The Bidirectional Polyomino Partitioned PPUF as a Hardware Security Primitive

James B. Wendt and Miodrag Potkonjak
Computer Science Department
University of California, Los Angeles
{jwendt, miodrag}@cs.ucla.edu

***Abstract*—Physical unclonable functions (PUFs) have demonstrated great potential as fast and robust hardware security primitives. Public physical unclonable functions (PPUFs) have removed the main conceptual limitations of PUFs by enabling the creation of public key protocols. Traditional methods of constructing PPUFs leverage intrinsic process variation in submicron integrated circuits. However, these implementations impose high power usage and require long simulation times, producing high latency protocols. We propose to use next generation CMOS-compatible technologies, such as memristors and nanowires, whose components exhibit non-linear circuit characteristics, for the creation of PPUFs. We utilize the bidirectional nature of these components and introduce a novel architecture of PPUF polyomino partitioning. Furthermore, we present new security protocols that authenticate orders of magnitude faster than their CMOS-based PPUF counterparts. We simulate the design and demonstrate its resilience to a host of attacks using SPICE circuit simulation.**

***Index Terms*—Hardware security, public physical unclonable function, nanotechnology.**

I. Introduction

Physical unclonable functions are multiple input physical systems that produce one or more outputs that are difficult to predict or calculate. In addition, physical characteristics inherent in the PUF make it exceedingly difficult to reverse engineer and reproduce. They have been used in numerous applications ranging from authentication to hardware metering and remote system control [1] [2] [3]. PPUFs are PUFs that are intentionally made easy to reverse engineer. The extracted physical parameters of the PPUF system serve as the public key and the input-output mappings serve as the secret key. The key observation is that only the owner of the PPUF can easily and rapidly compute the output signals (response) for a given input (challenge). PPUFs enable public key communication and authentication protocols that are resilient to physical and side channel attacks [4].

A major drawback in current CMOS-based PPUF designs is that in any two-party communication protocol, the party without the PPUF has a much higher message passing latency than the owner of the PPUF [5]. This is because message passing requires authentication of challenge-response pairs computed by the PPUF, which in turn requires simulation of the device on the receiving side. While it is often the case that the size and security of the PPUF scale in tandem, the larger the PPUF is, the harder it becomes to simulate and thus, the longer the authentication time grows. Attempts to attenuate this disparity in communication latencies include matched quantized PPUF solutions and digital bimodal functions [6] [7].

Until now, all PUFs and PPUFs are realized using application specific integrated circuits or FPGAs [8]. The high simulation cost of these devices can be attributed, in part, to the unidirectionality of gates in the integrated circuit. In all CMOS-based ICs, the input to output flow is unidirectional and cannot be reversed. This restriction dictates that—unless some set of inputs and outputs can be partitioned from the rest of the design—full simulation of the entire device is required in order to authenticate an input-output pair. So far, no such partitioning mechanisms have been proposed for IC PUF designs.

We use beyond-CMOS technology, modeled after nanotechnologies such as memristors and III-V nanowires, to realize a faster, lower energy, and more capable PPUF as a new hardware security primitive. Our design enables a participating party to authenticate a challenge-response pair returned by simulating only a small randomly chosen partition of the design space, a feature that is not possible in current PPUF designs. This is feasible primarily due to the bidirectional nature of these nanotechnologies. Coupled with our architecture, this enables the design of very large PPUFs, rendering full simulation of the device impossible, while maintaining that authentication time remains constant.

By introducing this new PPUF we resolve the architectural issues of current designs that limit authentication to full simulation. We contribute a novel design that allows partitioning of the PPUF for fast authentication through rapid simulation. Furthermore, we provide new protocols—either conceptually new or faster than previous CMOS-based PPUFs—and show their resilience to a host of attacks.

II. Preliminaries

A. (Public) Physical Unclonable Functions

A PUF is a physical system that has a complex but definite and stable mapping of inputs to outputs. A PUF can be generalized as a very complex mathematical function derived from the intrinsic physical behavior of its components and design. The mapping of challenges to responses must remain easy to evaluate but impossible to predict [9]. Additionally, as the name suggests, PUFs are unique and unclonable, difficult to physically copy and difficult to simulate.

