Exploiting Machine Learning Against On-Chip Power Analysis Attacks: Tradeoffs and Design Considerations

Farid Kenarangi, *Student Member, IEEE* and Inna Partin-Vaisband, *Member, IEEE*

***Abstract*— Modern power analysis attacks (PAAs) and existing countermeasures pose unique challenges on the design of simultaneously secure, power efficient, and high-performance ICs. In a typical PAA, power information is collected with a monitoring circuit connected to the compromised device. The non-typical voltage variations induced on a power distribution network (PDN) by such a malicious probing are sensed with on-chip sensors and exploited in this paper for detecting PAAs in real-time using statistical analysis. A closed-form expression for the voltage variations caused by malicious probing is provided. Guidelines with respect to the PDN characteristics and number of sensors are proposed for securing power delivery. The PAA detection system is designed in a 45-nm standard CMOS process. Based on the simulation results, a PAA on an IBM benchmarked microprocessor is detected with the accuracy of 88% with 30 on-chip sensors. Power overhead of 0.34% and 14.3% is demonstrated in, respectively, the IBM microprocessor and a typical advanced encryption standard system. In a practical cryptographic device, security sensitive PDN regions can be identified, significantly reducing the number of the on-chip sensors.**

***Index Terms*— Side-channel attack, power analysis attack, hardware security, cyber security, cryptographic devices, machine learning, logistic classifier, data analysis, on-chip power delivery.**

I. Introduction

HARDWARE security of integrated circuits (ICs) is a significant concern in many emerging market segments, such as intelligent transportation, innovative health care, sophisticated security systems, and smart energy applications. In a typical hardware attack, physical side-channel information, such as IC power and timing traces [1], [2], memory cache hits and misses [3], and electromagnetic (EM) characteristics [4] is exposed in a running device. The exposed side-channel information is a strong function of the operations executed within the device, and can be related to sensitive on-chip data. To extract valuable data from a device, the dependence of the data on physical IC characteristics is exploited by attackers as part of the statistical analysis performed on collected side-channel information. One of the commonly used side-channel attacks is a power analysis attack (PAA) which aims revealing sensitive data based on IC power consumption [5].

Power distribution network (PDN) is a primary IC component, comprising up to several millions of nodes in modern VLSI systems. Voltage variations at the individual on-chip power grid nodes depend upon the effective impedance among the nodes, non-linear current loads, and distributed on-chip voltage sources [6], [7]. Note, that only on-chip PAAs are considered and all the referenced PDNs and power grids should be interpreted as on-chip PDNs and on-chip power grids to avoid confusion with high-voltage power grids. In a typical power attack, these voltage variations are captured in real-time with a small resistor (of up to 50 Ohms [8]) externally connected between a power or ground (P/G) pin of the device and off-chip power supply (*e.g.,* battery). Power dissipated within the resistor provides valuable information on the on-chip switching activity of the device and, ultimately, the executed vulnerable data.

Since the introduction of power analysis in 1998 [1], different PAA countermeasures have been proposed. Hiding and masking techniques are common preventive measures for enhancing resilience of modern integrated systems to PAAs [9]–[30]. With existing hiding techniques, the IC power consumption is adjusted for each operation, forcing identical power to be dissipated for different tasks. Dual-rail pre-charge (DRP) techniques, such as sense amplifier-based logic (SABL), dual-spacer dual-rail (DSDR) logic, and three-phase dual-rail pre-charge logic (TDPL) are known for their symmetric differential nature, maintaining constant power over time [9]–[14]. Limited supply current fluctuations with current-mode logic (CML), such as MOS current mode logic (MCML) and dynamic current-mode logic (DyCML) have also been explored for hiding sensitive information [15]–[17]. Alternatively, a complementary data-independent switching scheme has been proposed in [10] and demonstrated in [18] and [19] for flattening power profiles with wave dynamic and differential (WDDL) logic. Another approach has been presented in [20] for eliminating power dependence on data by equalizing circuit current with switching capacitors. Intuitively, power profile flattening methods exhibit higher power consumption, trading off power for

Manuscript received March 9, 2018; revised July 4, 2018 and September 4, 2018; accepted September 16, 2018. Date of publication October 11, 2018; date of current version January 18, 2019. This paper was recommended by Associate Editor P. K. Meher. *(Corresponding author: Inna Partin-Vaisband.)*

The authors are with the Department of Electrical and Computer Engineering, The University of Illinois at Chicago, Chicago, IL 60607 USA (e-mail: fkenar2@uic.edu; vaisband@uic.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSI.2018.2872567

security. Each of the aforementioned techniques has been reported to exhibit individual limitations, such as low scalability, high power consumption, degradation of performance, and/or limited effectiveness against PAAs.

With masking approaches, the objective is to add random characteristics to power profiles, reducing the dependence of overall power consumption on processed data. For example, power profile have been randomized in [21]–[24] by scrambling valuable data with Boolean and/or arithmetic operations. Power consumption with this approach is a function of the masked data and not the original valuable data. A single mask is, however, often not sufficient for mitigating modern PAAs. Alternatively, managing multiple masks increases the computational load of cryptographic operation, decreasing the IC performance. Delay is one of the parameters that is widely used as a source of randomness in masking approaches. A flow for random insertion of delays within data paths has been presented in [25], randomizing power profiles at each cycle. Data dependent propagation delays have been introduced in [26] and exploited for randomizing arrival times of the individual output bits. In [27], power profiles have been randomized with ring oscillators, generating additional random power traces within AES core. Secure double rate register (SDRR) has been proposed in [28] for randomizing data within combinational and sequential logic at register transfer level (RTL). In [29], randomized multitopology logic (RMTL) has been described that exhibits a reconfigurable logic format and dynamic structure, further supporting the concept of randomized power profiles. Random pre-charge logic (RPL) has been presented in [30], inducing random noise to the actual power profiles.

The existing countermeasures have been demonstrated to reduce the vulnerability of modern systems to power attacks. These approaches do not, however, guarantee the ultimate security against advanced PAAs, yet exhibiting significant overhead in terms of power, performance, area, design complexity, and system scalability [9]–[30]. In addition, the existing countermeasures are designed as preventive measures, and are not capable of detecting security attacks at runtime. A set of power efficient IC design solutions for detecting and mitigating PAAs in real-time is required to effectively enhance resilience of integrated systems to security attacks.

Detecting PAAs at runtime is proposed in this paper. The proposed approach is based on the observation that an externally connected device (*e.g.,* resistor) affects the effective impedance of the integrated power grid, inducing non-typical voltage variations at the node of connection (*i.e.,* the compromised node), as shown for a typical PDN in Fig. 1. To detect power grid probing, we propose to reveal these fine, non-typical voltage variations with compact power efficient sensors integrated on-chip. PDN exploration at the circuit level is however highly complicated in VLSI systems and straightforward analysis of power profile is computationally infeasible in real-time [31]. To detect these fine voltage variations in modern power grids, advanced algorithms are required for classifying the captured power information into secure and compromised categories. Machine learning (ML) is a diverse set of statistical techniques for analysis and in particular for classification of complex patterns in large data

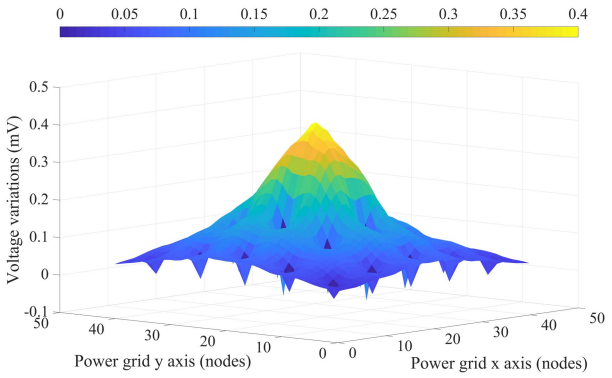

Fig. 1. Voltage variations across a typical 50×50 power grid caused by a probing circuit connected at the compromised node at $(25, 25)$. The highest variations are observed at the compromised node, as shown in light shade.

sets. Compact ML classifiers have recently been demonstrated for low power on-chip data classification [32] that can be exploited for providing an efficient alternative for the existing PAA countermeasures.

A theoretical framework that comprises circuits, models, and algorithms for on-chip ML-based prediction of the security level in a running device is proposed. With the proposed approach, PDN voltage variations are simulated during the system design stages in both the secure and compromised configurations and exploited for supervised training of a ML classifier. The trained ML classifier is designed on-chip for detecting non-secure probing of a power grid at runtime.

Given the distributed nature and complexity of modern power grids and process, voltage, and temperature (PVT) variations in advanced technology nodes, non-secure probing can be accurately detected with a single on-chip sensor only within a limited region surrounding the compromised node. To guarantee full security coverage in high-end VLSI systems, hundreds of compact, power efficient sensors should ultimately be integrated on-chip.

An analytic expression for voltage variations across power grid is derived that considers the effect of malicious probing. Based on the expression, the tradeoffs among the physical power grid characteristics, system security, and the power efficiency of the proposed ML approach are investigated and the related design considerations are described. By carefully choosing the power gird characteristics, a complex integrated system can be power-efficiently secured without compromising the performance and functionality of the system. The theoretical and simulation results provide an important, intuitive insight into the behavior of compromised power grids and the design of next generation secure and scalable power delivery systems.

