SN Computer Science (2023) 4:77
https://doi.org/10.1007/s42979-022-01491-5

ORIGINAL RESEARCH

# A Side-Channel Evaluation of On-chip Vdd Distribution Network with Decoupling Capacitance

**Ravikumar Selvam**[1] · **Akhilesh Tyagi**[1]

Received: 1 May 2022 / Accepted: 3 November 2022 / Published online: 2 December 2022


## Abstract

Design of an on-chip power (Vdd) distribution network (PDN) is an important step in modern-day integrated circuit design. Several algorithms and tools exist to enhance the performance of power distribution with reduced noise. Unfortunately, however, power distribution network is the primary target of power analysis side-channels. In this paper, the values and topologies of decoupling capacitance incorporated into on-chip power distribution network are analyzed for side-channel resistance. We show that an on-chip power distribution network with decoupling capacitance thwarts the power side-channel attacks. The proposed design uses multiple decoupling capacitances along the power lanes in a distributed fashion to suppress the data leakage from the sensitive logic blocks. Grid-style and tree style power distribution networks are designed and evaluated to assess the effect of decoupling capacitance on power side-channel resistance. A novel, approximate heuristics to extract the feature vector from the switching current ($I$) of the internal logic blocks is developed and applied in the analysis. Machine learning (ML) classifiers are used to quantify the side-channel effectiveness in terms of success rate for the power side-channel adversary. The test circuit for proposed techniques is implemented using FreePDK 45 nm technology library. Spice level simulations are conducted with various decoupling capacitance values. With only 39.33% area overhead, the power side-channel adversary success rate is reduced from 83 to 21% for tree-style PDN and from 68 to 17% for grid-style PDN with decoupling capacitance.

**Keywords** Side-channel attack · Power distribution network · Decoupling capacitance

## Introduction

When the secret values are stored in the hardware registers of a computing logic block, the data-dependent correlation between the power waveform features at the power pin and the secret data can be exploited by a side-channel attack (SCA). This poses a very potent threat to integrated circuits. Power analysis attack is a passive side-channel attack targeting integrated circuits by eavesdropping on the power pins using low impedance probes for data-dependent leakage. Kocher et al. reported the first successful side-channel attack in 1999 by analyzing the power consumption measurements from tamper-resistant devices [1]. Over the years, researchers have explored different techniques such as template attack [2], correlation-enhanced power analysis collision attack [3], and many more to exploit power measurements for data-dependent switching activity at the power pin. Data-dependence can be removed by making the power consumption profile statistically data value agnostic. Cryptographic techniques to encode data are also used to mitigate power side-channel attacks, but they are less effective at eliminating the data-dependent correlations with power. Side-channel leakage is addressed at various abstraction stages of the design hierarchy. The design abstractions broadly include architecture design stage, logic design stage, and physical design stage, as shown in Fig. 1.

In the architecture design stage, multiple functional blocks and the interconnections between the blocks are

Akhilesh Tyagi is contributed equally to this work.

This article is part of the topical collection "Smart and Connected Electronic Systems" guest edited by Amlan Ganguly, Selcuk Kose, Amit M. Joshi and Vineet Sahula.

✉ Ravikumar Selvam
rkselvam@iastate.edu

Akhilesh Tyagi
tyagi@iastate.edu

[1] Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, USA

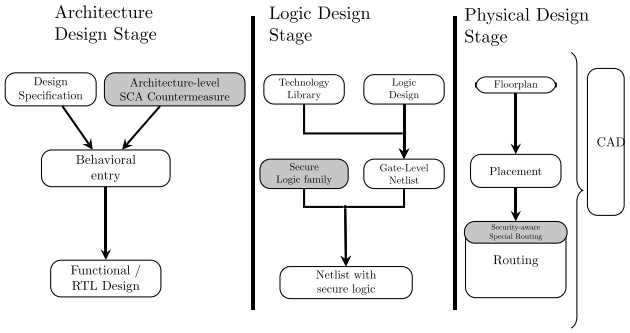


**Fig. 1** Design hierarchy and SCA countermeasures

specified at the system level. The random noise/delays can be inserted near the sensitive blocks as a side-channel countermeasure. The random noise block performs random operations to generate/add noise to the power while processing the secret, sensitive values [4, 5]. Other countermeasures such as shuffling [4] prevent power side-channel attacks by introducing a random delay between functions to break synchronization between the traces. This method increases the number of samples required to mount a successful attack significantly.

Several logic level side-channel prevention techniques have been proposed. These can be broadly classified into two categories: masking and hiding. Masking splits sensitive intermediate variables into multiple shares. One of the best-known masking countermeasures is a secret sharing scheme, which was initially developed by Ishai et al. [6]. Subsequently, several techniques like threshold implementation [7], RNS secure circuits [8] were developed to improve the resistance against the power side-channel attack against glitches. In the hiding techniques, the data dependence on power is reduced by equalizing the current draw for all circuit operations on all data values through customized logic circuits such as sense amplifier based logic (SABL) [9], wave dynamic differential logic (WDDL) [10, 11]. Most of these countermeasure techniques use resources proportional to statistical difficulty in detecting data-dependent correlations. They do provide considerable side-channel resistance to the naïve implementations. However, the available resources in practical circuit implementations limit the chip designers in achieving the ultimate power analysis resistance. The secure logic families use redundant functional blocks causing significant area overhead, which results in higher power consumption. The available design space to trade design resources with side-channel resistance is minimal.

The physical design stages are automated using computer-aided design (CAD) tools to run the following functions: partition, floorplan, placement, and routing. The CAD tools optimize the logic level design for better area, speed, and power. The physical CAD design flow can handle private information through cryptography, but the security is not considered. In this paper, we target the power distribution network design with decoupling capacitance to hide the power and data correlation leakage. The design schema distributes the capacitance on the power distribution network to suppresses the data-dependent power leakage. This technique breaks the synchrony between the logic block actions on secret data and the corresponding power observations at the Vdd pin. This reduces the data correlation impact at the pin observation point of the off-chip power network. The activation schedule of capacitance can also be randomized for additional power noise. These techniques tend to make the capacitance switching schedule combinatorially challenging for the adversary.

*Our contributions:* This paper demonstrates and quantifies the side-channel resistance of power distribution networks with decoupling capacitances against a power side-channel attack. We abstract the power signatures of the circuit operations as feature vectors both at the internal circuit node and external pin. Heuristics to propagate features from the logic blocks to the power pins with decoupling capacitance are developed and discussed in Ref. [12].[1] The LED block cipher round function is implemented as the test circuit in both tree-style and grid-style PDN using NCSU FreePDK 45nm technology library. The side-channel resistance is quantified by the success rates of machine learning classifiers such as Naive Bayes (NB), linear discriminant analysis (LDA), and quadratic discriminant analysis (QDA).