The scaling down of CMOS feature sizes has paved the way for leveraging intrinsic process variation in submicron integrated circuits in order to both accomplish non-linearity in challenge-response pairs as well as ensure unclonability. As the name suggests, PUFs are systems that are impossible to physically clone, thus a PUF is considered a unique key in itself.

The unpredictable challenge-response mappings coupled with the unclonability of the PUF make it a perfect candidate for secure cryptographic protocols. Low power and high speed PUFs are even more desirable.

B. Nanotechnology

New nanotechnologies have the potential to provide smaller form, lower power, and faster computation speeds than current technologies. We employ nanotechnologies which possess the additional following properties: (i) their material has a highly non-linear input-output response, and (ii) randomness is inherent in their synthesis.

The synthesis of self-assembled three-dimensional networks of III-V nanowires results in chaotic networks [10] [11]. Memristors exhibit inherent process variations similar to the process variations observed in transistors [12]. Memristors are of special interest because of their low leakage levels (10^{-7} less than current transistors), incredibly

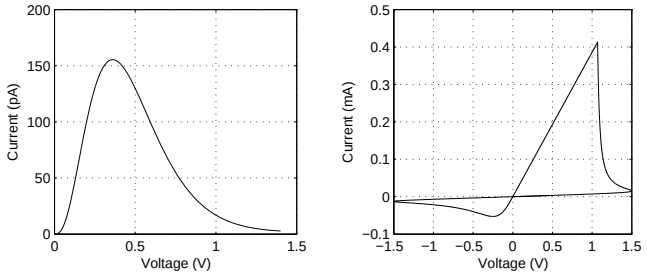

Fig. 1: I-V characteristics of III-V nanowires (left) and memristors (right).

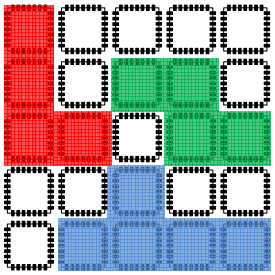

Fig. 2: Example of a 5×5 NanoPPUF grid. Adjacent NanoPPUF-cells are connected via adjacent pins. Example polyomino partitions of size 4, 4, and 5 are shaded.

fast switching speeds (latencies of $5ns$) and non-linear input-output response characteristics [13] [14].

Furthermore, these nanotechnologies are bidirectional in nature; their input and output pins do not need to be statically assigned. This introduces a new level of complexity and freedom in design. While ICs are inherently unidirectional in nature, nanotechnologies have the freedom to choose input vectors along with on which nodes to assign inputs. We demonstrate that this additional degree of freedom has a wildly non-linear affect on the output response of our nanotechnology-based PPUF.

Utilizing nanotechnologies for PUFs has been proposed recently [15] [16]. We progress this research by simulating our design and demonstrating its resilience to security attacks.

III. Design of the NanoPPUF

In its simplicity, the NanoPPUF is a random network of non-linear nanoelectronic components that, when applied some set of inputs at a set of boundary nodes, distributes those inputs non-linearly throughout the network and produces a random and unpredictable output at the remaining, unassigned boundary nodes. We formalize this notion by describing the NanoPPUF-cell and the larger NanoPPUF.

A. Microstructure: Non-linear Components

The geometry of the NanoPPUF-cell is modeled after the geometric random networks produced by the synthesis processes of III-V nanowires. We construct our NanoPPUF-cell network by uniformly randomly distributing nodes and connecting them within a threshold distance of one another corresponding to experimental measurements. While the random network synthesis benefits non-linearity, methodical placement of nodes still generates non-linear circuits due to non-linearity of the components. After the network is constructed, it is pruned of any disconnected subgraphs to ensure node voltage analysis convergence in the SPICE circuit simulator. The discrete components that make up the edges of the NanoPPUF-cell network are modeled after the current-voltage (I-V) characteristics depicted in Figure 1.

