, followed by PUF bit generation and attack resilience analysis.

The work of Y. Yona and S. Diggavi was supported in part by NSF grants 1321120 and 1423271.

978-1-5386-1421-1/17/$31.00 ©2017 IEEE

A. Gate Oxide Breakdown

The gate oxide breakdown is detrimental to metal-oxide-semiconductor (MOS) devices because it can cause significant drifts of transistor parameters. The breakdown can be categorized into two types: soft breakdown and hard breakdown, where both mechanisms introduce significant sudden increase of the leakage current [25]. For soft breakdown, the conducting path from gate to the substrate is formed by the charged traps in the gate oxide. Once there is conduction, new traps begin to accumulate due to thermal damage, which in turn increases the conductance. The positive feedback eventually leads to thermal runway and oxide is physically melt in the breakdown spot. This type of breakdown is called hard breakdown. The gate leakage current of an oxide with breakdown can be 100X larger than the leakage current of an oxide without breakdown [26].

B. Plasma Induced Gate Oxide Breakdown

During silicon wafer fabrication, radio frequency (RF) plasma processes are widely used for etching, photoresist stripping, or ion implantation [27]. In the plasma ambient, metal segments, VIAs, or polysilicon electrodes, which are the antenna segments, can be electrically charged by ions or electrons when the currents produced from the ion and electron do not cancel out with each other through each RF cycle [28], and therefore produce the antenna voltage. For the antenna segments connected to the gate inputs, the resulting electrical stress from the antennas can potentially damage the underlying gate oxide and create a conducting path from the gate to the substrate. The phenomenon is called plasma induced gate oxide breakdown, or the antenna effect.

Though the maximum voltage rise over half of the RF period can be modeled [27], the actual voltage still cannot be predicted because the exact motion and amounts of ions and electrons collected by the antenna segment are random and unpredictable. The higher the gate voltage is, the higher the probability for the gate oxide breakdown to occur, thus causing a device to fail. Also, systematic plasma variation across wafer does not have much impact on the local randomness because the variation is negligible to a die [27].

To avoid the antenna effect, design rules of the antenna ratio (AR) [29] as shown in equation (1) must be strictly followed during fabrication [30]. Practical design rules of AR range from 100 to 5000 depending on the process details [29].

$$AR = \frac{\text{exposed antenna area}}{\text{gate oxide area}} \tag{1}$$

Since both soft breakdown and hard breakdown can induce about 100X or more leakage current than a good oxide, they are both considered as breakdown in our proposed stable PUF construction. In [31], a device is considered as a failure if the gate leakage current is larger than 1nA, and based on the criterion the author proposed a failure probability prediction formula. However, the process parameters of our testchip fabrication are unknown prior manufacturing therefore we implemented a variety of antenna ratios to measure breakdown probabilities, which are presented in Section III-B.

Many techniques have been proposed to solve antenna effect, such as jumper insertion [32] or antenna-aware routing [33]. However, while foundries try to avoid antenna effect during manufacturing, we exploit the uncontrollable physical phenomena as another randomness source of a stable PUF.

C. Voltage Stressed Gate Oxide Breakdown

The purpose of antenna rules is to protect all transistors from having deviated parameters, for example 20% gate leakage increase at 1.4xVDD [34], which could be harmful for a normal fabrication but still far from causing a real breakdown. Therefore, to introduce a noticeable plasma induced breakdown (100X increase of leakage current) with 50% probability of a transistor, an AR larger than 1000X antenna rule may be required, which can result in large area overhead.

To avoid using large antenna segments, we propose to induce gate oxide breakdown post fabrication by applying high voltage stress to the gate of a transistor that essentially mimics the charge accumulation during the plasma process. By voltage stressing the gate terminal of a transistor, oxide breakdown can be introduced with small AR or even without violating the antenna rules. The advantage of voltage stressed induced breakdown is that large antenna segments are not required, while the uncontrollable process variation of gate oxide thickness is magnified to achieve a breakdown probability close to 50%, which is desirable as a source of randomness for PUFs. On the other hand, such a PUF construction requires an additional one-time stress step post manufacturing (or during PUF enrollment). Please note that our proposed voltage stressed gate oxide breakdown mechanism is different from the Erasable PUF proposed in [35], where oxide breakdown is introduced to erase targeted bit cells instead of being used as a stable source of randomness.