We also investigate the random activation schedule of decoupling capacitance in the power distribution network for further side-channel resistance enhancement. The results confirm that the distributed decoupling capacitance across the power path provides significantly better side-channel resistance compared to a single large decoupling capacitance. Finally, area overhead and security trade-offs for the power distribution network with decoupling capacitance are discussed.

This paper is organized as follows. "Related Work" briefly presents related background information, including a brief overview of the power distribution network and related work from the literature. The data leakage propagation on a power distribution network is discussed in "Side-Channel Perspective". "Experiments and Analysis" presents a practical implementation of power distribution networks with the corresponding results. Finally, "Conclusions" summarizes and concludes the paper.

[1] The cited paper is the conference version of current work with preliminary results which published in 7th IEEE International Symposium on Smart Electronic Systems, iSES 2021.

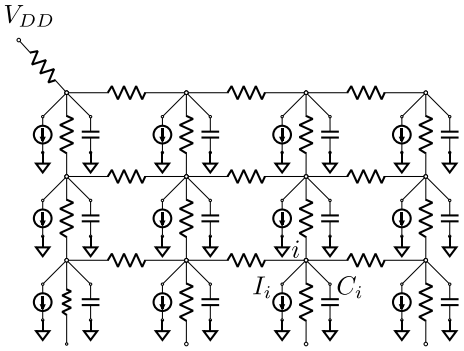


**Fig. 2** Power distribution network

## Related Work

With the ever-increasing net and gate counts, it is challenging to distribute Vdd and ground to all the transistor source and drain terminals with uniform and acceptable IR voltage drop. For this reason, a power distribution network is considered to be one of the important and primary components in modern ICs. The power/ground signals are distributed to interconnect the logic cells with the VDD/GND pin. The on-chip power and ground buses are routed in metal-1 or metal-2 layers. The metal interconnections are modeled with simple RC circuits, as shown in Fig. 2. The current symbol at each node represents the current demands of the corresponding logic block. The primary objectives of the PDN are to provide a stable voltage to all the logic cells with low noise to satisfy their peak and average power demands. The power distribution network is constructed with small resistance to reduce the IR loss over the metal layer.

Let us assume that the node $i$ undergoes a signal transition with the current drawn from the power supply denoted as $I_i$ in Fig. 2. The external pin connects to all internal nodes with minimal IR drop. Though the external power pins are held at voltage $V_{DD}$, the internal node voltage fluctuates according to the switching activity of the logic node. Large switching activity leads to higher current draw, which leads to the voltage $V_{DD}$ − IR at the logic block. This fluctuation in voltage and current over time can be observed at the external power pin through the propagation across the power distribution network. The trace path between the internal logic node $i$ and the external pin crosses the junction point from the other logic nodes. The power path between external $V_{DD}$ and the logic node $i$ can be designed in either a tree-style or grid-style. The structure of on-chip PDN design plays a significant role in achieving the designer's requirement on electrical characteristics. In tree-style PDN, the logic nodes are connected with a dedicated power path to the $V_{DD}$. Tree style PDN design supports limited interconnect resources requirements for low-cost ICs. On the other hand, the grid-style PDN connects the logic node and $V_{DD}$ through multiple paths. This provides robust and high-performance current distribution for high-speed integrated circuits.

Most of the research efforts in PDN design focus on modeling IR drop [13–15] and power switching noise [16, 17]. These models are used to iteratively refine the PDN design for low IR voltage drop and noise. Mayhew and Muresan [18, 19] proposed a countermeasure technique by placing an nMOS gate capacitance near the logic block. The nMOS capacitance momentarily supplies the current in-demand and decouples the logic block from the power distribution network. They performed a correlation power analysis attack and reported a success value indicator for the effectiveness of the countermeasure. Dofe and Yu [20] presented a correlation power analysis attack in 3D ICs by exploiting power distribution noise. Kenarangi et al. [21] presented an on-chip sensor to detect the real-time power analysis attack by monitoring the voltage variation induced on the power distribution network. Similarly, a CMOS-based self decoupling battery cell system was presented for powering the security-sensitive modules when the voltage drop reaches the minimum threshold [22]. Table 1 shows summary of the

**Table 1** Summary of the related work in PDN

| SI no. | Category | References | Summary |
|---|---|---|---|
| 1 | IR drop | [13–15] | This category reports the work related to IR drop analysis using spice simulation and analyzes silicon substrate for effective power distribution |
| 2 | Power switching noise | [16, 17] | In this category, the authors discuss simultaneous switching noise (SSN) through the parasitic present in the PDN rail and characterize it using lumped inductive-resistive-capacitive RLC model |
| 3 | 3D ICs and on-chip sensor | [18–22] | This category of the research paper discusses the security of 3D ICs and on-chip sensor for the real-time power analysis attack detection |
| 4 | Our proposed work | | In this paper, we analyze the power distribution network for data leakage and develop heuristics on feature vector propagation from the logic node to the external power pin. Decoupling capacitance is investigated with various configurations on grid-style and tree-style PDN |

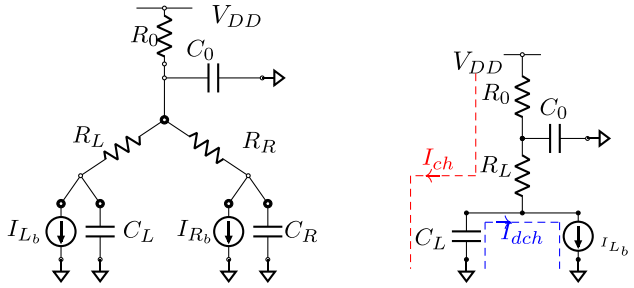


**Fig. 3** Switching current propagation

relevant research article published in analyzing the power distribution network. Several of these efforts focus only on the runtime detection of power analysis attacks. Our work investigates different styles of power distribution networks and the effect of distributed decoupling capacitance across the design against power side-channel attacks. Further, we also study the random scheduling of decoupling capacitance and the sampling requirements for power side-channel adversary to build machine learning models.

## Side-Channel Perspective

We have described the power distribution network from a high-level side-channel perspective. In this section, we amplify the PDN attributes that are relevant to the side-channel analysis. As mentioned in "Related Work", the power distribution networks are represented as an RC circuit, where $R$ and $C$ represent the parasitic resistance and capacitance of the metal wire in the PDN implementation. The $V_{DD}$ nodes of the internal circuits are branches of the power supply pin. Hence, we illustrate the activities of the power distribution network with a simple RC circuit, which is shown in Fig. 3. The current node $I_{R_b}$ and $I_{L_b}$ represent the current demand for the switching activity of the internal circuits $R_b$ and $L_b$, the left and right branches of a node in a PDN tree. The parasitic components of the power grid metal layer $R_R$, $R_L$, $C_R$, and $C_L$ are resistance and capacitance of left and right branches. The parasitic resistance and capacitance near $V_{DD}$ pin are denoted as $R_0$ and $C_0$, respectively. In power analysis attack, the adversary only have access to the external power supply pin $V_{DD}$. Hence, the power leakage captured at $V_{DD}$ by the adversary across the off-chip power network is equivalent to the current drawn across the resistor $R_0$. According to Kirchoff's current law, the current across the resistor $R_0$ is given by Eq. 1.

$$I_0 = I_L + I_R + I_{C_0} \tag{1}$$

The parasitic capacitances of the power grid $C_0$, $C_R$, and $C_L$ have a negligible effect on the overall power consumption. This causes the current $I_0$ to equal the sum of the current drawn by the circuits $R_b$ and $L_b$.

