This paper presents a strong silicon physically unclonable function (PUF) immune to machine learning (ML) attacks. The PUF, termed the subthreshold current array (SCA) PUF, is composed of a pair of two-dimensional transistor arrays and a low-offset comparator. The fabricated PUF chip allows 2^65 challenge-response pairs (CRPs) and achieves high reliability with an average bit error rate (BER) of 5.8% for temperatures -20 to 80°C and VDD ± 10%. The calibration-based CRPs filtering method effectively improves BER to 2.6% with a 10% loss of CRPs. When subjected to ML attacks, the PUF shows resilience that is 100X higher than known alternatives, with negligible loss in PUF unpredictability.

**Keywords**: PUF, SCA, security, nonlinearity and attacks

### Introduction

PUFs have great promise as hardware authentication primitives due to their physical unclonability, high resistance to reverse engineering, and difficulty of mathematical cloning. Strong PUFs are distinguished by an exponentially large number of CRPs, in contrast with weak PUFs that have a smaller CRP set. Because the adversary cannot create an enumeration clone (by recording all CRPs) even when in physical possession of a PUF, strong PUFs enable secure direct enumeration clone (by recording all CRPs) even when in physical possession of a PUF, strong PUFs enable secure direct cloning (by recording all CRPs) even when in physical possession of a PUF.

In this paper, we present the first ever silicon implementation of a strong PUF resilient to ML attacks. We fabricated a 2^65 CRP SCA-PUF in 130nm CMOS. When used in conjunction with reliability-driven dynamic thresholding, the BER of only 2.6% is achieved for temperatures -20 to 80°C and VDD ± 10% at 90% CRP utilization. When subjected to a suite of ML attacks, the PUF shows resilience that is 100X higher than known alternatives, with negligible loss in PUF unpredictability.

**SCA-PUF: Principle of Operation**

The basic building block of SCA-PUF is a pair of nominally identical n×k two-dimensional transistor arrays (Fig. 1) with all devices subject to stochastic variability operating in subthreshold region. Under impact of process randomness, the output voltages of two arrays differ and are converted to a binary output by a comparator. Each array consists of k columns and n rows of a unit cell (Fig. 2). A unit cell consists of a stochastic PFET (Mij) that always operates in the subthreshold region and a non-stochastic switch transistor (Mc) arranged in parallel to the stochastic PFET. The term “stochastic transistor” refers to a device with high amount of threshold voltage variability (achieved by using a minimum-size transistor). Switch transistors are controlled by an external challenge word, leading to 2^nk possible output voltages. The transistors Mc and Mij act as current sources and form the basis of the subthreshold current array.

### Proposed Circuit Design

We built a 65-bit-input SCA-PUF based on 5×13 array architecture (Fig. 2). The array analog outputs are Vout_a and Vout_b. The distribution of ΔVout = Vout_a - Vout_b determines the allowable maximum offset of the comparator. The measured ΔVout distribution has μ = 17.9mV, σ = 13.6mV (Fig. 5a). We design the comparator to have input offset Vio < 100μV, which guarantees that 99.6% of CRPs are correctly resolved. This low input offset is achieved by using an active offset cancellation scheme (offset is cancelled by adjusting Vcalp and Vcaln of the comparator (Fig. 6)).

An innovative circuit stabilizes PUF output. In Fig. 2, if node X is directly connected to VDD, the common-mode output voltage Vout_CM = (Vout_a + Vout_b)/2 is highly sensitive to the challenge (causing it to vary from 300mV to 900mV). This complicates comparator offset cancellation because it is feasible only when Vout_CM is well controlled. To achieve a fixed Vout_CM, we separate node X from VDD by inserting a PFET M3, whose gate voltage Vg is generated by a novel common-mode feedback circuit (Fig. 7a). With a proper size ratio between M1, M2, and M3, it is ensured that (Iout + Iout2)/2 = Iout, making (Vout_a + Vout_b)/2 = Vcalp. As a result, Vout_CM = (Vout_a + Vout_b)/2 ≈ Vref, regardless of the challenge (Fig. 5b).

### Results and Discussion

Inter-die Hamming distance (HD) and Hamming weight (HW), quantifying PUF output uniqueness and randomness, respectively, are measured across 50 dies with 500 challenges. For ideal PUFs, both inter-die HD and HW are 0.5. For our PUF, the average normalized inter-die HD is 0.499 (σ = 0.043) (Fig. 7a), and the average HW is 0.528 (σ = 0.109) (Fig. 7b).

Intra-die HD, used as a measure of BER to quantify temporal stability, is measured across 5 dies with 500 challenges, under operating conditions of -20 to 80°C and 1.08-1.32V VDD. The average intra-die HD is 0.058 (σ = 0.038) (Fig. 7a). Distinct effects of temperature and VDD on BER are studied in Fig. 8, which shows that high temperature affects bit reliability most. The average BER in the worst case is 9%.

Dynamic thresholding is used to further reduce BER. We developed an efficient indirect method to determine individual CRP reliability levels that allows discarding unstable CRPs. Rather than measure ΔVout(CRP), which is a direct measure of reliability, we measure response’s sensitivity to comparator offset (by changing the calibration voltage with ΔVcal and scanning it). This method can greatly reduce the average BER to 2.6% with a 10% loss of CRPs or < 0.1% (0.4% in the worst case) with a 42% loss of CRPs (Fig. 9).

To prove resilience to ML attacks, we apply a suit of standard ML techniques, including support vector machines (SVM), logistic regression (LR), and neural networks (NN), to predict PUF responses given a number of training set CRPs. The prediction error stays over 40% on 10^4 training points, which is 100X higher than a 65-bit arbiter PUF, and the error decreases slowly compared with the arbiter PUF (Fig. 10a).

The secrecy capacity [7] increases linearly with CRPs while it saturates for the arbiter PUF (Fig. 10b).

Table 1 summarizes measurement results and compares with prior work. The SCA-PUF generates response bits at 6Kb/s while consuming 68nW and 11pJ/bit. The design occupies 44700μm^2 (die photo in Fig. 11).

### References


Type SCA PUF Delay PUF RO PUF
Post-attack security (prediction error on 10^4 CRPs) 40% - 1% 1% [6]
Secrecy capacity (bits) (on 2^16 CRPs) -2^4 - -523776
Number of CRPs 3.7 \times 10^{11} -5.5 \times 10^{11} -1.8 \times 10^{11} 1.4 \times 10^{10}
BER in worst case 4.5% 4.8% 0.48%
Power (µW) / Energy (pJ/bit) 0.068 / 11.0 28.4 / 17.75 - -
V_{DD} range 1.08-1.32V 0.7-1.2V 1.2-1.44V ±2% 1.08-1.2V
Temperature range -20-80°C -25-125°C -40-85°C 27-67°C 20-120°C