The proposed ML-based approach is demonstrated based on IBM benchmark suite circuits [33], [34] that model typical microprocessor power grids. Based on SPICE circuit level simulation results, PAAs can be accurately detected with the proposed approach in modern microprocessors. To the best of authors knowledge, the set of the proposed circuits, models, and algorithms is the first IC design framework for accurately and *proactively detecting* side-channel attacks at runtime and

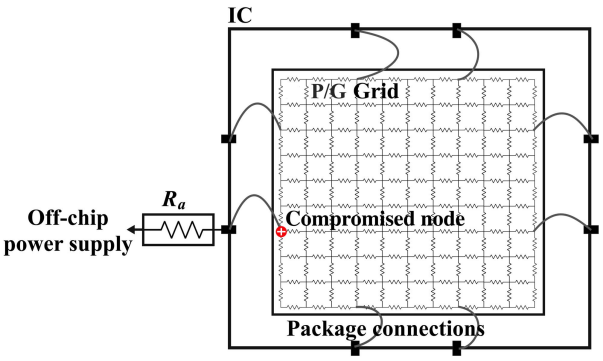

Fig. 2. Power analysis setup for capturing power profiles within a power grid.

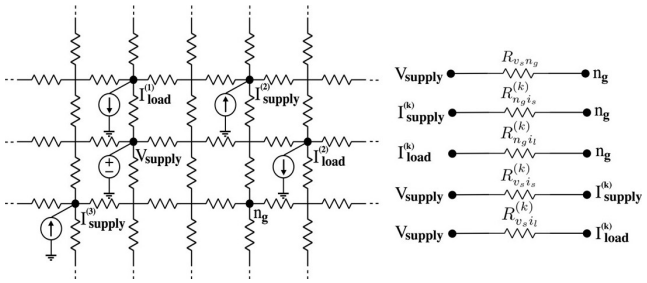

Fig. 3. A power grid comprising three voltage sources, and two current loads, $I^{(k)}_{load}$. All of the voltage sources except for one, V_{supply}, are converted to equivalent current sources, $I^{(k)}_{supply}$. The effective resistance between power grid nodes is also illustrated.

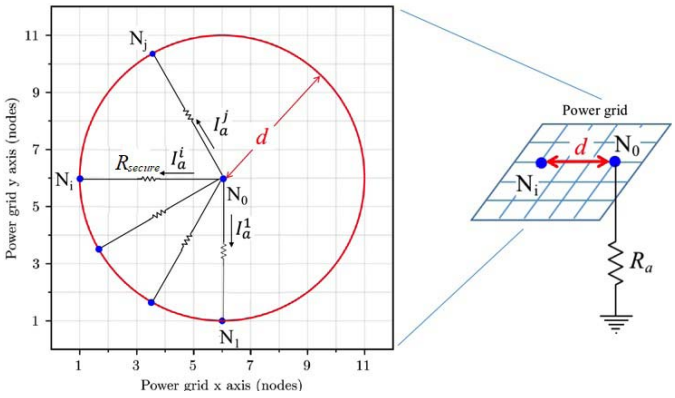

Fig. 4. Effective impedance and currents in a uniformly resistive compromised grid.

is expected to provide an efficient alternative for the existing preventive power attack countermeasures.

The rest of this paper is organized as follows. In Section II, a typical PAA is described and closed-form expressions for voltage variations caused by an off-chip malicious probing are derived. Based on the analytic expression, an ML framework is proposed for training and classifying power traces, as described in Section III. Power delivery design tradeoffs related to hardware security are discussed in Section IV. Simulation results are presented in Section V. The paper is summarized in Section VI.

II. PAA Analytic Description

Power attacks are known for their low cost, reliable, and non-invasive nature [5]. Several approaches exist for excerpting the power information from a running device [8], [35]–[37]. A typical PAA setup is illustrated in Fig. 2 with a small resistor (R_a) connected between the power line of the compromised device and the external power supply. In this configuration, the on-chip power signal at the compromised node is mirrored across the resistor R_a and exploited for revealing valuable physical information and consequently secret data. A power grid configuration with an access probe connected between an off-chip power supply and a PDN node is referred to as a compromised configuration throughout this paper. Alternatively, the configuration without an external resistor is referred to as secure configuration.

The effect of this resistive non-secure connection on the physical characteristics of a power grid is analyzed in this section. A PDN, designed as a grid of vertical and horizontal interconnections, is a primary IC component which physically distributes power supply voltages and currents in modern VLSI systems. Voltage at the individual power grid nodes is determined by the difference between the nominal on-chip voltage, $V_{nominal}$, and voltage drops across the power grid due to parasitic RLC impedances and switching of the distributed non-linear loads.

Voltage drop at the individual grid nodes has been formulated in [6] for an infinite uniform resistive grid with multiple voltage sources and current loads. With this approach, all of the voltage sources except for one are converted to equivalent current sources, $I^{(k)}_{supply}$, supplying current with opposite direction with respect to the current loads, $I^{(k)}_{load}$. A power grid with three voltage sources and two current loads (*i.e.*, $I^{(k)}_{load}$) is illustrated in Fig. 3. All of the voltage sources except for one (*i.e.*, V_{supply}) are converted to equivalent current sources (*i.e.*, $I^{(k)}_{supply}$). Principle of superposition is utilized to determine the voltage drop at any arbitrary node n_g in a power grid with N voltage sources and M current loads [6],

$$\begin{aligned} V_{n_g} = {} & V_{nominal} \\ & - \frac{1}{2}\sum_{k=1}^{M}[I^{(k)}_{load}(R_{v_s n_g} + R^{(k)}_{v_s i_l} - R^{(k)}_{n_g i_l})]) \\ & + \frac{1}{2}\sum_{k=2}^{N}[I^{(k)}_{supply}(R_{v_s n_g} + R^{(k)}_{v_s i_s} - R^{(k)}_{n_g i_s})], \end{aligned} \tag{1}$$

where the indices $n = \{n_g, v_s, i_s, i_l\}$ identify the power grid nodes associated with, respectively, the arbitrary power grid node, the voltage source, a current source ($k = 2, \ldots, N$), and a current load ($k = 1, \ldots, M$). $R_{n_1 n_2}$ is the effective resistance between the nodes n_1, $n_2 \in n$ (see Fig. 3). The negative and positive summation terms in (1) describe the contribution of, respectively, the original system current loads, $I^{(k)}_{load}$, and those current sources converted from voltage sources, $I^{(k)}_{supply}$. The voltage drop determined by these summation terms is a function of the effective resistances $R_{n_1 n_2}$. Hence, changes in

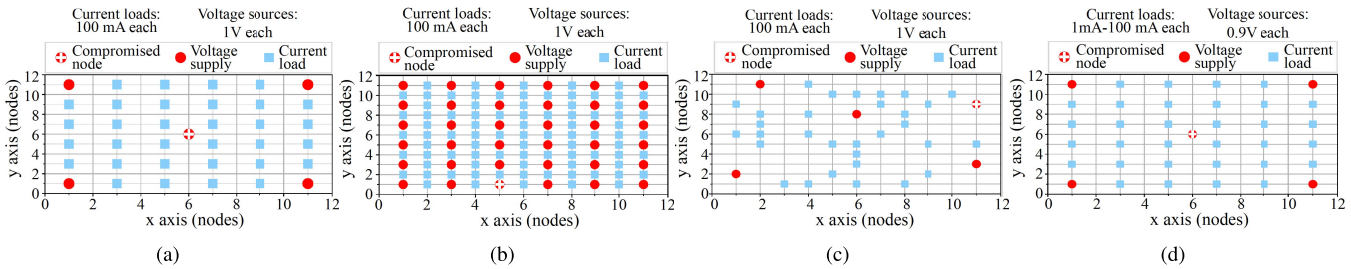

Fig. 5. Power grid configurations with a monitoring circuit and various number, magnitude, and placement of voltage sources (red circles) and current loads (blue squares). Adjacent resistance of $R_{adj} = 0.4\ \Omega$ and access resistance of $R_a = 1\ \Omega$ is considered in all these configurations. The voltage variations at each node within the individual compromised grids are simultaneously evaluated based on (4) and in SPICE. Maximum error of (a) 0.37%, (b) 0.59%, (c) 0.67%, and (d) 0.31% is reported with (4) as compared to SPICE results.

the effective power grid impedance due to malicious probing are expected to induce changes in power grid voltage levels.

A primary objective is to find the effective resistance between the compromised node N_0 and a power grid node N_i located at a distance d from the N_0. Here, the distance d is the number of nodes between the N_0 and N_i and R_{secure} and $R_{compromised}$ is the effective resistance between two nodes in, respectively, secure and compromised configuration. First, a continuous uniform resistive medium is assumed. Note that only portion of the total current, I_{tot}, flowing through R_a reaches the node N_i. Consider a circle centering at the node N_0 with the node N_i located on the circle perimeter, as shown in Fig. 4. Recalling the symmetric nature of the uniform power grid around the N_0, nodes N_i and N_j on the perimeter are provided with the same share ($I_a{}^i = I_a{}^j = \frac{I_{tot}}{2\pi d}, \forall i, j$) of the total current, I_{tot}, flowing from the N_0. On the other hand, R_a, can be modeled as $2\pi d$ parallel connected resistors, $R'_a = R_a \cdot 2\pi d$. Applying Kirchoff's law at N_0, the current at each power grid node on the perimeter of the circle shown in Fig. 4 is the current through a single resistor, R'_a. In a practical on-chip power grid, nodes are located in discrete locations. Thus, the node N_i is provided with the additional share of the total current, I_{tot}, from the adjacent virtual nodes (*i.e.*, those nodes existent in continuous, but not in discrete configuration), practically decreasing the effective resistance between N_0 and N_i by a factor of $4d$. The series effective resistance introduced by the probe is therefore,

$$R_{probe} = R_a \cdot 2\pi d \cdot \frac{1}{4d} = \frac{\pi}{2} R_a. \tag{2}$$

The effective resistance between two nodes in a uniform resistive secure grid can be analytically expressed as a function of the vertical (n) and horizontal (m) distance between the power grid nodes [38],

$$R_{secure} = \left[\frac{1}{2\pi} \ln(n^2 + m^2) + 0.51469\right] R_{adj}, \tag{3}$$

where R_{adj} is the resistance between two adjacent nodes. Thus, the effective resistance between two nodes in a compromised PDN is formulated based on (2) and (3) as,

$$\begin{aligned} R_{compromised} &= R_{secure} + R_{probe} \\ &= \left[\frac{1}{2\pi} \ln(n^2 + m^2) + 0.51469\right] R_{adj} + \frac{\pi}{2} R_a. \end{aligned} \tag{4}$$