D. Stable Signal Unit Construction

The permanent gate oxide breakdown mechanism, which can be caused by plasma damage or voltage stressed damage, is used to construct a Stable Signal Unit (SSU) as a source of permanent defectivity. A SSU is a p-MOS transistor designed to violate antenna rules, and its drain, source, and bulk terminals are connected to capture the effect of the gate oxide breakdown at all possible locations. Similar to a gate oxide breakdown model given in [36], the SSU is attached in series to a precision resistor as given in Fig. 1, where Fig. 1 (a) shows a SSU without oxide breakdown and Fig. 1 (b) shows a SSU with oxide breakdown. If no breakdown occurs as depicted in Fig. 1 (a), the device is essentially a capacitor or a resistor much larger than the precision resistor, thus the output voltage would be lower than 50% VDD when the evaluation signal EVA is VDD; if a breakdown happens, as shown in Fig. 1 (b), the device can be seen as resistors much smaller than the precision resistor, thus the output voltage would be higher than 50% VDD when EVA is VDD. The resistance of the precision resistor (10MΩ) is determined by actual measurements from 99 testchips as described in Section III-B. Different from the bit generation units in [37], our SSU does not suffer from potential response time latency due to the limited leakage current when no breakdown occurs.

E. Attack Resilience

It is worth mentioning that the SSU is more secure than an antifuse cell because an antifuse cell is programmed with

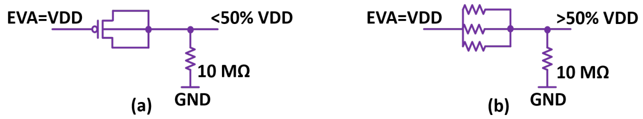

Fig. 1. Schematic of antenna SSU attached to a precision resistor.

hard breakdown only, while the output of the SSU is decided by both soft breakdown and hard breakdown, and a soft breakdown is much harder to detect than a hard breakdown (albeit possible for a very resourceful attacker). For probing attack, the efficiency is limited by the mechanical constraints. For imaging attacks, such as Scanning Electron Microscopy (SEM), Transmission Electron Microscopy (TEM), or Electron Beam Induced Currents (EBIC), it is difficult to efficiently identify a soft breakdown for several reasons:

1) It is difficult to detect a soft breakdown because its physical appearance is very similar to a fresh gate oxide without any visible holes. Furthermore, SEM has limited ability to observe traps inside the oxide, therefore it is difficult to see if a conducting path formed by traps, or a soft breakdown, exists. It is also challenging for EBIC to identify a soft breakdown because the limited current of a soft breakdown can induce measurement noises [38], and the throughput of the electron beam is low.
2) It is difficult to observe a soft breakdown from a top-down or cross-section TEM because the image does not effectively tell the depth of the traps [39]. In addition, to obtain a cross-section TEM, the chip has to be vertically cut into thin films, which will destroy the neighboring SSUs. Therefore, even if a hard breakdown information might be retrieved from a cross-section view, the attacker cannot obtain the secrets of all SSUs of a same PUF because of the destructive observation.

III. Testchip Fabrication and Measurement Results

A. SSU Implementations

The proposed SSUs are implemented and fabricated on 99 testchips with commercial 65nm GP 1P9M_6X1Z1U CMOS bulk technology with 1V nominal voltage. The smallest gate size (0.0072μm^2) of the technology is used for all the SSUs. In our testchips the fabricated SSUs intentionally violate antenna rules by a few hundred times to a few thousand times on different layers.