To study the power side-channel attributes of PDN, we assume the circuit block $L_b$ as a sensitive circuit block that processes secret data values. Circuit block $R_b$ performs non-sensitive circuit operations and presents noise to the power side-channel adversary. The adversary's goal is to capture the power-data correlation leakage caused by the switching activity of the internal circuit block $L_b$ through the off-chip power supply network. The current trace contains all the information for the extraction of a specific feature vector, such as peaks and time slope. The peak correlates with the state of an internal circuit that has maximum switching during computation. The peak values in the current profile vary with the amount of data-dependent charging and discharging of the transistors. All logic gates are designed to produce stable outputs within their time constraints based on the clock. The scheduling specification allows the logic gates to switch at different times, which determines the peak current and the current slopes. The slope specifies the rate of change in the current between successive stages. The resistive path of the power distribution network only scales down/up the amplitude of the current peak through the resistive spread, but the number of peaks remains invariant over PDN RC parameters.

The CMOS-based decoupling capacitor can hold sufficient charge to satisfy the current demand of the logic block $L_b$ over a small time window. The decoupling capacitance allows the sensitive circuit to decouple from the $V_{DD}$ pin virtually. During this stage, the capacitor charges/discharges the logic block $L_b$. A minimal current is drawn from the $V_{DD}$ pin, which affects the propagation of feature vectors to the $V_{DD}$ pin.

The decoupling capacitor charge/discharge cycle is illustrated in Fig. 3. When the circuit powers on, the decoupling capacitance is charged in a short time. Later, this charge is used to supply the current to logic cells. The power distribution network is continuously supplied through the $V_{DD}$ pin as per the circuit specification. The stability of the input power supply allows for constant current flow in every branch of the power distribution network. When the switching current demand is larger than the constant branch current, i.e., $I_{L_b} > I_{R_L}$, the capacitor starts to discharge. The discharging rate is defined based on active connections of NMOS and PMOS transistors at the logic cell.

Side-channel leakage is measured by the current demand at $V_{DD}$ node, which is caused by the propagation of switching current $I_{L_b}$. With presence of decoupling capacitance, the propagation of the current profile is severely affected

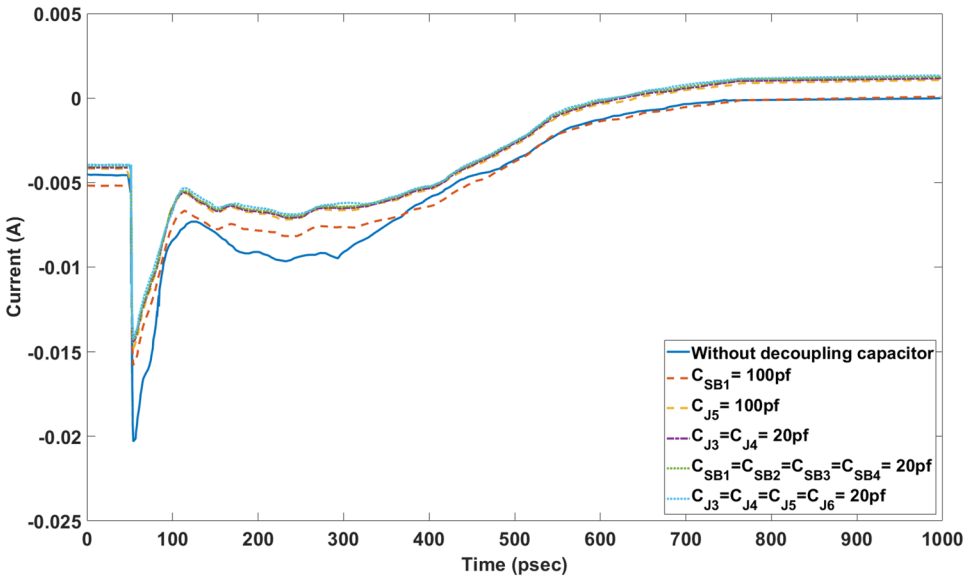


**Fig. 4** Switching current at $V_{DD}$

by the discharge current of $C_L$. The discharge cycle of the decoupling capacitance is data-dependent in which the discharge current profile looks similar to the switching current profile of the logic cell. Then the branch current $I_{R_L}$ can be written as:

$$I_{R_L} = I_{L_b} - \frac{\Delta V_{L_b}}{R_{L_b}}(e^{-t_d/(R_{L_b} * C_L)}), \tag{2}$$

where $t_d$ is the time interval for decoupling capacitance discharge cycle. $\Delta V_{L_b}$ is voltage drop at logic node $L_b$.

After time $t_d$, the switching current of the logic block $I_{L_b}$ becomes less than the branch current $I_{R_L}$ and the decoupling capacitance shifts mode to the charging cycle. In this time period, The capacitor draws current from the $V_{DD}$ pin to recharge at a constant rate. Now the total branch current is given as:

$$I_{R_L} = I_{L_b} + \frac{\Delta V_{L_b}}{R_L + R_0}(e^{-t_c/((R_L+R_0) * C_L)}) \tag{3}$$

Where,

$t_c$ is the total time period of the charging cycle.
$R_{t_L b}$ is the total resistance of the path from $V_{DD}$ to logic block $L_b$.

In the interest of a power analysis attack, the side-channel adversary continuously observes the $V_{DD}$ pin for the feature vectors generated from the switching current of the sensitive logic block. The capacitance charge/discharge cycle severely affects the propagation of peak and slope observed at $V_{DD}$ pin. The sensitivity of decoupling capacitance on power tree is illustrated in Figs. 4 and 5. During the discharge cycle, the decoupling capacitance suppresses the current peak values within its branch. Further, the propagation of peak value through the power distribution network adds more switching noise from the other logic cells. The switching noise makes the feature set observed at $V_{DD}$ pin unreliable to an adversary for mounting the attack.

From Eq. 2, it is clear that the discharge current entirely depends on the logic cells. Hence, the rate of change in the current of the logic blocks $L_b$ appears to be compensated by the discharge current of the decoupling capacitance. This action ensures minimal variation in the branch-current propagation that results in a low time slope. The charging phase of the decoupling capacitor is ultimately data independent. The charging rate of the decoupling capacitance is constant over PDN RC values, whereas the period of the discharge cycle is highly data-dependent.