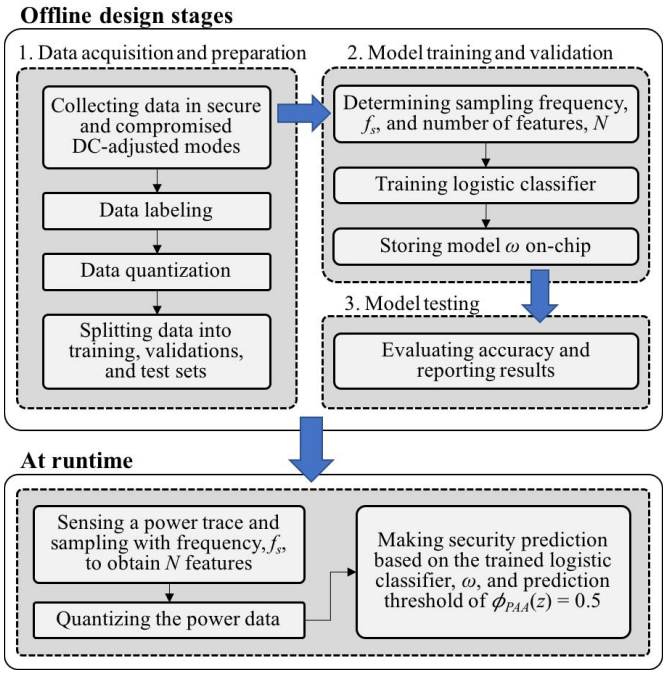

Fig. 6. Proposed ML-based training flow for detecting PAAs.

To evaluate the analytic expression in (4), four power grid configurations are considered. Each power grid comprises eleven vertical and eleven horizontal power lines, a monitoring circuit, and various number, magnitude, and configurations of power components (voltage sources and current loads), as shown in Fig. 5. The accuracy of the resulting on-chip voltage levels obtained based on (4) is verified with SPICE simulations for each of these power grid configurations, yielding a maximum voltage error of less than 7 mV (*i.e.*, 0.7% of the nominal supply voltage) as compared to SPICE. In these power grid configurations, $R_{adj} = 0.4\ \Omega$ and $R_a = 1\ \Omega$ is considered. The design flow of the proposed ML framework is derived based on the theoretical foundation presented in this section, as described in the following section.

III. ML Design Flow and Parameters for PAA Detection

In this section, a ML framework for detecting PAAs is described. The detection flow is illustrated in Fig. 6,

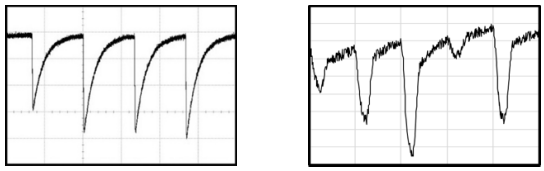

Fig. 7. Typical power waveforms extracted from (a) an experimental AES circuit on FPGA board SAKURA-G [39], and (b) a realistic industrial-size IBM benchmark.

comprising data acquisition and preparation, model training and validation, model testing, and at runtime PAA detection, as described in Sections III-A-C. ML parameters and Python simulation results are provided in Section III-D.

A. Data Acquisition and Preparation

Data collection and feature selection are primary factors in efficiently increasing the accuracy of the ML-based classification. With the proposed design flow, power traces (*i.e.*, continuous voltage levels across a PDN) are collected as part of the design process of a cryptographic system in secure and artificially designed compromised configurations, and labeled, respectively, '0' and '1'. To simplify the process of designing a realistic industrial-size system, power traces in this work are modeled based on IBM power grid benchmarks in Cadence. To illustrate the practicality of the IBM power grids [33], [34], an IBM power trace with 10% variations is shown in Fig. 7 along with an experimentally measured power trace of a common Advanced Encryption Standard (AES) system [39]. Note the similar periodic behavior of these power traces. Periodicity of a typical power trace is exploited to obtain numerous ML observations (*i.e.*, each power signal period is used as a single power observation). To account for system variations and generate a practical data set, a randomly distributed noise is added to the power traces, yielding a balanced data set of 2,000 secure and 2,000 compromised power observations. Out of the 4,000 unique observations 70%, 15%, and 15% are used for, respectively, training the classifier, validating ML parameters, and testing the proposed system. The individual power trace observations are sampled and used as ML features for PAA detection.

1) Sparse Sampling: Sampling frequency of the power traces and number of ML features are both important for determining the overall power consumption of the proposed PAA detection ICs. To extract a sensitive data in a typical PAA, thousands of power traces are commonly required [5]. The necessity for the numerous repetitions of the cryptographic operation is exploited for reducing the power consumption of the proposed PAA detection system with sparsely sampled power traces. With this approach, power traces are sampled at lower frequency over numerous cycles of the PAA operation, yielding longer detection time and lower average power consumption. To illustrate the proposed approach, consider a typical periodic voltage signal with a period of 6 ns and six features, as shown in Fig. 8. Note that feature (1) acquired at $t = t_1$ can also be obtained at time $t_1 + K \cdot 6$ ns ($K \in \mathbb{N}$). For example, features similar to those selected within the first period can be collected over ten periods, as shown in Fig. 8, significantly reducing the sampling frequency and thus, the power consumption of the proposed PAA detection system.

2) DC-Shifted PAA: An example of typical PDN voltage variations captured in secure and compromised configurations with nominal supply voltage of 1.8 V is illustrated in Fig. 9 in the form of transient signal (Fig. 9(a)) and histogram (Fig. 9(c)) of a single ML feature. In this example, the maximum voltage droop due to power grid parasitic impedance is 125 mV in secure configuration. Alternatively, in the compromised configuration a larger voltage droop is observed due to the increased grid impedance. Hence, there is an apparent difference of 34 mV in the on-chip voltage DC component in secure and compromised configurations. ICs operated in secure and compromised configurations can therefore be distinguished using simple averaging methods, not requiring utilization of more advanced ML approaches. This DC shift in a compromised PDN can be significantly reduced by an attacker through properly adjusting the DC voltage level of the off-chip power supply. Classification of PDN security level in these advanced optimally adjusted PAA setups is considered in this paper. Power traces captured in secure and compromised DC-adjusted systems are shown in Fig. 9(b). A histogram of a single ML feature is shown in Fig. 9(d), exhibiting significantly lower variance between the secure and compromised data as compared with the non-adjusted PAA configuration (see Fig. 9(a) and Fig. 9(c)). Classifying the PDN security level under a DC-adjusted power attack is impractical with the straightforward averaging methods. Alternatively, ML classifiers can be exploited to accurately distinguish a secure IC operation from a compromised operation under DC-adjusted PAAs. All the results are reported based on ideally DC-adjusted PAAs. Note, that in practical cryptographic devices, ideally compensating for the DC shift at a compromised node is a challenging and often infeasible task. Thus, the accuracy of PAA detection in those practical, partially DC-adjusted systems is expected to be higher than the worst-case accuracy reported in the following sections.

B. Model Training and Validation

To design an effective integrated solution for PAA mitigation, design complexity, power, area, and detection accuracy of the proposed ML circuits are simultaneously considered. First, the collected power profiles are projected onto a two-dimensional space using principle component analysis (PCA). The visualized projected data exhibits linearly separable input features. Thus, both linear and non-linear classifiers can be exploited for classifying secure and compromised configurations. Existing non-linear classifiers are typically computationally expensive and exhibit significant power and area overheads. Thus, linear classifiers should be preferred for on-chip integration in real-time applications, subject to performance constraints. The performance of the proposed solution for PAA detection is evaluated in Python with logistic regression, support vector machine (SVM), and

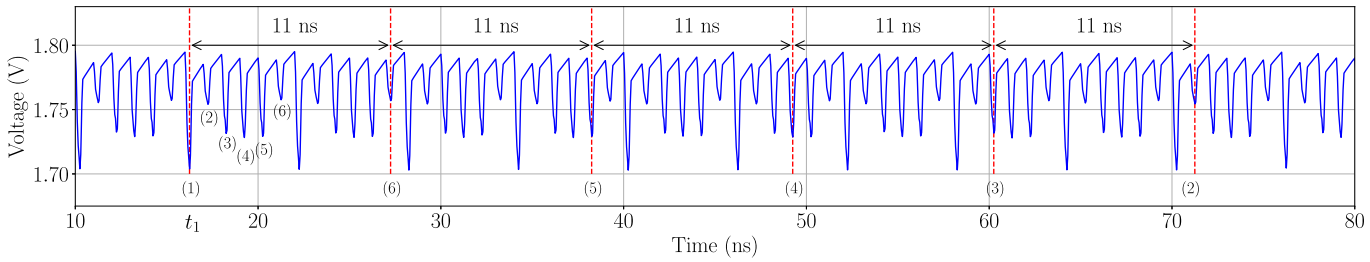

Fig. 8. Sparsely sampled ML features are collected over multiple periods (*i.e.*, observations) of a power trace, lowering the average power consumption of the proposed PAA detection approach. Collecting two sets of identical features is demonstrated within a single period (sampling frequency of 1 GHz) and over ten periods (sampling frequency of 90 MHz).