B. Macrostructure: NanoPPUF Design

The NanoPPUF consists of a grid of NanoPPUF-cells in which adjacent cells are connected via their matching pins. These connections are multiplexed for reasons we describe in the following section.

The analog nature of the device requires minimal power and energy consumption since the application of input voltages and measurement of output voltages occurs almost instantaneously. The total time taken to apply a challenge to the system consists of the delay in input vector application, network convergence (on the order of nanoseconds), and output readings.

When applying inputs to the NanoPPUF, both an input value vector (set of input voltages) and an input set vector (set of pins to apply those input voltages to) constitute an input or challenge. This ability to choose input sets increases the input and output space exponentially as compared to traditional CMOS which is limited to static input sets which must always be set high or low.

C. Polyomino Partitioning

We define a valid partition of the NanoPPUF device as a set of adjacent and connected NanoPPUF-cells. This entails that a path can be drawn from one NanoPPUF-cell to all others in the partition.

Pins connecting neighboring NanoPPUF-cells are multiplexed to facilitate application of input voltages as well as to facilitate voltage measurements. This enables enumeration of the full set of boundary conditions that surrounds a partition. Due to the bidirectional property of the nano-components, this partition is an effective sub-NanoPPUF that can be simulated and verified independently of the rest of the device provided that all boundary conditions are known.

As we demonstrate in the following section, simulation of the NanoPPUF grows exponentially with its size, ultimately rendering it too computationally expensive to complete in a reasonable amount of time. By partitioning the design space, we enable simulation of a much smaller portion of the NanoPPUF in order to validate that the boundary inputs and outputs of that partition are consistent with its internal circuitry. In Section V we use this property in our public key cryptographic protocols.

The gridded structure of the NanoPPUF allows us to partition the space using polyomino shapes. A polyomino is defined as a two-dimensional geometric figure formed by joining one or more equally sized squares edge to edge. Possibly one the more popular polyomino shapes are tetrominoes (polyominoes of size four) which are found in the game of Tetris. The number of fixed polyominoes, A_n, that can be constructed from n cells is estimated using the following formula:

$$A_n \sim \frac{c\lambda^n}{n} \tag{1}$$

where $\lambda = 4.0626$ and $c = 0.3169$ are estimates [17].

IV. Security of the NanoPPUF

In this section we discuss and analyze the security properties of our NanoPPUF design. Herein, we define an input set as the list of pins used as inputs for the NanoPPUF. We define an input vector as the vector of voltages assigned to those input pins. The mapping of an input vector to an input set constitutes a challenge. Similarly, we define the remaining pins as a part of the output set, and the output vector as the values at those pins. The mapping of an output vector to an output set constitutes a response.

A. Simulation Effort

The unique non-linearity expressed by the components in the NanoPPUF network is crucial to the application of this device as a PUF. The unpredictable challenge-response mappings that result from this non-linearity ensure that full simulation and prediction of the NanoPPUF remains impractical and infeasible.

We measure the simulation time across different NanoPPUF-cell sizes, comparing between using non-linear and linear components. Results of the experiments are depicted in Figure 3. While simulation

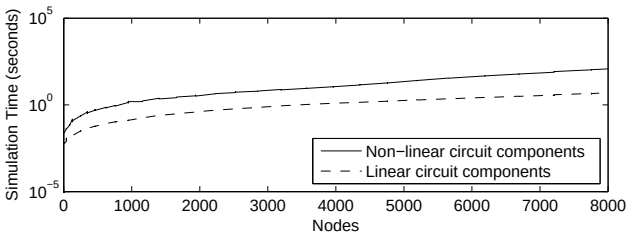

Fig. 3: Simulation time using the SPICE circuit simulator.