The newly attached decoupling capacitance detaches the logic block from the power distribution path and supplies the current in demand. Thus the decoupling capacitance influences the propagation of feature set vectors over the power distribution network. The power trace observed at external $V_{DD}$ pin with decoupling capacitance is difficult to decode. The feature vector, such as local maximum and minimum values, are extracted from the voltage values to evaluate the effect of decoupling capacitance on the power tree. These feature vectors combined with the feature set extracted from the switching current form the data for detailed analysis.

Multiple capacitors at various junction points muddle up the relationship between the current draw as observed at the target logic block and the current draw observed at the $V_{DD}$ pin. Furthermore, conceptually if the current draw of the logic block from the decoupling capacitors could be broken up into random epochs through a switch connecting

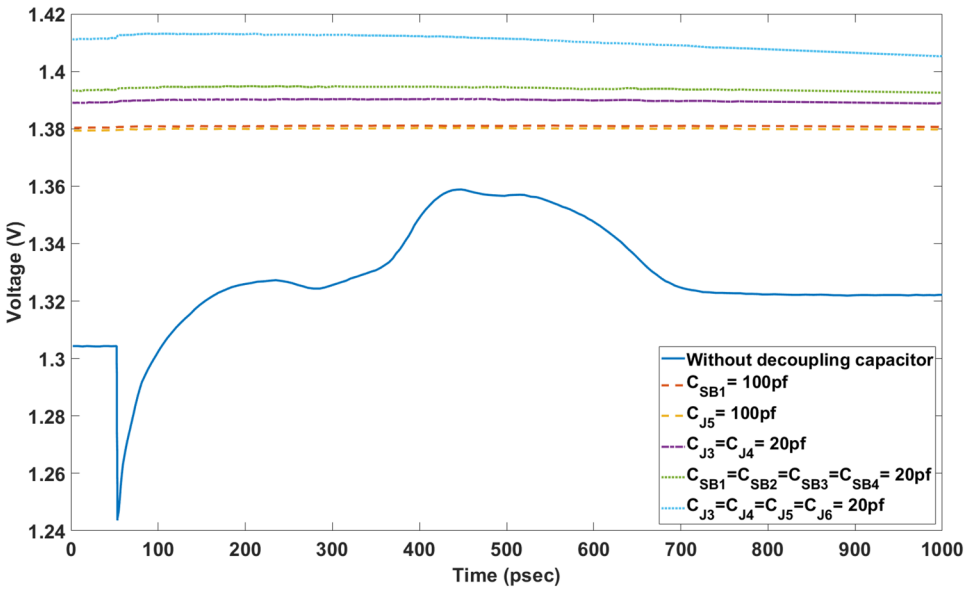


**Fig. 5** Voltage drop at $V_{DD}$

or disconnecting the decoupling capacitor, the synchrony between the current draw at the logic block and $V_{DD}$ pin would be further broken. Randomized connectivity of multiple ($m$) decoupling capacitors creates a random schedule potentially increasing the dimensionality of the machine learning model at $V_{DD}$ pin by a factor of $2^m$. We explore these scenarios in "Experiments and Analysis".

## Experiments and Analysis

This section discusses the different styles of power distribution networks and evaluates their side-channel resistance with decoupling capacitances. The logic nodes of the power distribution network are designed using an AES-like encryption function called LED-64. The LED-64 block cipher is a lightweight encryption function that encodes 64-bit plaintext using a 64-bit key. The block cipher is constructed using AES-like round functions such as AddConstant, AddRoundKey, S-box, and Mixcolumn, which are denoted as ARC, ARK, SB, and MXC, respectively. These functional blocks form a pipelined architecture, where during each cycle, all functional blocks are active in parallel on the power tree. For all our experiments, each functional block has separate inputs for the computation, and the outputs are connected to a fixed load capacitance of 0.5 nf. The primary goal of our experiment is to investigate the impact of on-chip decoupling capacitance against power side-channel attacks, which does not necessarily evaluate the strength of the block cipher implementation.

### Analysis on Tree-Style PDN

The tree-style PDN is a commonly utilized for connecting the external power source to the logic nodes. In this study, A six-layer deep power tree is constructed as shown in Fig. 6. Each layer has one functional block and one junction point. The functional blocks on the power tree are constructed using the round functions of the LED block cipher. Each branch in the power tree is designed with its own individualized set of RC values.

In this experiment, the adversary's goal is to infer the secret data values from the feature vectors extracted from the voltage drop and switching current caused by the circuits' transitions on the power distribution network. The S-box

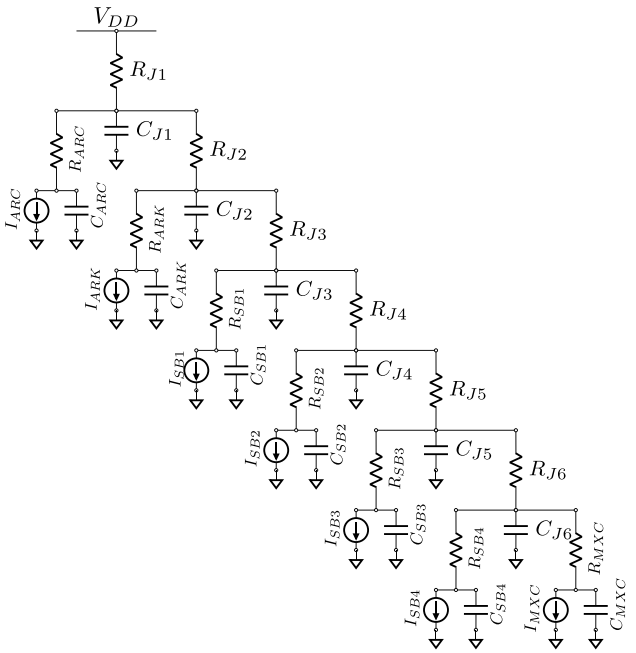


**Fig. 6** Power distribution network with LED round functions

**Table 2** Success rate of the PDN without decoupling capacitance

| Classifier | NB | LDA | QDA |
|---|---|---|---|
| Success rate | 87.32% | 81.82% | 80.37% |

function is considered to be a critical function in any encryption for the side-channel analysis; for this reason, the S-box function (SB1) is assumed to be a target function for an adversary. The effect of the on-chip decoupling capacitor is studied under the following scenarios.

1. Single large decoupling capacitance.
   (a) Adding capacitance close to the S-box.
   (b) Adding capacitance at the junction.
2. Distributed decoupling capacitance.
   (a) Adding two decoupling capacitances at the junction .
   (b) Adding four decoupling capacitances close to the S-boxes .
   (c) Adding four decoupling capacitances at the junction .