On each chip, 29 SSUs are implemented with 17 different ARs, therefore the total number of SSU implementations is 2871 from 99 chips. For each of the SSUs, the cell area and detailed antenna violation report are given in Table I, where a zero indicates that there is no antenna rule violation on such layer. The antenna rule violation reports are provided to the foundry to skip such design rule checks without extra cost for the foundry. The M_T, V_T, and P_T structures test the effects of metal, VIA, and polysilicon layers from small AR to large AR, respectively. For each of the M_T, V_T, and P_T, two SSUs with same AR are implemented, therefore 24 bits of responses are obtained from these SSUs on a chip. The remaining five test structures are of various combinations of the violating layers, and one SSU is implemented for each of the five test structures. In summary, on each chip, 29 bits are measured, and 24 bits of them are obtained from the duplicated 12 structures of M_T, V_T, and P_T.

TABLE I
Cell area, accumulated areas of VIA, metal, polysilicon, and polysilicon perimeter of SSUs fabricated. The numbers are in μm^2. A zero indicates no antenna rule violation on such layer.

	Cell	VIA	Metal	Poly	Poly Perim. (μm)
M_T1	36	0.87	1144.57	0.00	0.00
M_T2	360	1.17	1468.57	0.00	0.00
M_T3	1200	0.00	4398.88	0.00	0.00
M_T4	4800	0.16	36781.89	0.00	0.00
V_T1	2.4	0.87	1108.57	0.00	0.00
V_T2	8	2.31	1108.57	0.00	0.00
V_T3	90	15.27	1185.66	0.00	0.00
V_T4	804	144.91	1895.05	0.00	0.00
P_T1	4.8	1.26	1917.53	0.00	0.00
P_T2	27	1.26	1917.53	18.17	55.59
P_T3	203	1.26	1917.53	180.07	128.43
P_T4	1800	1.26	1917.53	1800.07	222.46
Test1	804	1071.86	5631.11	0.00	0.00
Test2	4.7	1.86	0.00	0.00	0.00
Test3	80	0.26	299.20	0.00	0.00
Test4	60	20.84	318.78	28.07	83.81
Test5	118	54.40	617.25	56.39	164.72

B. Breakdown Probability Evaluation

To determine the gate oxide breakdown of a SSU, we use Agilent 34411A Digital Multimeter to measure the equivalent resistance R_{eq} of each SSU, and from the distribution of R_{eq} we choose a proper precision resistor as shown in Fig. 1 to determine whether or not an oxide breakdown has occurred. Fig. 2 shows R_{eq} distribution of a SSU implementation (V_T1) with plasma induced and voltage stressed breakdown on 99 chips in an increasing order at 25°C, 1V. For both distributions, the R_{eq} of a SSU implementation without oxide breakdown is at least 100X larger than a SSU with oxide breakdown. After voltage stress, the R_{eq} are in general smaller and much more oxide breakdowns are introduced. The results are similar for all SSUs. The large gap in the figure can be effectively exploited to generate stable digital signals from SSUs. Therefore, we choose, according to the R_{eq} measurements, a 10MΩ precision resistor to measure the gate oxide breakdown of each SSU.

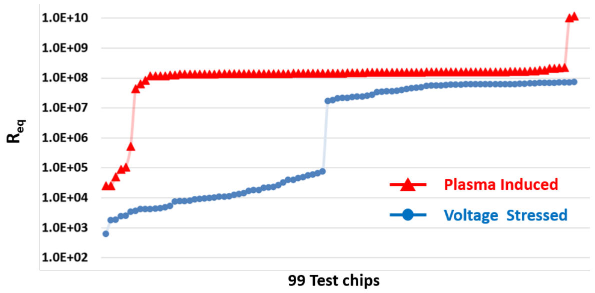

Fig. 2. The R_{eq} distribution of a SSU implementation (V_T1) with plasma induced and voltage stressed oxide damage on 99 chips at 25°C, 1V.

1) Plasma Induced Breakdown: For the plasma induced breakdown, the results of breakdown probabilities of SSU implementations on 99 chips are shown in Table II. From the

TABLE II
BREAKDOWN PROBABILITY OF 17 AR IMPLEMENTATIONS ON 99 TESTCHIPS.