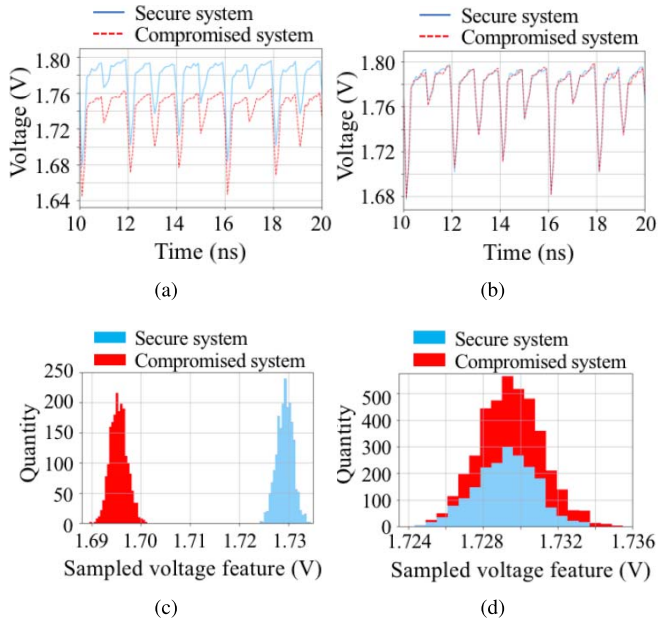

Fig. 9. Voltage signals at secure and compromised PDN nodes, (a) typical transient voltage profile, (b) DC-adjusted transient voltage profile, (c) histogram of a single ML feature, as sampled based on (a), and (d) histogram of a single ML feature, as sampled based on (b).

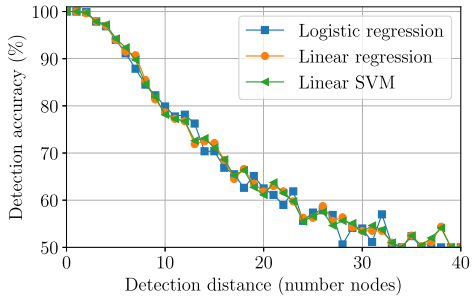

Fig. 10. Accuracy of PAA detection as function of the distance between the sensor and compromised node with logistic regression, linear regression, and SVM models.

linear regression. While all the three models demonstrate similar performance (see Fig. 10), logistic model exhibits lower design complexity in binary classification problems due to its simpler threshold function, as described in the next subsection (see (8)). Thus, the logistic regression is preferred, owing to its high accuracy, low design complexity, and superior power and area characteristics.

Logistic regression (LR) is a common supervised ML model known for its high performance and simple implementation. The conventional LR model outputs a continuous probability range of prediction variable. Alternatively, a probability threshold can be used to divide the continuous output range into two discrete classes. Such a model is designed based on a logistic regression and a probability boundary of 0.5 [40]. The proposed model is referred to as logistic classifier. The logistic classifier is trained based on data acquired from SPICE, x_i^k. In this notation, the indices i and k are, respectively, the power trace index (*i.e.*, observation number) and the sample index within a power trace (*i.e.*, feature number). For example, x_2^3 is the third feature of the second collected observation.

The response of the proposed logistic classifier is determined based on a linear combination of features $x_i = ({x_i}^1, {x_i}^2, \ldots, {x_i}^N)$ and model weights $w = (w^1, w^2, \ldots, w^N)$,

$$Z = w^T x_i. \tag{5}$$

The training objective is to find the weights w that minimize the cost function $J(y, z)$ across all of the observations [41],

$$J(y, z) = -y \log(\phi(z)) - (1 - y) \log(1 - \phi(z)), \tag{6}$$

where y is the truth label and $\phi(z)$ is the sigmoid function that outputs the probability of power data x_i to be classified as compromised with model weights w,

$$\phi(z) = \frac{1}{1 + \exp(-z)}. \tag{7}$$

Note that $\phi(z)$ is a number between 0 and 1 which in fact corresponds to the probability of x_i being compromised. For example, for $\phi(w^T x_i) = 0.8$, there is an 80% likelihood for a system to be compromised with a power trace which corresponds to x_i.

C. Model Testing and Operation at Runtime

To avoid the on-chip implementation of non-linear sigmoid function (see (7)), a probability boundary of 0.5 is used for testing the proposed classifier and in a running device. The prediction of the logistic classifier is based on,

$$\begin{aligned} \phi_{PAA}(z) &= \text{sign}(z) \\ &= \text{sign}\Big(\sum_{i=1}^{m} w_i x_i\Big) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0, \end{cases} \end{aligned} \tag{8}$$

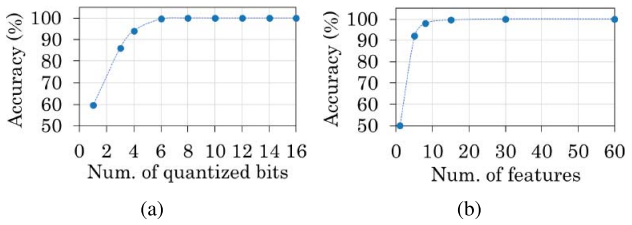

Fig. 11. Accuracy of the PAA detection with the proposed logistic classifier in a practical power delivery system, considering (a) various quantization levels, and (b) various number of features.

where the decision value of '0' or '1' is the indicator of the system operated in, respectively, secure or compromised configurations. The accuracy of the proposed logistic classifier is evaluated on a test set as a percentage of all the correct predictions out of the total number of test predictions,

$$\alpha = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%, \tag{9}$$

where, TP is the number of compromised (*i.e.*, attack-positive) configurations that are correctly classified as compromised, TN is the number of secure (*i.e.*, attack-negative) configurations that are correctly classified as secure, FP is the number of secure configurations incorrectly classified as compromised, and FN is the number of compromised configurations incorrectly classified as secure.

D. ML Parameters and Simulation Results

With the proposed sparsely sampled power traces, PAA can ultimately be detected with high accuracy at very low sampling frequencies. Alternatively, the number of features, and quantization level of power data are key system parameters that set the upper limit for PAA detection accuracy. Within the proposed framework, these parameters are determined iteratively based on training and validation sets. All the simulation results in this subsection are obtained based IBM benchmark suite circuits [33] and the proposed logistic classifier trained with $70\% \cdot 4,000 = 2,800$ observations in Python with scikit-learn ML library [42].

To determine the preferred A/D resolution, the proposed classifier is trained and validated with varying quantization levels. The accuracy of PAA detection in a noisy power delivery system is illustrated in Fig. 11(a) as a function of A/D resolution. Based on these results, the 60% detection accuracy with a single bit resolution approaches the performance of a random guess. Alternatively, the accuracy rapidly increases with the increasing number of quantization bits and saturates at the theoretical limit of 100% for a practical noisy power delivery system. To meet the accuracy requirement of PAA detection in the IBM power grid, eight bits are assigned for encoding the sampled power trace on-chip. The preferred number of features is determined as 30 with a similar iterative process of training and validation (see Fig. 11(b)). In a practical design, the accuracy of PAA detection is a strong function of physical system characteristics, affected by the PDN, on-chip sensors, power supplies, and current loads. Primary system parameters that affect the performance of the proposed ML framework are discussed in the following section.

IV. PAA Detection - Physical Characteristics and Design Tradeoffs

A typical IC has multiple external P/G supply pins that can ultimately be used for connecting PAA equipment. The location of the PAA compromised node is therefore unknown during the IC design stages. Yet, power grid should be secured at all potentially compromised P/G nodes. Compact integrated on-chip sensor is a primary component of the proposed approach exploited for sensing non-typical voltage variations within a PDN. The PDN voltage variations can, however, be accurately sensed only within a certain effective radius from the physical location of an on-chip sensor. Thus, multiple integrated sensors are considered for maintaining the system-wide security of cryptographic circuits. The effective radius of the individual sensors is a strong function of power grid impedance, physical location of the off-chip and distributed on-chip power supplies, and sensor electrical characteristics. The tradeoffs among physical and electrical system parameters are investigated in this section for effectively detecting PAA with minimum number of on-chip sensors.

A. The Effect of Number of Sensors on System Security

The effective radius of a sensor, δ_{eff}, is defined here as the maximum distance of the compromised node, n_c, located at (x_c, y_c) from the sensing node, n_s, located at (x_s, y_s) such that the sensor at n_s can detect a PAA at n_c with accuracy, $\alpha_{PAA}(n_s, n_c)$, that is equal or higher than a threshold accuracy, α_{th}. Considering the discrete nature of PDN, the effective radius is reported as the effective number of power grid nodes, $|n_s - n_c|$, between n_s and n_c,

$$\delta_{eff} = \max_{(x_s, y_s)} \{|n_s(x_s, y_s) - n_c(x_c, y_c)| : \alpha_{PAA}(n_s, n_c) \geq \alpha_{th}\}. \tag{10}$$

To accurately detect a PAA, multiple compact sensors should be integrated on-chip in ultra large scale integrated systems. Alternatively, those smaller systems such as smartcards, RFID tags, FPGAs, and microcontrollers can ultimately be secured with a single sensor. A system is secure at a certain point of time if and only if all its sub-regions (as determined by the individual on-chip sensors) are secure at this point of time. Based on findings collected from numerous power grid simulations, the number of required sensors increases linearly with the power grid size and limited by the number of the system power pins. Alternatively, the accuracy of PAA detection within a single sensing sub-region is completely independent of the on-chip power grid size, demonstrating high scalability of the proposed ML framework.

To illustrate the effect of threshold accuracy on the number of sensors, an IBM power grid comprising 10,000 nodes, 169 distributed power supplies, 5,387 non-linear current loads, and a resistive probe connected at the center of the power grid is considered. The sensitivity radius, δ_{eff}, is obtained in this configuration based on (10) for threshold accuracy, α_{th},

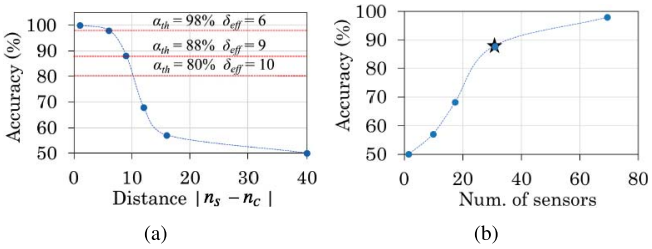

Fig. 12. Detection accuracy in an IBM power delivery system, (a) at different distances from the sensor, and (b) with multiple sensors.

of 80%, 88%, and 98%. The PDN is designed and simulated in Cadence. The simulation results of the PAA detection accuracy and number of sensors required for securing the system are shown in Fig. 12. The results exhibit lower accuracy with the increasing distance between the sensing and compromised nodes, $|n_s - n_c|$, and smaller number of sensors. Alternatively, power consumption, footprint, and design complexity also decrease with smaller number of sensors. For those applications with limited power and area resources, a knee point at lower accuracy threshold should be considered. For example, a PAA in the simulated IBM power grid can be detected with accuracy of 88% with 30 sensors.