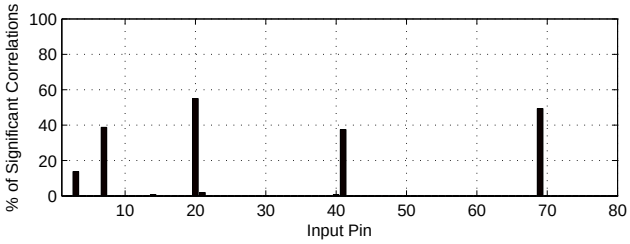

Fig. 4: Percentage of significant input-output correlations on a randomly generated NanoPPUF for an input set size of 2.

time increases exponentially with size in both the linear and non-linear cases, simulation time grows an order of magnitude faster for the non-linear circuit than for the linear circuit. Execution time of the physical NanoPPUF grows linearly with its size.

B. Input to Output Correlation

Figure 4 depicts the percentage of significant correlations between inputs and outputs for a randomly generated NanoPPUF. The application of input voltages is applied to two input pins. We choose two input pins because it is the smallest input set size that can be chosen without rendering output prediction trivial. One pin is kept static while the other varies; the varying pin is labeled on the x-axis. We iterate over a range of inputs on both pins and measure the outputs. We then calculate the input to output correlations for that given input set and range of inputs. The results highlight the low percentage of correlations that are statistically significant per input set.

C. Input Set to Output Correlation

Input to output prediction becomes even more difficult with the introduction of the additional degree of freedom that input sets are variable. A NanoPPUF-cell alone may have n pins. This constitutes an input set space of $O(2^n)$ possible sets coupled with $O(2^n)$ possible input values. Even if input-output statistical mappings are discovered by an adversary, they are limited to only exist per input set. Appending more NanoPPUF-cells to the design exponentially increases this space and ultimately refutes the possibility of enumerating all combinations.

D. Output to Output Correlation

Figure 5 depicts a subset of output correlations from an 80 pin NanoPPUF. Figures 5a and 5b are computed from the same NanoPPUF, but with different input sets. While some correlations exist in each experiment, a simple change in input sets from Figure 5a to Figure 5b shows a dramatic change in output to output correlation. The argument holds that as we increase the size of our NanoPPUF the number of mappings of input set to output-output correlation matrices grows exponentially large, thus eliminating the possibility of enumerating all combinations.

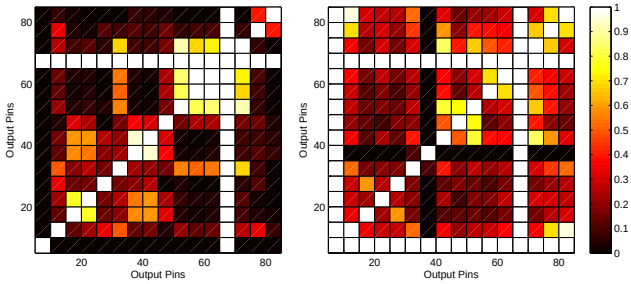

Fig. 5: R^2 correlations between a subset of NanoPPUF output pins. Both figures depict the same NanoPPUF with different input sets. Pin 65 and has perfect correlation because it returns a constant value due to process variation at network synthesis.

V. Protocols

We present two security protocols that leverage the design of the NanoPPUF. Recall that the design and characteristics of a PPUF are publically known such that anyone can simulate any polyomino partition of the design, however the device is large enough that full simulation is impossible.

We define the challenge, C, as the vector of inputs and the response, R, as the vector of outputs such that $C \rightarrow R$ when C is applied to the NanoPPUF. Recall that R consists of all output values in the NanoPPUF for the purposes of partitioning. The challenge set X represents the list of pins on which the challenge vector C is applied; the length of C and X are equal. The communicating parties, Alice and Bob, are represented by A and B, and their respective NanoPPUFs are denoted by PPUF_A and PPUF_B.

A. Authentication

We prevent adversarial spoofing attacks and identity theft through a simple and elegant NanoPPUF-based authentication protocol. In this scenario, Alice wishes to authenticate that she is communicating with Bob and not a malicious adversary pretending to be Bob. Alice begins by issuing a challenge C to Bob. Bob applies that challenge to his physical NanoPPUF, PPUF_B, and returns the computed result, $C \rightarrow R$, to Alice.