	Plasma Induced	Voltage Stressed
M_T1	0.5%	57.6%
M_T2	0.5%	51.5%
M_T3	2.5%	57.1%
M_T4	2.0%	51.0%
V_T1	0.5%	50.0%
V_T2	6.1%	54.0%
V_T3	0.0%	64.7%
V_T4	0.0%	58.6%
P_T1	1.0%	50.5%
P_T2	2.5%	51.5%
P_T3	1.0%	58.6%
P_T4	1.0%	60.0%
Test1	16.2%	N/A
Test2	2.0%	N/A
Test3	5.1%	N/A
Test4	1.0%	N/A
Test5	3.0%	N/A

table we see that the breakdown probability of each SSU after plasma induced oxide damage is well below 50%. The Test1 SSU implementation has the highest breakdown probability of 16%, which means the responses of SSUs are highly biased. This is undesirable for its low randomness in each response bit. Using larger AR to further increase the breakdown probability may not be a proper approach due to large area overhead. Also, as seen from Table I and Table II, the breakdown probability does not increase dramatically as the AR increases. Our results show that even when the AR is more than 1000X larger than the antenna rule, the breakdown probability is still much lower than 50%.

2) Voltage Stressed Breakdown: For the voltage stressed breakdown, we stress 24 SSUs (M_T, V_T, and P_T groups) on each testchip by applying 5.5V to the EVA for 10 seconds. The results of the stress are shown in Table II. From the table we can see that breakdown probabilities, which are only slightly correlated with the ARs, are elevated to at least 50% even for the SSUs with the smallest ARs. Different stress voltages have been tried in our experiments, but only when the voltage is 5.5V will the breakdown probability be elevated to 50%. These results show that more unbiased responses compared to plasma induced breakdown can be achieved by using small SSUs such as V_T1. Therefore, a SSU can be implemented with much smaller area, possibly even without violating the antenna rule, than the plasma induced breakdown approach.

C. Stability Evaluation

To evaluate the stability of the SSUs, we measure all SSU responses from 99 chips at 6 corners: temperatures at 25°C and 100°C with ±20% voltage variation at 0.8V, 1V, and 1.2V.

1) Plasma Induced Breakdown: For the plasma induced breakdown, all SSUs from 99 chips (total 2871 bits generated) are completely stable at all corners during multiple measurements. This can be explained by the fact that the change of R_{eq} at different corners are limited. Fig. 3 shows the change of R_{eq} of a SSU (Test1) under voltage and temperature variations. In Fig. 3 (a), the R_{eq} of the SSU with breakdown is only a few KΩ and the changes under extreme temperature and voltage variations are limited. On the other hand, Fig. 3 (b) shows a SSU without oxide breakdown, where the R_{eq} remains at less than 45MΩ, which is still orders of magnitude larger than the SSU with oxide breakdown.

TABLE III
BIT ERROR RATES OF 2376 SSUS OF THE VOLTAGE STRESSED BREAKDOWN AT 6 CORNERS.

Corners	0.8V	1V	1.2V
25°C	0.04%	0.00%	0.12%
100°C	0.08%	0.08%	0.08%

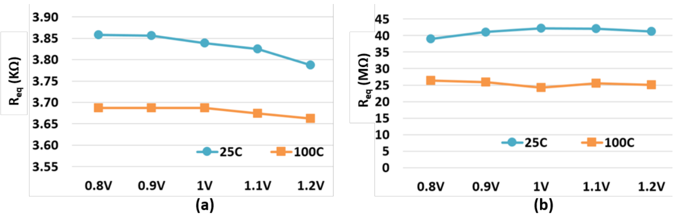

Fig. 3. Equivalent resistance under extreme voltage and temperature variations. (a) SSU with oxide breakdown. (b) SSU without oxide breakdown.

2) Voltage Stressed Breakdown: Unlike the plasma induced breakdown, for the voltage stressed breakdown, an extremely small portion of the SSUs are not completely stable. To quantize the results of stability evaluation for the voltage stressed breakdown, each SSU is measured 10 times at each corner and we define the responses measured at 25°C with 1V, where all responses are consistent, as the reference responses. A SSU is unstable at a corner if at least one of its values from the 10 measurements is different from the reference response. We define bit error rate (BER) the number of unstable bits divided by 2376, which is the total number of SSUs stressed (24 SSUs on each of the 99 chips). Table III shows the BER at each corner. We found that at several corners, 1 to 3 SSUs out of 2376 SSUs implemented are unstable for the voltage stressed breakdown. Since most responses of unstable SSUs are still consistent with the reference responses, instead of performing a ”afterburn” phase to all broken oxides, where additional hardware and calibration are required [40], we take majority vote of multiple measurements to effectively eliminate the erroneous responses.