B. The Effect of PDN Impedance on System Security

The effective resistance between PDN nodes, as defined by (4), comprises two terms. The first term in the equation describes the impedance between nodes in a typical secure power grid. This impedance increases with longer distance between the nodes and higher resistance between adjacent nodes, R_{adj}, decreasing quality of power (QoP) at the loads. Alternatively, the QoP increases with multiple, distributed on-chip power supplies. The effect of malicious probing on the effective PDN impedance is determined by the second term in (4), and increases with higher resistance of the access probe R_a, resulting in larger voltage variations across the power grid. The voltage variations due to malicious probing (*i.e.,* PAA signal) in those systems with higher resistance R_a and lower resistance R_{adj} are therefore more apparent to an on-chip sensor. Intuitively, the capacity of a sensor to detect malicious voltage variations is higher in those power grids with higher QoP. The optimum number of distributed sensors is determined based on the resistance ratio R_a/R_{adj} and QoP. Note that the amplitude of the PAA signal is determined by the impedance characteristics of the monitoring circuit and cannot be controlled during the design process. Alternatively, the secure power signal is determined by power supplies and physical power grid characteristics.

To illustrate the effect of power grid impedance on the sensitivity radius δ_{eff}, a typical 100×100 uniform PDN comprising 169 uniformly distributed power supplies, 5,387 non-linear current loads, and a resistive probe connected at $n_c = (50, 50)$ is simulated in Cadence, as demonstrated in Fig. 13. A three-dimensional map of absolute voltage variations in this system is shown in Fig. 14 for different values of the grid resistance ratio (R_a/R_{adj} = 2, 5, and 10). The PAA signal,

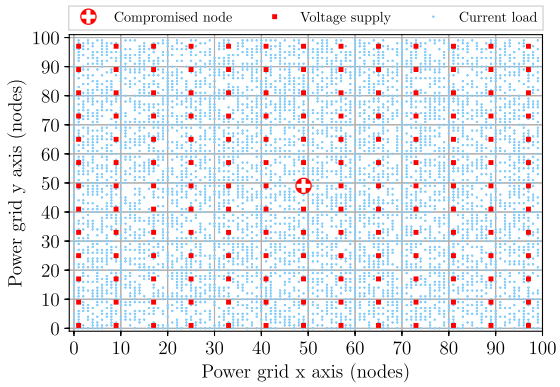

Fig. 13. Schematic of a PDN with 169 uniformly distributed voltage sources, 5,387 non-linear current loads, and a resistive probe connected at (50, 50).

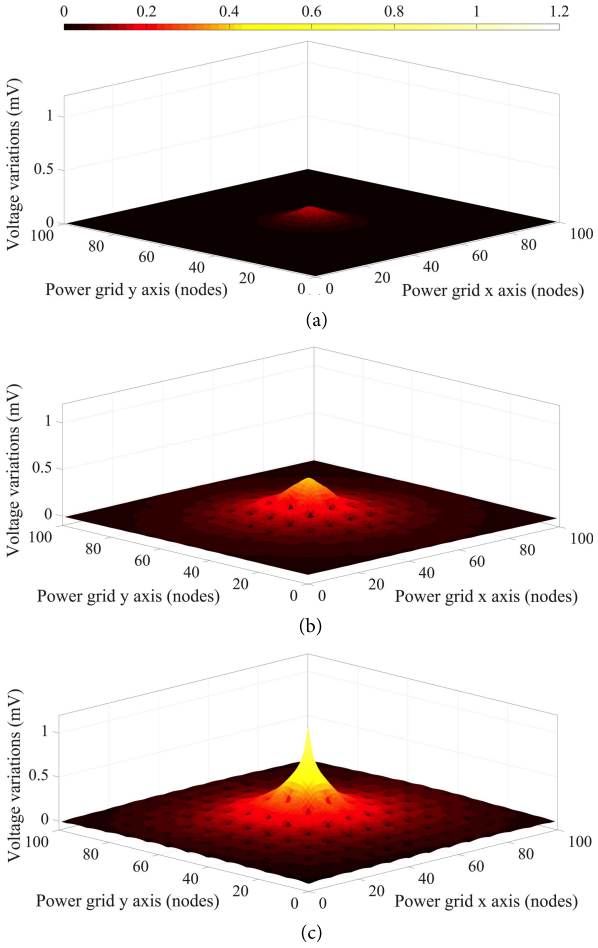

Fig. 14. Voltage variations in a compromised power grid with an access probe resistance of R_a= 1 Ω, and resistance between adjacent power grid nodes of (a) R_{adj} = 0.5 Ω (R_a/R_{adj} = 2), (b) R_{adj} = 0.2 Ω (R_a/R_{adj} = 5), and (c) R_{adj} = 0.1 Ω (R_a/R_{adj} = 10).

as shown in the figure, propagates to a greater distance in those systems with higher resistance ratio R_a/R_{adj}, increasing the effective radius of the individual integrated sensors, as shown in Fig. 15. Thus, to increase the resilience of ICs to PAAs with less sensors, low impedance PDN should be preferred for

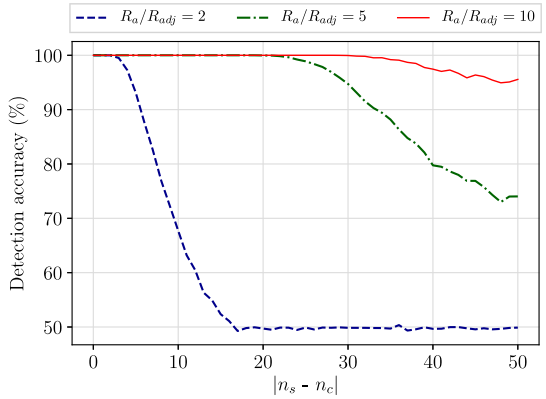

Fig. 15. PAA detection accuracy as function of PDN impedance.

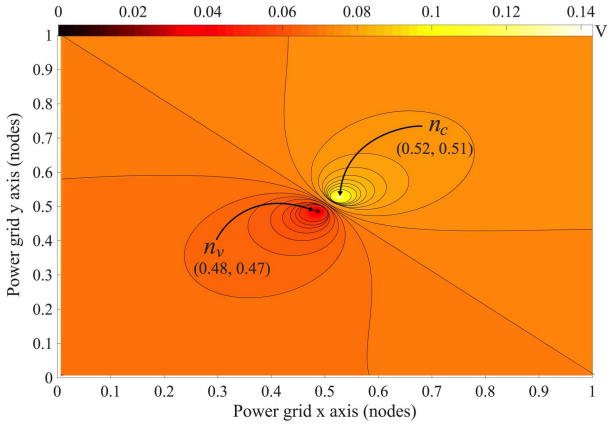

Fig. 16. Voltage map of a typical PDN. Shade intensity illustrates the absolute amplitude of voltage variations. Voltage variations are lowest (dark dot) in close proximity to voltage source and highest (light dot) at the PAA compromised node.

simultaneously enhancing and securing the power signal. The design of low-impedance PDN is an important cornerstone to the process of securing modern integrated systems. With the proposed approach, lower impedance power grids can ultimately be secured with a few integrated sensors.

C. The Effect of Physical Location of the Distributed Power Supplies on System Security

Based on the principle of spatial locality in power delivery systems, highest share of load current is supplied by the nearest power supplies due to the lower effective impedance [43]. A two-dimensional normalized map of absolute voltage variations in a typical PDN is shown in Fig. 16 with a power supply (located at $n_v(0.48, 0.47)$), an externally connected resistive probe (located at $n_c(0.52, 0.51)$), and distributed current loads. As a result of spatial locality of power distribution, the lowest voltage variations (dark dot) are exhibited in close proximity of the on-chip power supply, n_v, and the highest voltage variations (light dot) are observed at the compromised node, n_c. If the effective resistance between n_v and a sensor located at node n_s is lower than the effective resistance between n_c and n_s, a higher portion of the current is supplied to n_s by the voltage source located at n_v than by the malicious voltage source at n_c. As a result, the voltage variations at n_s are dominated by the power supply at node n_v and those voltage variations due to probing decrease at close proximity with n_v. The PAA detection sensors should therefore be integrated farther from the on-chip power supplies. The resilience of ICs to PAAs is expected to increase in those power delivery systems with sparsely distributed power supplies. To demonstrate the proposed method at the circuit level, this ML framework is designed in Cadence. The details of the circuit level implementation are described in the following section.

V. Design of the PAA Detection System

The proposed ML-based framework is designed at the 45 nm technology node. The area occupied by a single PAA detection circuit is 76 μm^2, as estimated based on transistor count in SPICE. A total of 30 PAA detection circuits should be integrated on-chip for detecting PAA across the IBM power grid with accuracy of above 88%. The ML classifier is designed with 30 features for PAA detection at 85 MHz sampling frequency, exhibiting an average power consumption of 34.71 μW. As compared with power consumption of hundreds of miliwatts (and up to several watts) reported for the existing state-of-the-art PAA countermeasures [9]–[30], the proposed system exhibits significantly lower power consumption. The security level of the system is determined with accuracy of 88% within 31 clock cycles of the system operation. The design of the proposed ML classifier is described in Section V-A. Simulation results for real-time PAA detection are presented in Section V-B.