Given the challenge and response pair, Alice is able to validate the authenticity of Bob's response. Due to the partitionable design of the NanoPPUF and the bidirectionality of the NanoPPUF-cell, Alice can simulate a partition of the NanoPPUF and validate that the inputs and outputs along the boundaries of the partition converge to Bob's response, R.

An adversary pretending to be Bob would have to simulate the entire NanoPPUF design since he must respond to Alice with a full output response, R, and he cannot guess which polyomino partition Alice will choose to authenticate.

For additional security, Alice could validate two separate partitions in parallel or even request two challenge response pairs from Bob. The computation costs are negligible for Bob but prohibitively expensive for an adversary who would have to fully simulate the two challenges.

B. Remote Secret Key Exchange

The remote secret key exchange protocol allows Alice and Bob to securely communicate with one another by encrypting their messages with a secret key exchanged securely over an unsecured channel. In the case of the NanoPPUF, the secret key is a polyomino partition chosen at runtime.

Alice decides to send a message to Bob. Alice simulates a secret key, C_B, on a partition of PPUF_B and calculates R_B. Bob receives a copy of R_B, the encrypted message, $M = C_B \oplus m$, and the input set X, on which the challenge was applied. Bob is able to discover the

Protocol 1 Authentication

1: A sends challenge C to B
2: B applies challenge C to PPUF$_B$ and records response R
3: B sends R to A
4: A picks a random polyomino partition P of PPUF$_B$
5: A simulates the boundary conditions (C and R) on P to validate that node voltage analysis converges correctly

Protocol 2 Remote Secret Key Exchange

1: A simulates a secret challenge $C_B \rightarrow R_B$ on a partition P of PPUF$_B$
2: A computes $M = C_B \oplus m$, where m is the message
3: A sends M, R_B, and the challenge set X' to B
4: B iterates through challenges on X' until C_B is found
5: B computes $m = M \oplus C_B$

secret key to decrypt this message since he owns the NanoPPUF that originally constructed R_B. He accomplishes this by iterating through all possible combinations of inputs on the given input set until the challenge, C, is found such that $C \rightarrow R_B$.

An attacker snooping M, R_B, and X is unable to find the secret key since C_B can only be calculated from M, R_B, and X by iterating over an exponential number of possible input vectors on X.

We can render guessing C_B even more difficult by sending a set of input sets, X', which contains X and an excess of random input sets to increase the search space for the attacker. This has a minimal affect on the NanoPPUF owner, since the physical response of the device is ultra fast.

VI. Resiliency to Attacks

Correlations between input and output vectors in the NanoPPUF are dwarfed by the NanoPPUF's large input set space. Furthermore, the effect of swapping a single input pin in an input set drastically alters the output response. Even if a prediction mapping can be found between an input set and output vector, an attacker would require a unique prediction model for every possible input set. This task grows exponentially with the size of the NanoPPUF as the number of pins grows with each additional NanoPPUF-cell.

Correlations between outputs in the NanoPPUF are unpredictable and non-trivially dependent on the input values. Additionally, we demonstrated that output-output correlations are also unpredictable and non-trivially dependent on the input set.

While it seems unlikely, it is not yet provable that a statistical mapping from input set to output-output correlations does not exist. However, even if such a model is discovered, an attack would require enumerating all possible input set to output-output correlation mappings. Since the number of input sets is exponential with respect to the number of input pins, this attack is rendered infeasible. Furthermore, as previously discussed, the adversary has no knowledge of an input to output prediction model and thus has little to gain from an output-output correlation model since the input vector space grows exponentially with the inputs.

An attack via simulation is infeasible since simulation time increases exponentially with the size of the NanoPPUF. Constructing a larger device will render simulation attacks impossible while retaining the efficiency of the security protocols which rely on polyomino partitioning.

However, an attacker might attempt to leverage the speed and performance gained from special purpose hardware, such as FPGAs or ASICs, in order to simulate a NanoPPUF faster than in software. However, this effort is easily thwarted by increasing the size of the NanoPPUF, thus increasing the simulation time exponentially while even special purpose hardware cannot scale at that rate.