D. Uniqueness Evaluation

The inter-Fractional Hamming Distance (FHD) [41] is calculated as the uniqueness evaluation of SSUs. Consider the 24 voltage stressed SSUs on each chip as a 24-bit weak PUF [42], the distribution of inter-FHD of 99 chips are presented in Fig. 4. The average of inter-FHD is 51.7% and the standard deviation is 11.4%, where for an ideal Binomial distribution with success probability P=0.5, the mean is 50% and the standard deviation is 10.2%. Please note that the results of uniqueness evaluation are focused on the voltage stressed breakdown SSUs because for the plasma induced breakdown SSUs, the responses are highly biased and post processing would be required to extract randomness, for example using OR gates at the outputs of multiple SSUs to generate an unbiased bit as explained in Section IV-A.

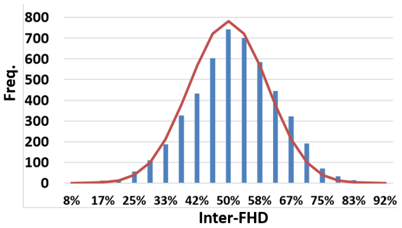

Fig. 4. Inter-FHD distribution of voltage stressed SSUs on 99 chips overlaid with an ideal Binomial distribution curve with success probability P=0.5.

TABLE IV

STATISTICAL DISTANCES BASED ON THE COLLECTED DATA. IN EACH ENTRY THE LEFT SIDE REPRESENTS THE STATISTICAL DISTANCE OF BITS THAT ARE LOCATED NEXT TO EACH OTHER, WHEREAS THE RIGHT SIDE REPRESENTS THE DISTANCE OF BITS THAT HAVE THE SAME ANTENNA RATIO.

Statistical Distance	Max	Min	Mean
KL	0.11/0.057	0.0002/0.0001	0.022/0.015
TVD	0.19/0.13	0.009/0.007	0.07/0.05
GW	0.06/0.029	0.0001/0.00009	0.011/0.008

E. Statistical Analysis of the PUF Responses

In this section we provide a statistical analysis for the data of the fabricated SSUs after voltage stressed oxide breakdown as presented in Section III. We evaluate the statistical dependence between pairs of bits using various statistical distance measures. We consider pairs as we have only 99 bits per location, and so going beyond the pairwise probability mass function can lead to more noisy and less reliable evaluation. We are interested in the level of independence because the more independent the bits are, the more secure the PUF is.

Essentially, we use that data to evaluate the pairwise probability mass functions of bits under the following two restrictions: The pairwise probability mass function of bits that have the same antenna ratio; the pairwise probability mass function of bits that are located next to each other. This in turn enables us to evaluate the statistical dependence of element that are more likely to be statistically dependent, that is, statistical dependence due to similar design rules as well statistical dependence between PUFs that are close together.

We calculate the distance between the evaluated probability mass function (i.e., $P_{X,Y}\left(x,y\right)$) and an independent one with the same marginal probability mass functions (i.e., $P_X\left(x\right) \cdot P_Y\left(y\right)$) by assigning them to various statistical distance measures. This enables us to demonstrate the level of independence between pairs of bits. The results are presented in Table IV for the following statistical distance measures: The Kullback-Leibler (KL) divergence [43]; total variation distance (TVD) [44]; and guesswork (GW) [22].

Table IV shows that the average statistical distance between $P_{X,Y}\left(x,y\right)$ and $P_X\left(x\right) \cdot P_Y\left(y\right)$ is very small across measures, which indicates that this PUF response is very close to being statistically independent.