A. Design of the On-Chip ML Classifier

The proposed logistic classifier with the probability threshold of 0.5 is designed for detecting PAAs at runtime. The classifier is trained using an error adaptive algorithm for enhancing classification performance in presence of noise [44]. A schematic representation of the integrated system is illustrated in Fig. 17, comprising summation, multiplication, and memory circuits. Mirror adder topology is chosen for ML summation as a power efficient alternative for the conventional adder. Transistor level schematic of the mirror adder is shown in Fig. 18(a). The mirror adder blocks are also utilized for designing the digital multiplication unit. To maintain high detection accuracy, 8 bits and 16 bits are assigned for encoding, respectively, the sampled P/G traces and individual feature weights. A 8×16 bit multiplication unit is designed, yielding the most power consuming component of the classifier. By reducing the dimensionality of the proposed logistic model, higher detection accuracy can be traded off for lower power consumption. A Master-Slave Flip-Flop (MS-FF) circuit is used for storing data in the system. The transistor level schematic of the MS-FF is shown in Fig. 18(b). Similar MS-FF components are used as digital buffers for synchronizing the classifier operations.

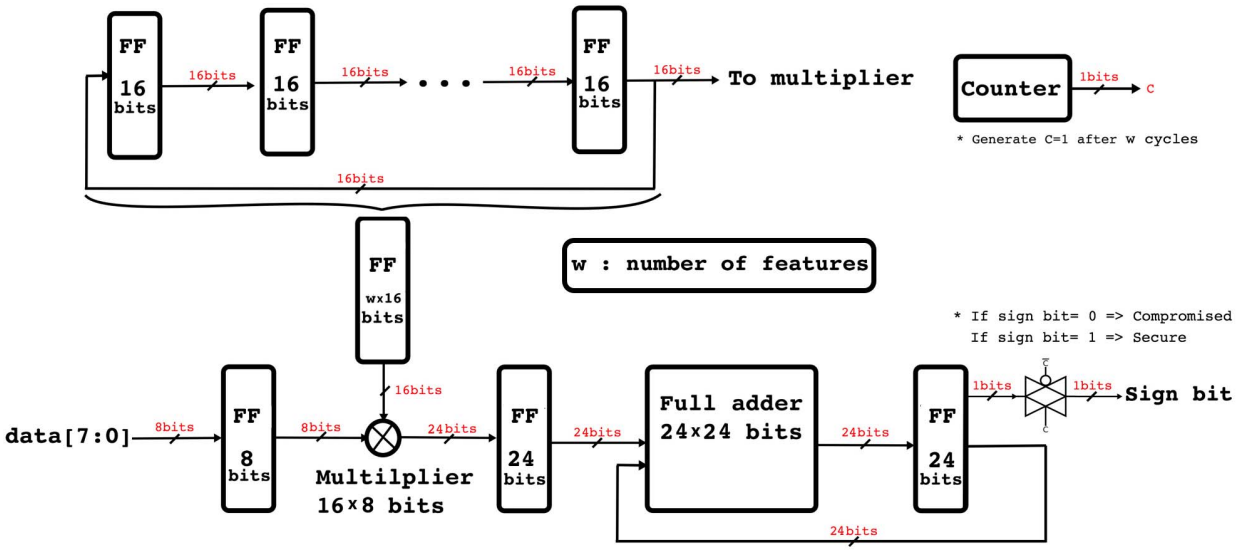

Fig. 17. Schematic of the proposed ML system for detecting malicious power grid activity.

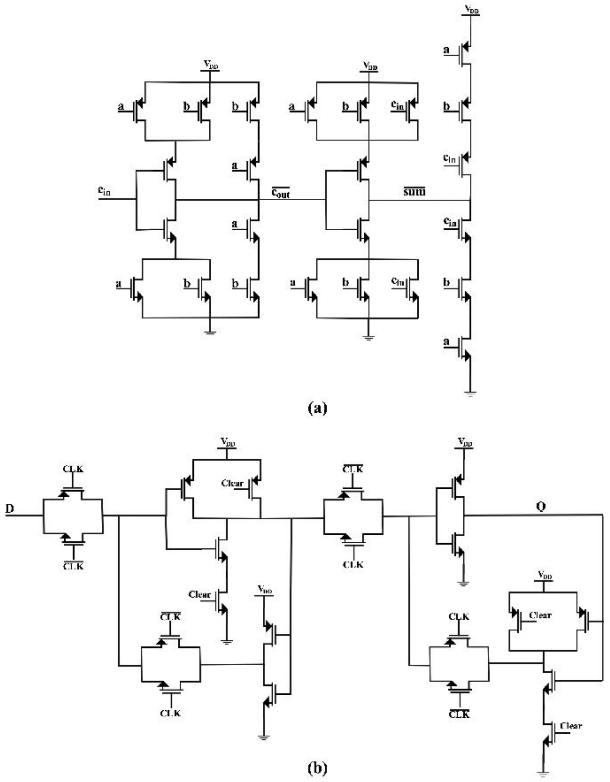

Fig. 18. Circuits for realizing the proposed ML flow, (a) mirror-adder, and (b) Master-Slave Flip-Flop.

B. Simulation Results

Successful classification of the security level is demonstrated in Fig. 19 for four consecutive prediction periods. The proposed ML classifier is trained offline and the model weights are stored in the MS-FFs. At 494 ns, the system is fully on and detection mechanism is activated. In the first clock cycle, the first feature is multiplied by a corresponding weight. The first multiplication result is stored in the decision register and simultaneously the second sampled feature is linearly transformed in the second cycle. Multiplication results for the remaining features are accumulated in the decision register for 30 subsequent cycles. Based on (8), a positive and negative value of the decision register yields, respectively, a compromised and secure prediction. Thus, at the end of each 31-cycle decision period, the PDN security level is determined based on the sign bit ('1' for secure and '0' for compromised). The decision register is reset between every two successive decisions. Decision length for the proposed system is 354 ns.

The average power consumption of the detection system decreases linearly with lower sampling frequency. At 85 MHz, malicious power activity can be detected within a short period of time of 354 ns with accuracy of 88%. Alternatively, in a successful PAA, the probe remains connected to PDN for long periods of time (in the order of magnitude of seconds or even minutes). By lowering the sampling frequency, the power consumption of the proposed detection system can be significantly reduced, while maintaining the accuracy of the predictions. The tradeoff between detection time and power consumption is illustrated in Fig. 20 based on two data points of sampling frequency and corresponding power consumption ((56 MHz, 22.12 μW) and (85 MHz, 34.71 μW)), as simulated in SPICE. Backward projection based on these data points yields the average power consumption of only 32 nW with decision time of 30 μs, and prediction accuracy of 88%.

The effect of the proposed PAA countermeasure on power, area, and throughput of the system is listed in Table I along with the existing state-of-the-art solutions against PAA [19], [20], [26]–[28]. The primary objective of the existing hardware security methods is to reduce the correlation between the system operations, (*e.g.*, secret key processing) and related power profiles. These preventive means increase the measurement to disclosure (MTD) metric which is defined as the number of power traces required for executing a successful PAA (*e.g.*, extracting the correct secret key).

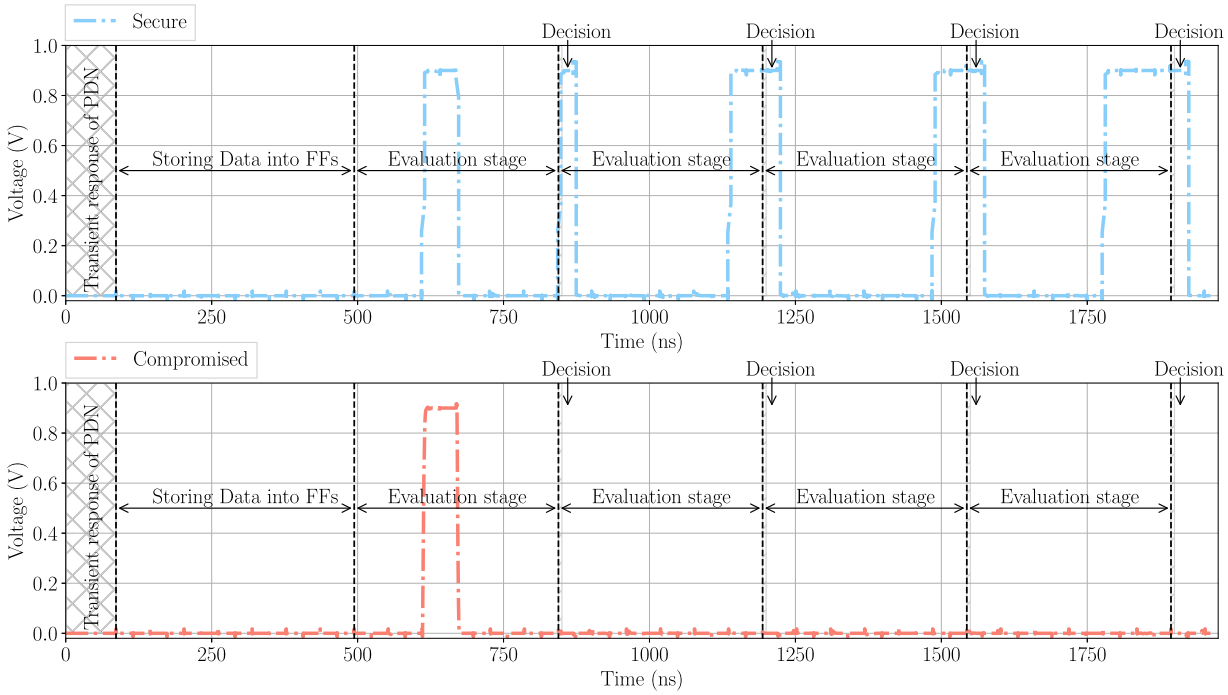

Fig. 19. PAA detection in an IBM microprocessor with decision throughput of 2.8 MHz. Circuit level simulation shows the decision bit over four consequent detection periods in secure and compromised power grids.