In Experiment 1, the decoupling capacitance is chosen within the range of 0.1–100 pf. This capacitance is placed at different positions in the power distribution tree. Similarly, for Experiment 2, the decoupling capacitance is chosen within the range of 0.1–20 pf. The presence of decoupling capacitance significantly reduces the propagation of feature vectors and loweres its observability at root node $V_{DD}$, which limits the adversary's secret data extraction and modeling abilities.

The feature set vectors of the switching current and the voltage drop values are described in "Side-Channel Perspective". All the experiments are conducted with the same 10,000 inputs, which are randomly generated. For the machine learning classifier training, the feature vectors are constructed with the peak from the voltage, and the current values observed at the $V_{DD}$ pin, labeled with the secret data values, which is the input value of S-box, SB1. The success rate of ML adversary is higher for a power distribution network that propagates the power leakages from S-box to the $V_{DD}$ node without any attenuation; for suppressed leakage through decoupling capacitors or other mechanisms, the ML adversary success rate results in lower values. Base success rates without any attenuation are given in Table 2. These are from the power tree simulations without any decoupling capacitance.

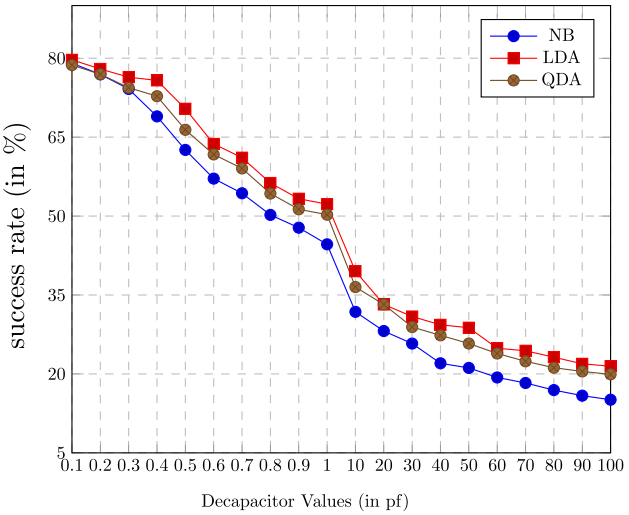


**Fig. 7** Capacitance near S-box

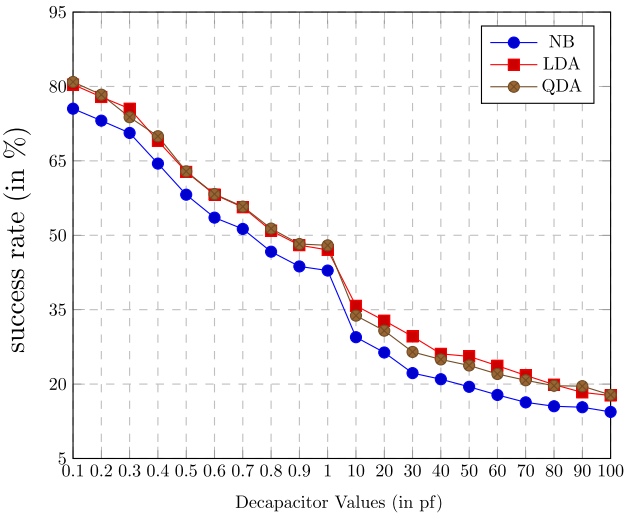


**Fig. 8** Capacitance at the junction

For Experiment (1.a), the decoupling capacitance near the S-box $C_{SB1}$ is raised from 0.1 to 100 pf gradually, and corresponding success rates are recorded. For detailed analysis, the experiment is continued through the depth of the power distribution tree for other S-box functions such as $C_{SB2}$, $C_{SB3}$, and $C_{SB4}$. The success rate of ML adversary is recorded and plotted in Figs. 7 and 8. The success rate for the largest decoupling capacitor values of 100 pf is approximately 21%, which is a 74.07% reduction compared to the success rate values without any decoupling capacitors on the power distribution network. The adversary success rate is reduced by approximately 58.75% with decoupling capacitance range between 0.1 to 40 pf. For decoupling capacitance of 40 pf, the success rates are 22.87% for NB classifier, 28.49% for LDA, and 27.33% for QDA.

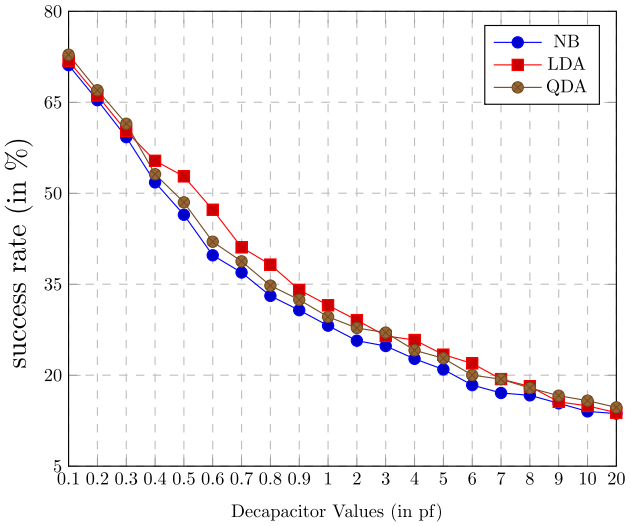


**Fig. 9** Decoupling capacitance at $C_{J3}$ and $C_{J4}$

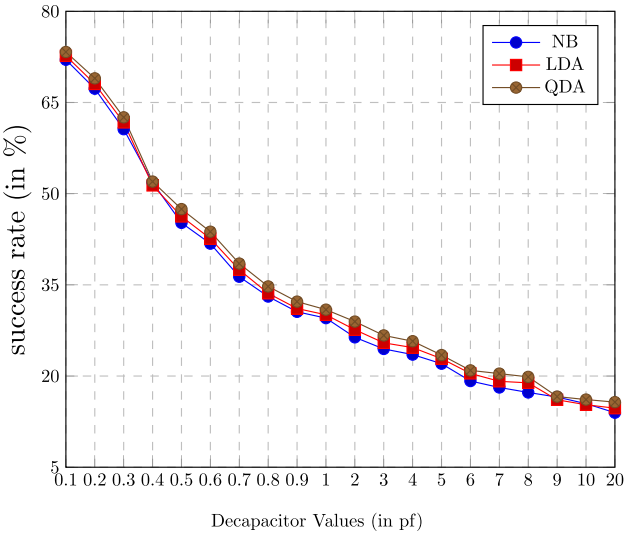


**Fig. 10** Decoupling capacitance at $C_{J5}$ and $C_{J6}$

Every branch in the power distribution tree is connected to other branches or the root node through the junction points. These junction points divide input voltage to the branches based on the current draw of functional blocks at the branches. Hence, it is vital to study the behavior of the decoupling capacitors at the junction. Experiment (1.b) increases the junction capacitance $C_{J3}$, $C_{J4}$, $C_{J5}$, and $C_{J6}$ one at a time. The adversary success rates are recorded, and results are shown in Fig. 8. The success rate value for 100 pf decoupling capacitance at the junction is about 18%, which is a 77.77% reduction compared to the base case power tree with no decoupling capacitance.