IV. GATE OXIDE BREAKDOWN PUF IMPLEMENTATIONS

A. Plasma Induced Breakdown PUF

Our measurement results show that the probability of plasma induced breakdown due to antenna rule violation is much lower than the ideal 50%, which means that most responses are zeroes. To reduce the bias, we propose to use OR gates at the output of SSUs as a more area-efficient approach than using even larger antenna segments, which shows limited impact on increasing the breakdown probability. Fig. 5 (a) shows an exemplary implementation of plasma induced breakdown PUF. The 10MΩ precision resistor is shared between two SSUs, where only one of EVA$_1$ and EVA$_2$ will be asserted. Please note that a precision resistor can be shared by more than two SSUs but only one of the SSUs is asserted at a time. The outputs of buffer gates are determined by the breakdown of the SSU.

Take Test3 as an example. When 11 Test3 SSUs are ORed together, the probability of generating a zero is $(1-5.1\%)^{11} = 56\%$, and the area is $880\mu m^2$, which is still more area-efficient than a practical SRAM PUF implementation where (511,19,119)-BCH is suggested to correct 15% error probability at different corners [45]. For such SRAM PUF to generate 19 information bits, the estimated BCH implementation is 12000 XOR gates [46] or an area of $54000\mu m^2$ for the 65nm technology we used. To generate the same number of 19 bits of response with Test3, the estimated area is about $16720\mu m^2$. The comparison shows that the *SRAM PUF is more than 3X of size of the plasma induced breakdown PUF*. In addition, the ECC execution latency is eliminated for the plasma induced breakdown PUF.

B. Voltage Stressed Breakdown PUF

The probability of voltage stressed breakdown is much higher than the plasma induced breakdown, therefore no OR gates are needed to reduce the response bias, but a stress path for each SSU is required. Fig. 5 (b) shows an exemplary implementation of voltage stressed breakdown PUF. A precision resistor is shared by 3 SSUs. Before response generation, the PUF is stressed through the stress path and outputs of SSUs are connected to GND with all EVA signals set to zero. Once SSUs are stressed, a normal voltage is applied to the stress path and one of the EVAs is asserted at a time for evaluation. To generate a bit, approximately 1 inverter and 4 transistors are needed, which translates to an area of only $4\mu m^2$ for 65nm technology. The PUF can be stressed on chip, for example with a charge pump with an area overhead of $12200\mu m^2$ [47]. Therefore, to generate 19 bits of response, the total area is approximately $12276\mu m^2$, which is about 30% smaller than the plasma induced breakdown PUF. As the number of bits increases, the area reduction becomes more evident since the charge pump is shared among multiple bits. The PUF can also be stressed from outside of the chip to save even more area, but an antifuse cell may be needed at the stress path. To stress the PUF, the antifuse cell has to be permanently programmed to closed state. Therefore, if the antifuse cell is already in closed state before stress, it means that the PUF has been contaminated and should be discarded. Please note that if the PUF is stressed from outside of the chip, an attacker may destroy the PUF or introduce more breakdowns by further stressing the PUF, but the PUF is not programmable or clonable because the breakdown of each transistor cannot be controlled.

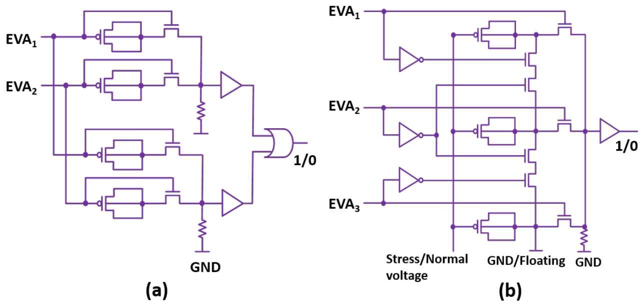

Fig. 5. (a) Plasma induced breakdown PUF implementation. (b) Voltage stressed breakdown PUF implementation.

V. Conclusion

In this paper we implement highly stable PUFs exploiting uncontrollable plasma induced and voltage stressed gate oxide damage. The proposed SSUs are fabricated and measured from 99 testchips. Measurement results show that the SSUs are highly stable, therefore significant area reduction can be achieved by eliminating ECC implementation. We show that the responses are unbiased and unique, and we analyze the data of our testchips using various statistical distance measures to show that these bits are independent.