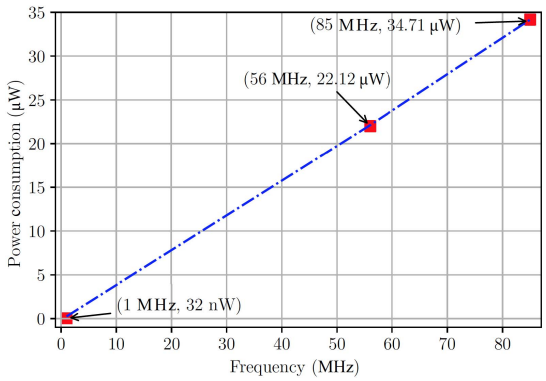

Fig. 20. Backward projection of the average power consumption with the proposed sparsely sampled power traces. Power estimation at low sampling frequency is based on two data points, as simulated in SPICE, (56 MHz, 22.12 μW) and (85 MHz, 34.71 μW).

TABLE I

SYSTEM CHARACTERISTICS OF THE EXISTING PAA COUNTERMEASURES. THE EFFECTIVENESS OF THE EXISTING AND PROPOSED TECHNIQUES IS REPORTED, RESPECTIVELY, AS $MTD > N_{protect}X$ AND PAA DETECTION ACCURACY, α_{th}.

		[19]	[20]	[26]	[27]	[28]	Current work
Technology		180nm	130nm	65nm	65nm	65nm	45nm
Method		Hiding	Hiding	Masking	Masking	Masking	Detection
Effectiveness		>156X	>2500X	*N/A*	>1000X	>1388X	up to 100%
Cost	Power	270%$^{(a)}$/1250%$^{(b)}$	33%$^{(a)}$	100%$^{(a)}$	51%$^{(a)}$	200%$^{(a)}$	14.3%$^{(a)}$/0.34%$^{(b)}$
	Area	210%$^{(a)}$/200.5%$^{(b)}$	7.2%$^{(a)}$	125%$^{(a)}$	62% $^{(a)}$	33%$^{(a)}$	7.2%$^{(a)}$/0.06%$^{(b)}$
	Throughput	*N/A*	-50%	*N/A*	*N/A*	-74.21%	0%

$^{(a)}$ Overhead is listed relatively to an unprotected AES IC.
$^{(b)}$ Overhead is listed relatively to an unprotected overall system.

Consequently, the performance of existing PAA countermeasures is typically reported as $MTD > N_{protect}X$, where X is the MTD of an unprotected system and $N_{protect}$ is the factor by which the security of the protected system is increased. Alternatively, with the proposed method, a PAA is actively detected at runtime and ultimately completely prevented. The effectiveness of the proposed method is evaluated statistically as an accuracy of PAA detection, α_{th}.

Power and area overheads of the proposed system are evaluated based on the unprotected power grid of a high-performance IBM microprocessor, *ibmpg1t* [34]. The power consumption of 27 W is determined by summing the transient current-voltage products at the individual power supply nodes in *ibmpg1t*. To detect PAAs in *ibmpg1t* with 88% accuracy, 30 PAA detection ICs should be integrated on-chip (see Fig. 12(b)). A single ML classifier (see Fig. 17) has been simulated in Cadence, exhibiting a total power consumption of 34.71 μW. To evaluate the overall power overhead of all the 30 PAA detection ICs, a typical power consumption of 3.1 mW is considered for an 8-bit ADC [45], yielding an increase of only 30 $\times$ (34.71 μW + 3.1 mW) = 94 mW (*i.e.*, 94 mW / 27 W $\times$ 100 = 0.34%) in the overall power consumption.

The physical size of the unprotected system, *ibmpg1t*, is evaluated based on the total vertical and horizontal metal lines within the top metal layer (*i.e.*, 20,679 $\times$ 20,937 lines) and a typical top metal pitch of 0.56 micrometers in 45 nm technology node [46], yielding a total chip area of (20,679 $\times$ 0.56 μm) $\times$ (20,937 $\times$ 0.56 μm) = 1.36 cm^2. Note the

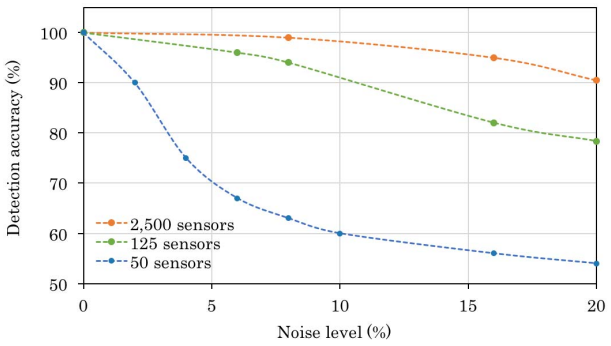

Fig. 21. Accuracy of PAA detection with the proposed logistic classifier in a noisy power delivery system.

typical for a high-performance microprocessor power density of (27 W / 1.36 cm^2) = 0.2 W/mm^2. Alternatively, based on Cadence circuit level design and typical ADC form factor [45], the proposed PAA detection ICs, including the 8-bit ADC, occupy 3,036 μm^2, resulting in a system-wide area overhead of 30 $\times$ 3,036 μm^2 = 91,080 μm^2 (*i.e.*, 0.06% of total chip area).

Securing smaller cryptographic systems, such as AES cores, with the proposed ML approach is also investigated. A typical AES core is considered that exhibits a power consumption of 33.32 mW and an area of 0.35 mm^2 [20]. Note that these smaller systems can be secured with a single on-chip sensor, yielding power and area overhead of, respectively, 14.3% and 7.2% with the proposed circuits. These costs can be further reduced by replacing the ADC-based digital classifiers with more compact and efficient analog classifiers [32]. The design of these ICs is, however, out of the scope of this work.

Both power and area overheads with the proposed security method are significantly lower than with existing approaches often due to the preventive nature of those existing methods that require the entire system to be redesigned for improving the resilience against PAAs. Alternatively, with the proposed approach, PAAs are actively detected at runtime with a few low-complexity ICs.

To validate the system-wide accuracy of PAA detection across the IBM microprocessor grid under PVT variations, the proposed system is simulated with different number of sensors. The SPICE simulation results are shown in Fig. 21, exhibiting detection accuracy of above 90% under 2%, 11%, and 20% variations with, respectively, 50, 125, and 2,500 sensors. For those high-end VLSI systems, larger number of sensors should be preferred for accurately detecting PAAs across the device under high PVT variations (*e.g.*, 125 sensors for above 90% accuracy under 10% variations). Alternatively, certain security sensitive regions within high-end devices or within smaller integrated systems with less variations can ultimately be secured with high accuracy and few sensors.

VI. Conclusions

While numerous preventive countermeasures against power attacks exist, detection of system intrusion has not been previously explored. A ML-based countermeasure against PAAs is proposed that allows to detect malicious probing of a power grid in a running device. Within the proposed framework, non-typical voltage variations induced within power grid by a malicious probe are sensed on-chip, digitized, and analyzed using machine learning techniques. A logistic classifier is chosen due to its low design and training complexity, efficient power characteristics, and high PAA detection performance. With the proposed approach, the required number of features and resolution of the sampled power traces are selected based on the desired security level. The effect of a PAA on power grid characteristics is, for the first time, analytically formulated and the tradeoffs between the security level of an integrated system and power grid characteristics are investigated based on the proposed closed-form expressions. ML classification is demonstrated as a robust, low design complexity technique for accurately and power efficiently detecting power attacks. The effectiveness of the proposed method increases in those low impedance power grids with sparsely placed power supplies.

The PAA detection system is designed in SPICE and simulated in a 45 nm standard CMOS process. The simulation results validate the performance and functionality of the proposed approach. A power attack on a benchmarked IBM microprocessor is successfully detected within 354 ns in 88% of PAA configurations, exhibiting low power consumption of 34.71 μW. The proposed framework does not affect the performance of the cryptographic device and can be configured to consume nanowatts with longer detection periods. This paper is the first to introduce PAA detection at runtime and demonstrate the effectiveness of power efficient, compact ML classifiers for increasing the resilience of modern ICs for side-channel attacks.

References

[1] P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," in *Advances in Cryptology* (Lecture Notes in Computer Science), vol. 1666. Berlin, Germany: Springer, Aug. 1999, pp. 388–397.

[2] P. C. Kocher, "Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems," in *Advances in Cryptology—CRYPTO*, vol. 1109. Berlin, Germany: Springer, Aug. 1996, pp. 104–113.

[3] D. A. Osvik, A. Shamir, and E. Tromer, "Cache attacks and countermeasures: The case of AES," in *Proc. Cryptographers Track RSA Conf. (CT-RSA)*, vol. 3860, Feb. 2006, pp. 1–20.

[4] J.-J. Quisquater and D. Samyde, "ElectroMagnetic analysis (EMA): Measures and counter-measures for smart cards," in *Smart Card Programming and Security—E-Smart)*, vol. 2140. Berlin, Germany: Springer, Sep. 2001, pp. 200–210.

[5] P. Kocher, J. Jaffe, B. Jun, and P. Rohatgi, "Introduction to differential power analysis," *J. Cryptograph. Eng.*, vol. 1, no. 1, pp. 5–27, Apr. 2011.

[6] S. Köse and E. G. Friedman, "Efficient algorithms for fast *IR* drop analysis exploiting locality," *Integr., VLSI J.*, vol. 45, no. 2, pp. 149–161, Mar. 2012.

[7] I. Vaisband and E. G. Friedman, "Stability of distributed power delivery systems with multiple parallel on-chip LDO regulators," *IEEE Trans. Power Electron.*, vol. 31, no. 8, pp. 5625–5634, Aug. 2016.