The side channel attack is simply not a threat since the complete characterization of the NanoPPUF is made public and any power or electromagnetic profile of execution would reveal no new information about the device or the mapping of inputs to outputs.

A look-up table attack would require enumeration of all possible inputs for all possible input sets along with the corresponding responses. While it would be near impossible to store this many entries, it would still require that the adversary have access to the NanoPPUF for a long enough period of time to fill even a fraction of such a database.

VII. Conclusion

In this paper we proposed the use of the non-linear and bidirectional characteristics in emerging nanotechnologies as the security mechanism behind the NanoPPUF. These nanotechnologies have demonstrated the potential to improve power, energy, and performance over existing CMOS-based PPUF designs. We have provided an answer to the existing issues in PPUF architecture that make rapid authentication difficult by exploiting the bidirectional nature of these nanotechnologies and building an architecture that compliments this characteristic. We designed an intelligent and elegant polyomino partitioning scheme that allows the NanoPPUF to grow to an arbitrarily large size while maintaining that authentication remains ultra fast and secure. We have analyzed the input and output relationships of the device through SPICE circuit simulation and have demonstrated its resilience to a host of attacks.

References

[1] S. Trimberger, "Trusted design in FPGAs," *DAC*, pp. 5-8, 2007.
[2] F. Koushanfar and M. Potkonjak, "CAD-based security, cryptography, and digital rights management," *DAC*, pp. 268-269, 2007.
[3] Y. M. Alkabani and F. Koushanfar, "Active hardware metering for intellectual property protection and security," *Proceedings of the USENIX Security Symposium*, pp. 1-16, 2007.
[4] N. Beckmann and M. Potkonjak, "Hardware-based public-key cryptography with public physically unclonable functions," *Information Hiding*, pp. 206-220, 2009.
[5] M. Potkonjak, S. Meguerdichian, A. Nahpetian, and S. Wei, "Differential public physically unclonable functions: architecture and applications," *DAC*, pp. 242-247, 2011.
[6] S. Meguerdichian and M. Potkonjak, "Using standardized quantization for multi-party PPUF matching: foundations and applications," *ICCAD*, pp. 577-584, 2012.
[7] T. Xu, J. B. Wendt, and M. Potkonjak, "Digital bimodal function: an ultra-low energy security primitive," *ISLPED*, pp. 292-297, 2013.
[8] S. Katzenbeisser et al., "PUFs: Myth, fact or busted? A security evaluation of physically unclonable functions (PUFs) cast in silicon," *Cryptographic Hardware and Embedded Systems*, pp. 283-301, 2012.
[9] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld, "Physical one-way functions," *Science*, vol. 297, no. 5589, pp. 170-178, 2002.
[10] K. A. Dick et al., "Directed Growth of Branched Nanowire Structures," *MRS Bulletin*, vol. 32, no. 2, pp. 127-133, 2007.
[11] L.-E. Wernersson, C. Thelander, E. Lind, and L. Samuel-son, "III-V nanowires-extending a narrowing road," *Proceedings of the IEEE*, vol. 98, no. 12, pp. 2047-2060, 2010.
[12] D. Niu, Y. Chen, C. Xu, and Y. Xie, "Impact of process variation on emerging memristor," *DAC*, pp. 877-882, 2010.
[13] D. B. Strukov et al., "The missing memristor found," *Nature*, vol. 453, pp. 80-83, 2008.
[14] R. Williams, "How we found the missing Memristor," *IEEE Spectrum*, vol. 45, no. 12, pp. 28-35, 2008.
[15] J. B. Wendt and M. Potkonjak, "Nanotechnology-based trusted remote sensing," *IEEE Sensors*, pp. 1213-1216, 2011.
[16] J. Rajendran, G. S. Rose, R. Karri, and M. Potkonjak, "Nano-PPUF: a memristor-based security primitive," *VLSI*, pp. 84-87, 2012.
[17] I. Jensen and A. J. Guttmann, "Statistics of lattice animals (polyominoes) and polygons," *Journal of Physics*, vol. 33, pp. 257-263, 2000.