References

[1] B. Gassend, D. Clarke, M.V. Dijk, and S. Devadas. Silicon physical random functions. In *Proc. CCSC*, 2002.

[2] A. R. Krishna et al. MECCA: a robust low-overhead PUF using embedded memory array. In *CHES*, Oct 2011.

[3] M. Tehranipoor and F. Koushanfar. A Survey of Hardware Trojan Taxonomy and Detection. *IEEE Design and Test of Computers*, Jan 2010.

[4] K. Yang, D. Forte, and M. M. Tehranipoor. UCR: An unclonable chipless RFID tag. In *IEEE HOST*, May 2016.

[5] Yu Zheng, A. Basak, and S. Bhunia. CACI: Dynamic current analysis towards robust recycled chip identification. In *Proc. DAC*, June 2014.

[6] W. Che, F. Saqib, and J. Plusquellic. PUF-based authentication. In *Proc. ICCAD*, Nov 2015.

[7] M.T. Rahman, D. Forte, J. Fahrny, and M. Tehranipoor. ARO-PUF: An aging-resistant ring oscillator PUF design. In *Proc. DATE*, March 2014.

[8] C. Herder, M. Yu, F. Koushanfar, and S. Devadas. Physical Unclonable Functions and Applications: A Tutorial. *Proc. of the IEEE*, Aug 2014.

[9] C. E. D. Yin and G. Qu. LISA: Maximizing RO PUF's secret extraction. In *IEEE HOST*, June 2010.

[10] R. Maes and I. Verbauwhede. Physically Unclonable Functions: A Study on the State of the Art and Future Research Directions. In *Towards Hardware-Intrinsic Security*. 2010.

[11] K. Lofstrom, W. R. Daasch, and D. Taylor. IC identification circuit using device mismatch. In *Proc. ISSCC*, Feb 2000.

[12] J.W. Lee et al. A technique to build a secret key in integrated circuits for identification and authentication applications. In *IEEE International Symposium on VLSI Circuits*, 2004.

[13] G.E. Suh and S. Devadas. Physical Unclonable Functions for Device Authentication and Secret Key Generation. In *Proc. DAC*, 2007.

[14] A. Maiti and P. Schaumont. Improving the quality of a Physical Unclonable Function using configurable Ring Oscillators. In *International Conference on FPL*, Aug 2009.

[15] D.E. Holcomb, W.P. Burleson, and K. Fu. Power-Up SRAM State as an Identifying Fingerprint and Source of True Random Numbers. *IEEE Transactions on Computers*, 2009.

[16] C. Jaeger et al. Random pn-junctions for physical cryptography. *Applied Physics Letters*, April 2010.

[17] L. Portmann and Teresa H.-Y. Meng. Metastability in CMOS library elements in reduced supply and technology scaled applications. *IEEE JSSC*, 1995.

[18] Lang Lin, S. Srivathsa, D.K. Krishnappa, P. Shabadi, and W. Burleson. Design and Validation of Arbiter-Based PUFs for Sub-45-nm Low-Power Security Applications. *IEEE TIFS*, 2012.

[19] D. Ganta and L. Nazhandali. Study of IC Aging on Ring Oscillator Physical Unclonable Functions. In *IEEE ISQED*, 2014.

[20] M. Majzoobi, F. Koushanfar, and S. Devadas. FPGA PUF using programmable delay lines. In *IEEE International Workshop on WIFS*, Dec 2010.

[21] J. Delvaux and I. Verbauwhede. Key-recovery attacks on various RO PUF constructions via helper data manipulation. In *Proc. DATE*, March 2014.

[22] W. Wang, Y. Yona, S. Diggavi, and P. Gupta. LEDPUF: Stability-Guaranteed Physical Unclonable Functions through Locally Enhanced Defectivity. In *IEEE HOST*, May 2016.