[8] S. Mangard, E. Oswald, and T. Popp, *Power Analysis Attacks: Revealing the Secrets of Smart Cards*. New York, NY, USA: Springer, 2008.

[9] K. Tiri, M. Akmal, and I. Verbauwhede, "A dynamic and differential CMOS logic with signal independent power consumption to withstand differential power analysis on smart cards," in *Proc. Eur. Solid-State Circuits (ESSCIRC)*, Sep. 2002, pp. 403–406.

[10] K. Tiri and I. Verbauwhede, "A logic level design methodology for a secure DPA resistant ASIC or FPGA implementation," in *Proc. Design, Automat. Test Eur. Conf. Exhib. (DATE)*, vol. 1, Feb. 2004, pp. 246–251.

[11] D. Sokolov, J. Murphy, A. Bystrov, and A. Yakovlev, "Design and analysis of dual-rail circuits for security applications," *IEEE Trans. Comput.*, vol. 54, no. 4, pp. 449–460, Apr. 2005.

[12] D. Sokolov, J. Murphy, A. Bystrov, and A. Yakovlev, "Improving the security of dual-rail circuits," in *Proc. Int. Workshop Cryptograph. Hardw. Embedded Syst. (CHES)*, vol. 3156, Aug. 2004, pp. 282–297.
[13] M. Bucci, L. Giancane, R. Luzzi, and A. Trifiletti, "Three-phase dual-rail pre-charge logic," in *Proc. Int. Workshop Cryptograph. Hardw. Embedded Syst. (CHES)*, vol. 4249, Oct. 2006, pp. 232–241.
[14] M. Bucci, L. Giancane, R. Luzzi, and A. Trifiletti, "A flip-flop for the DPA resistant three-phase dual-rail pre-charge logic family," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 11, pp. 2128–2132, Nov. 2012.
[15] H.-T. Ng and D. J. Allstot, "CMOS current steering logic for low-voltage mixed-signal integrated circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 5, no. 3, pp. 301–308, Sep. 1997.
[16] Z. Toprak and Y. Leblebici, "Low-power current mode logic for improved DPA-resistance in embedded systems," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, vol. 2, May 2005, pp. 1059–1062.
[17] F. Mace, F.-X. Standaert, I. Hassoune, J.-D. Legat, and J.-J. Quisquater, "A dynamic current mode logic to counteract power analysis attacks," in *Proc. 19th Int. Conf. Design Circuits Integr. Syst. (DCIS)*, Nov. 2004, pp. 186–191.
[18] K. Tiri and I. Verbauwhede, "A digital design flow for secure integrated circuits," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 25, no. 7, pp. 1197–1208, Jul. 2006.
[19] D. D. Hwang *et al.*, "AES-based security coprocessor IC in 0.18-μm CMOS with resistance to differential power analysis side-channel attacks," *IEEE J. Solid-State Circuits*, vol. 41, no. 4, pp. 781–792, Apr. 2006.
[20] C. Tokunaga and D. Blaauw, "Securing encryption systems with a switched capacitor current equalizer," *IEEE J. Solid-State Circuits*, vol. 45, no. 1, pp. 23–31, Jan. 2010.
[21] T. Popp, M. Kirschbaum, T. Zefferer, and S. Mangard, "Evaluation of the masked logic style MDPL on a prototype chip," in *Proc. Int. Workshop Cryptograph. Hardw. Embedded Syst. (CHES)*, vol. 4727, Sep. 2007, pp. 81–94.
[22] E. De Mulder, B. Gierlichs, B. Preneel, and I. Verbauwhede, "Practical DPA attacks on MDPL," in *Proc. IEEE Int. Workshop Inf. Forensics Security (WIFS)*, Dec. 2009, pp. 191–195.
[23] A. Moradi, M. Salmasizadeh, and M. T. M. Shalmani, "Power analysis attacks on MDPL and DRSL implementations," in *Proc. Int. Conf. Secur. Cryptol. (ICISC)*, vol. 4817, Nov. 2007, pp. 259–272.
[24] T. Popp and S. Mangard, "Implementation aspects of the DPA-resistant logic style MDPL," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2006, pp. 2913–2916.
[25] M. Bucci, R. Luzzi, M. Guglielmo, and A. Trifiletti, "A countermeasure against differential power analysis based on random delay insertion," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 4, May 2005, pp. 3547–3550.
[26] I. Levi, O. Keren, and A. Fish, "Data-dependent delays as a barrier against power attacks," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 8, pp. 2069–2078, Aug. 2015.
[27] S.-C. Chung, C.-Y. Yu, S.-S. Lee, H.-C. Chang, and C.-Y. Lee, "An improved DPA countermeasure based on uniform distribution random power generator for IoT applications," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 9, pp. 2522–2531, Sep. 2017.
[28] D. Bellizia, S. Bongiovanni, P. Monsurrò, G. Scotti, A. Trifiletti, and F. B. Trotta, "Secure double rate registers as an RTL countermeasure against power analysis attacks," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 7, pp. 1368–1376, Jul. 2018.
[29] M. Avital, H. Dagan, O. Keren, and A. Fish, "Randomized multitopology logic against differential power analysis," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 4, pp. 702–711, Apr. 2015.
[30] M. Bucci, M. Guglielmo, R. Luzzi, and A. Trifiletti, "A power consumption randomization countermeasure for DPA-resistant cryptographic processors," in *Proc. Int. Workshop Power Timing Modeling, Optim. Simulation (PATMOS)*, vol. 3254, Sep. 2004, pp. 481–490.
[31] I. Vaisband and E. G. Friedman, "Energy efficient adaptive clustering of on-chip power delivery systems," *Integr., VLSI J.*, vol. 48, pp. 1–9, Jan. 2015.
[32] Z. Wang and N. Verma, "A low-energy machine-learning classifier based on clocked comparators for direct inference on analog sensors," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 11, pp. 2954–2965, Nov. 2017.
[33] S. R. Nassif, "Power grid analysis benchmarks," in *Proc. Asia South Pacific Design Autom.*, Sep. 2008, pp. 376–381.
[34] P. L. Nassif Sani, Zhuo Li. (2011) *IBM Power Grid Benchmarks*. [Online]. Available: http://dropzone.tamu.edu/~pli/PGBench
[35] M. Bucci, L. Giancane, R. Luzzi, M. Marino, G. Scotti, and A. Trifiletti, "Enhancing power analysis attacks against cryptographic devices," *IET Circuits, Devices Syst.*, vol. 2, no. 3, pp. 298–305, Jun. 2008.
[36] E. De Mulder, "Electromagnetic techniques and probes for side-channel analysis on cryptographic devices," Ph.D. dissertation, Katholieke Univ. Leuven, Leuven, Belgium, 2010.
[37] S. Skorobogatov, "Using optical emission analysis for estimating contribution to power analysis," in *Proc. IEEE Workshop Fault Diagnosis Tolerance Cryptograph. (FDTC)*, Sep. 2009, pp. 111–119.
[38] G. Venezian, "On the resistance between two points on a grid," *Amer. J. Phys.*, vol. 62, no. 11, pp. 1000–1004, Nov. 1994.
[39] H. Guntur, J. Ishii, and A. Satoh, "Side-channel attack user reference architecture board SAKURA-G," in *Proc. IEEE 3rd Global Conf. Consum. Electron. (GCCE)*, Oct. 2014, pp. 271–274.
[40] F. E. Harrell, Jr., "Ordinal logistic regression," in *Regression Modeling Strategies*. New York, NY, USA: Springer, 2015, pp. 311–325.
[41] S. Raschka and V. Mirjalili, *Python Machine Learning*, 2nd ed. Birmingham, U.K.: Packt, 2017.
[42] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.
[43] I. P.-Vaisband *et al.*, *On-Chip Power Delivery and Management*, 4th ed. Cham, Switzerland: Springer, 2016.
[44] Z. Wang, R. E. Schapire, and N. Verma, "Error adaptive classifier boosting (EACB): Leveraging data-driven training towards hardware resilience for signal inference," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 4, pp. 1136–1145, Apr. 2015.
[45] B. Murmann. *ADC Performance Survey 1997–2018*. [Online]. Available: http://web.stanford.edu/~murmann/adcsurvey.html
[46] P. Moon *et al.*, "Process and electrical results for the on-die interconnect stack for intel's 45 nm process generation," *Intel Technol. J.*, vol. 12, no. 2, pp. 87–92, May 2008.

Farid Kenarangi (S'18) received the B.Sc. degree in electrical engineering from the University of Tabriz, Tabriz, Iran, in 2015. He is currently pursuing the Ph.D. degree in electrical engineering with The University of Illinois at Chicago, Chicago, IL, USA, under the supervision of Prof. I. Partin-Vaisband. His current research interests include hardware security, machine learning integrated circuits, analog design, and on-chip power delivery and management. He was a recipient of the 2017 University of Illinois at Chicago Chancellor's Graduate Research Award.

Inna Partin-Vaisband (S'12–M'15) received the B.Sc. degree in computer engineering and the M.Sc. degree in electrical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, in 2006 and 2009, respectively, and the Ph.D. degree in electrical engineering from the University of Rochester, Rochester, NY, USA, in 2015.

Between 2003 and 2009, she held a variety of software and hardware research and development positions with Tower Semiconductor Ltd., Israel; G-Connect Ltd., Israel; and IBM Ltd., Israel. She is currently an Assistant Professor with the Department of Electrical and Computer Engineering, The University of Illinois at Chicago. Her research is currently focused on innovation in the area of distributed power delivery and locally intelligent power management that facilitates performance scalability of heterogeneous ultra-large-scale integrated systems. Special emphasis is placed on statistical analysis, hardware security, and emerging technologies such as wireless power transfer and simultaneous wireless information and power transfer. She served on the technical program and organization committees for various conferences. She is an Associate Editor of the *Microelectronics Journal*.