Experiment (2) analyzes the power tree with the distributed capacitance in the range 0.1–20 pf. The junction point connected to the S-box is considered to be of primary interest. In Experiment (2.a), the decoupling capacitance of the two junction points near S-boxes is varied. The decoupling capacitance $C_{J3}$ is chosen for all the simulations along with $C_{J4}$, $C_{J5}$, and $C_{J6}$. The results are reported in Figs. 9 and 10. For maximum decoupling capacitance value of 20 pf i.e. $C_{\mathrm{SB1}} = C_{\mathrm{SB2}} = C_{\mathrm{SB3}} = C_{\mathrm{SB4}} = 20$ pf, the success rate of power analysis rootkit is 13.67% for NB classifier, 13.82% for LDA classifier and 14.69% for QDA classifier. Similarly, We also experimented with four decoupling capacitances on the power tree near junctions and S-boxes, and the results are given in Figs. 11 and 12. The success rate of Experiment (2.c) is the same as Experiment (2.b). With the distributed decoupling capacitance, the adversary success rate is reduced by 87.65% compared to the success rate of the power tree without any decoupling capacitor.

The results show that the reduction in the adversary success rate with distributed decoupling capacitance is higher than a single large decoupling capacitance at the power tree.

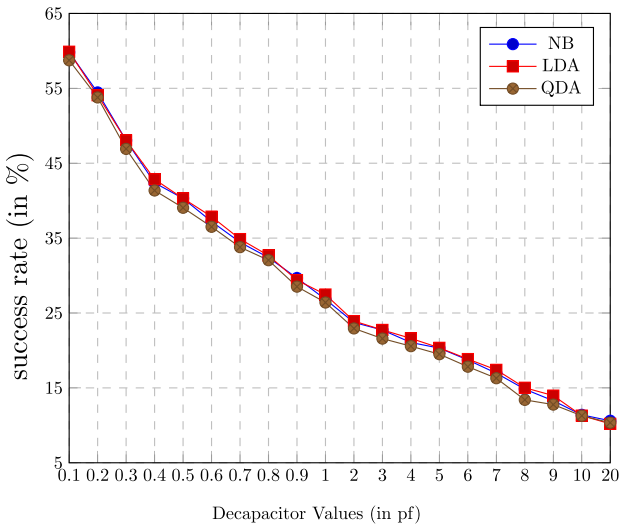


**Fig. 11** Decoupling capacitance distributed near the S-box

The distribution of decoupling capacitance along the power tree momentarily supplies the logic node and minimizes the feature vector visibility at the $V_{\mathrm{DD}}$. Compared to the single capacitor model, the current draw of other logic block junctions is more interleaved with the randomness of the current draw from the multiple capacitors. This causes further confusion for the adversary and a corresponding reduction in the adversarial success rate. An additional confusion for the adversary is created by arbitrarily activating the decoupling unit for a fixed capacitance value on the power path to further randomize the power profile. We evaluate such randomized decoupling schedules as well.

The random activation schedule of multiple capacitors adds to the confusion for the adversary. It increases the

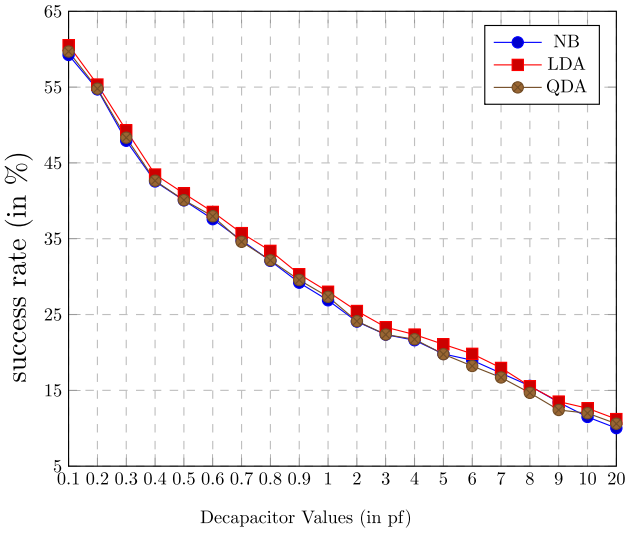


**Fig. 12** Decoupling capacitance distributed on the junctions

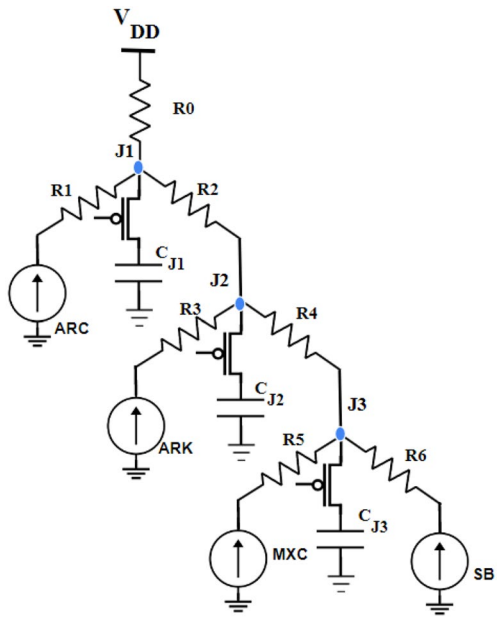


**Fig. 13** Grid-style PDN with pMOS enabled decoupling capacitance

model complexity proportionately. For a random scheduling of the decoupling capacitances, we designed a power tree structure with pMOS transistor enabled decoupling capacitances, as shown in Fig. 13. In this experiment, the junction nodes J1, J2, and J3 have decoupling capacitances and a pMOS transistor activation switch. We randomly activate two out of three decoupling capacitances during the simulation by enabling corresponding switches. The total decoupling capacitance activated on the power path is denoted as $t_{decap}$, and the corresponding ML results are given in Table 3.