[23] R. Liu, H. Wu, Y. Pang, H. Qian, and S. Yu. A highly reliable and tamper-resistant RRAM PUF: Design and experimental validation. In *IEEE HOST*, pages 13–18, May 2016.

[24] Mudit Bhargava and Ken Mai. A High Reliability PUF Using Hot Carrier Injection Based Response Reinforcement. In *CHES*, Aug 2013.

[25] J. Sune, G. Mura, and E. Miranda. Are soft breakdown and hard breakdown of ultrathin gate oxides actually different failure mechanisms? *IEEE Electron Device Letters*, April 2000.

[26] Sang U. Kim. Analysis of Thin Gate Oxide Degradation During Fabrication of Advanced CMOS ULSI Circuits. *IEEE Transactions on Electron Devices*, 1998.

[27] H. C. Shin and Chenming Hu. Thin gate oxide damage due to plasma processing. *Semiconductor Science and Technology*, Apr. 1996.

[28] Sychyi Fang and James P. McVittie. A model and experiments for thin oxide damage from wafer charging in magnetron plasmas. *IEEE Electron Device Letters*, June 1992.

[29] Z. Wang et al. Strategies to cope with plasma charging damage in design and layout phases. In *International Conference on ICICDT*, May 2005.

[30] P.J. Liao et al. Physical origins of plasma damage and its process/gate area effects on high-k metal gate technology. In *IEEE IRPS*, April 2013.

[31] Zhichun Wang. Detection of and Protection against Plasma Charging Damage in Modern IC Technology. In *Ph.D. thesis, Enschede*, Sep. 2004.

[32] Jia Wang and Hai Zhou. Optimal Jumper Insertion for Antenna Avoidance Considering Antenna Charge Sharing. *IEEE TCAD*, Aug. 2007.

[33] Tsung-Yi Ho, Yao-Wen Chang, and Sao-Jie Chen. Multilevel routing with jumper insertion for antenna avoidance. In *Proc. IEEE International SOC Conference*, Sept 2004.

[34] P. J. Liao others. Physical origins of plasma damage and its process/gate area effects on high-k metal gate technology. In *IEEE IRPS*, April 2013.

[35] Ulrich Rührmair, Christian Jaeger, and Michael Algasinger. An attack on puf-based session key exchange and a hardware-based countermeasure: Erasable PUFs. In *International Conference on FC*, 2011.

[36] Kyung Ki Kim. Reliable CMOS VLSI Design Considering Gate Oxide Breakdown.

[37] F. Tang et al. CMOS On-Chip Stable True-Random ID Generation Using Antenna Effect. *IEEE Electron Device Letters*, Jan 2014.

[38] Jun Chen et al. Electron-Beam-Induced Current Study of Breakdown Behavior of High-K Gate MOSFETs. *Solid State Phenomena*, 2010.

[39] Nobuo Tanaka. Present status and future prospects of spherical aberration corrected TEM/STEM for study of nanomaterials. *Science and Technology of Advanced Materials*, 2008.

[40] N. Liu, S. Hanson, D. Sylvester, and D. Blaauw. OxID: On-chip one-time random ID generation using oxide breakdown. In *Symposium on VLSI Circuits*, June 2010.

[41] V. Gunreddy A. Maiti and P. Schaumont. A Systematic Method to Evaluate and Compare the Performance of Physical Unclonable Functions. In *Embedded Systems Design with FPGAs*, 2013.

[42] U. Rührmair and D. E. Holcomb. PUFs at a glance. In *Proc. DATE*, March 2014.

[43] T.A. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley& Sons, 2006.

[44] J. E. Kennedy and M. P. Quine. The total variation distance between the binomial and poisson distributions. *Ann. Probab.*, Jan 1989.

[45] J. Guajardo et al. FPGA Intrinsic PUFs and Their Use for IP Protections. In *CHES*, Sep 2007.

[46] X. Zhang. VLSI Architectures for Modern Error-Correcting Codes. 2015.

[47] X. Li, H. Zhong, Z. Tang, and C. Jia. Reliable Antifuse One-Time-Programmable Scheme With Charge Pump for Postpackage Repair of DRAM. *Proc. VLSI Design*, Sep 2015.