**Table 3** Success rate for the tree-style PDN

| $t_{decap}$ | NB (%) | LDA (%) | QDA (%) |
|---|---|---|---|
| 0 pf | 86.17 | 83.95 | 84.72 |
| 1.2 pf | 51.56 | 52.31 | 52.98 |
| 3.6 pf | 38.29 | 37.14 | 38.06 |
| 6.0 pf | 29.73 | 29.56 | 27.32 |
| 12 pf | 20.91 | 21.71 | 20.38 |

Recall that if randomly activate some $k$ out of $m$ capacitors, the additional $\binom{m}{k}$ combinations have to be accounted for in the model by the adversary. This should increase the model complexity by $\left(\frac{em}{k}\right)^k$. If all the combinations of $m$ capacitors are randomly chosen for activation schedule, this complexity goes up by $2^m$. This increased complexity should result in a need for a larger number of samples to build a good statistical model for the adversary. We studied the required sampling needs to build an ML model, which is shown in Fig. 14. The minimum number of samples required to build a stable model in random activation is 6000 samples. The base case for a tree PDN with fixed decoupling capacitance at the junction node. In the power tree, there is only a single path between the sensitive logic block and external $V_{DD}$. In all the random activations, it will still contain 2* $t_{decap}$ decoupling capacitance. Hence, the random activation of decoupling capacitance does not impact the sample count requirement or success rate in the tree-style PDN.

## Analysis on Grid-style PDN

In grid-style PDN, the logic blocks are arranged in a grid with each PDN grid junction supplying $V_{DD}$ and ground. The grid structure is considered robust and has lower noise compared to the power tree. To study the side-channel resistance for this style, we constructed a $4 \times 4$ power grid with its path resistance of 1 Ω as shown in Fig. 15. The functional blocks such as AddRoundKey (ARK), AddConstant (ARC), MixColumn (MXC), and S-box (SB) are connected to the junction nodes J6, J7, J10, and J11, respectively. Similar to Fig. 13, the blue-colored junction nodes in Fig. 15 contains a pMOS-enabled decoupling capacitance.

The primary objective of our proposed approach is to cut off the external power supply to the logic node, i.e., place the decoupling capacitors at the junctions of a cut. In grid-structured PDN, there are multiple power paths between the target logic node J11 and external $V_{DD}$. All the power paths need to have decoupling capacitances in order for it to be effective. Even a single path without a decoupling capacitance transmits the logic block J11 power wave form unaltered to the $V_{DD}$ pin. The proposed countermeasure must

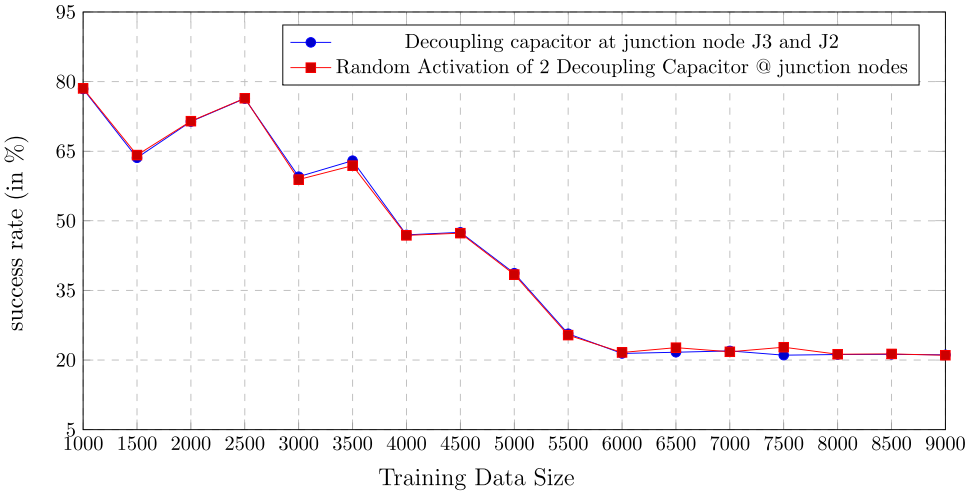


**Fig. 14** Model evaluation of tree-style PDN

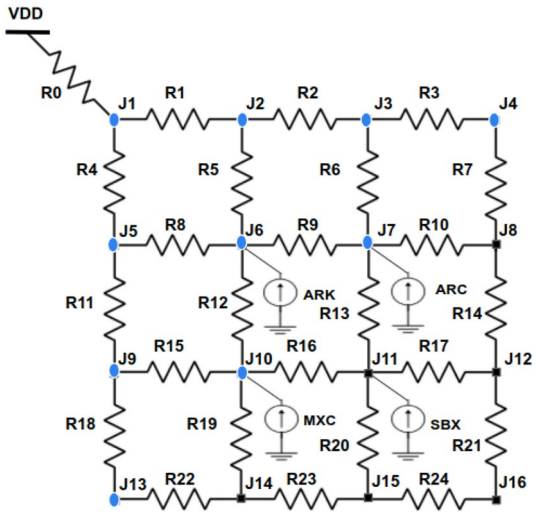


**Fig. 15** Grid-style PDN with pMOS enabled decoupling capacitances

identify all the junction nodes involved in the cuts of the power paths between the target logic node J11 and $V_{DD}$, which are captured as tuples. The tuples are created based on the minimum number of junction nodes to decouple all possible power paths between the logic node and $V_{DD}$. The tuples are represented as $T_i$, where $'i'$ denotes the node number of the junction node.

The junction nodes {J2, J5} form a cut, and are key junction nodes to distribute the power across the grid. Adding decoupling capacitances at junction nodes J2 and J5 decouples the logic node from the external $V_{DD}$ to degrade the feature vector propagation in the power path. Hence, we create a tuple $T_2$ with junction nodes J2 and J5. Similarly, the tuples $T_3$ and $T_4$ contain the junction nodes {J3, J6, J9} and {J4, J7, J10, J13} respectively, which also form cuts of the target logic node and $V_{DD}$. Similar to the previous experiments, We randomly generated 10,000 inputs and recorded the corresponding current and voltage drop at $V_{DD}$ node. The ML results for the grid-style analysis are reported in Table 4 and Table 5.

**Table 4** Success rate for the grid-style PDN without decoupling capacitance

| Classifier | NB | LDA | QDA |
|---|---|---|---|
| Success rate | 68.96% | 67.44% | 67.51% |

Comparing Tables 3 and 5 success rates, it is clear that the grid-style PDN shows better side-channel resistance compared to base tree-style PDN without any decoupling capacitance. With 12 pf decoupling capacitance, the cut tuple $T_4$ has a success rate of 16.04% for NB, 17.43% for LDA, and 15.34% for QDA, which is slightly better compared to other cut tuples in the grid-style PDN. This may indicate, similar to the tree PDN, that a cut closer to the target node, J11 in this case, yields the best side-channel resistance.

We also performed the random activation experiments for 12 pf decoupling capacitance in the grid-style PDN similar to tree-style PDN analysis, in order to assess increased sampling needs. In this experiment, a 2 pf decoupling capacitance is placed in all the junction nodes with the pMOS transistor. We randomly enable six decoupling capacitances for every sample run and record the feature vectors. We computed the minimum sample requirements to build a stable ML model. The results are shown in Fig. 16. For every cut tuple configuration in the grid style PDN, the minimum sample count required to build the ML model is 5500. With

**Table 5** Success rate for the grid-style PDN with decoupling capacitance

| $t_{decap}$ | Success rate in (%) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | NB | | | LDA | | | QDA | | |
| | $T_2$ | $T_3$ | $T_4$ | $T_2$ | $T_3$ | $T_4$ | $T_2$ | $T_3$ | $T_4$ |
| 1.2 pf | 51.57 | 50.81 | 48.73 | 51.41 | 49.05 | 51.06 | 51.29 | 48.96 | 51.13 |
| 3.6 pf | 40.31 | 40.08 | 39.67 | 39.63 | 39.62 | 40.67 | 39.32 | 41.11 | 40.85 |
| 6.0 pf | 20.18 | 21.07 | 19.25 | 21.72 | 20.98 | 21.38 | 21.31 | 21.93 | 21.46 |
| 12 pf | 17.26 | 16.61 | 16.04 | 18.97 | 18.74 | 17.43 | 16.49 | 15.75 | 15.34 |

**Table 6** Summary of area overhead

| Logic | Gate | Cell | Area ($\mu m^2$) |
|---|---|---|---|
| S-box | 488 | 592 | 1374.10 |
| MixColumn | 389 | 284 | 1096.30 |
| AddRoundKey + AddConstant | 192 | 128 | 540.60 |
| Total (w/o decoupling capacitance) | 1069 | 1004 | 3011.00 |
| Decoupling circuit (2 pf) | – | – | 118.12 |

random activation, the minimum count of samples required to build the model is 8000. Ideally, the sampling needs for random activation should go up by $\binom{12}{2} = 924$. But there are many other factors such as the topology that determine the model dimensionality increase. The minimum sample count required to build the ML model in larger power grids could be larger due to increased number of cuts leading to higher randomization potential.

The decoupling capacitors cost additional area. The area overhead is given in Table 6. The decoupling capacitance occupies 118.12 $\mu m^2$ of silicon area for 2 pf. In the random activation setting, we implement 2 pf decoupling capacitance in ten different junction nodes across the PDN grid, which sums up to 1181.2 $\mu m^2$ in total. In this scenario, the overall area overhead is 39.33% compared to the base LED round functions area.

## Conclusions

In this paper, we present a detailed analysis of the power distribution network with decoupling capacitance as a countermeasure to mitigate the power analysis attacks. Our approach is to supply a portion of the switching current for sensitive circuits as locally as possible to effectively decouple the circuit for a short time span. We illustrate the propagation of the feature set vectors on the power distribution network with and without decoupling capacitance. The discharge cycle of decoupling capacitance suppresses the peak switching activity to constrain the feature vectors from propagating to the $V_{DD}$ pin. We have discussed the effect of parasitic resistance and capacitance on feature vector propagation and temporal alignment.

The design space for the decoupling capacitors is further explored by (1) placing multiple decoupling capacitors at various PDN points, and (2) exploring a random activation schedule of distributed decoupling capacitors. We consider the potential placement of decoupling capacitances on two different styles of the power distribution network through various experiments. We have demonstrated the side-channel resistance in terms of success rate metrics using machine learning classifiers such as QDA, LDA, and NB. The results show that the grid-style PDN has higher side-channel resistance than the tree-style PDN with decoupling capacitors.

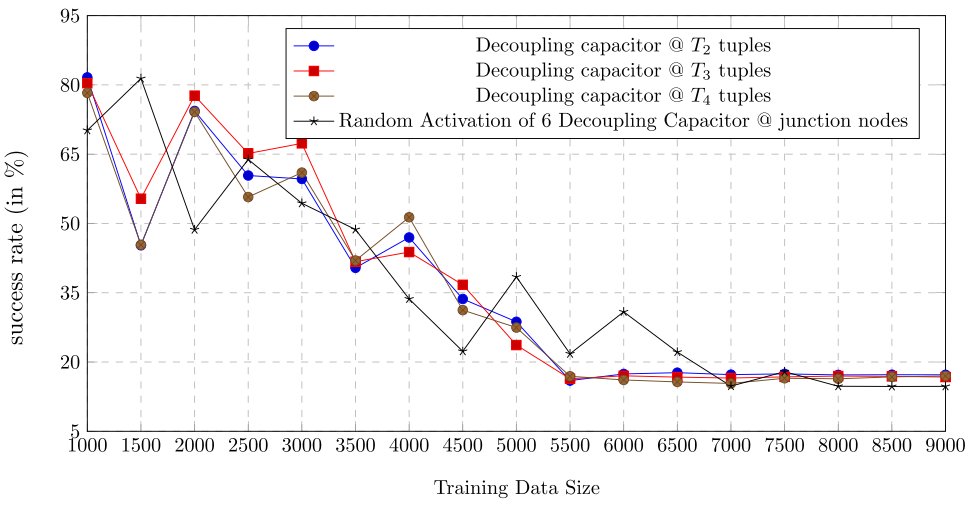


**Fig. 16** Model evaluation of grid-style PDN

The random activation approach shows that the minimum sample count required for building an ML model is increased by 45% compared to fixed decoupling capacitor activation at a cut tuple in the grid-style PDN. Unlike logic-style countermeasures, the area overhead of our proposed approach is very minimal and independent of the functionality embedded in logic cells. The proposed technique can be applied to any circuit during the physical design stage through CAD tools.

## Declarations

**Conflict of interest** Both authors declare no conflict of interest.

**Code availability** Not applicable.

**Author contributions** RS and AT conceived, designed the experiments and wrote the paper; RS carried out the experiments and analyzed the data. Both authors have read and agreed to the published version of the manuscript.

## References


1. Kocher PC. Timing attacks on implementations of Diffie–Hellman, RSA, DSS, and other systems. In: Advances in cryptology—CRYPTO '96, 16th annual international cryptology conference, Santa Barbara, California, USA, August 18–22, 1996, Proceedings, 1996. p. 104–13.
2. Chari S, Rao JR, Rohatgi P. Template attacks. In: Cryptographic hardware and embedded systems—CHES 2002, 4th international workshop, Redwood Shores, CA, USA, August 13–15, 2002, revised papers; 2002. p. 13–28.
3. Moradi A, Mischke O, Eisenbarth T. Correlation-enhanced power analysis collision attack. In: Cryptographic hardware and embedded systems, CHES 2010, 12th international workshop, Santa Barbara, CA, USA, August 17–20, 2010. Proceedings; 2010. p. 125–39.
4. Coron JS, Kizhvatov I. An efficient method for random delay generation in embedded software. In: Cryptographic hardware and embedded systems—CHES 2009, 11th international workshop, Lausanne, Switzerland, September 6–9, 2009, Proceedings; 2009. p. 156–70.
5. Benini L, Galati A, Macii A, Macii E, Poncino M. Energy-efficient data scrambling on memory-processor interfaces. In: Proceedings of the 2003 international symposium on low power electronics and design, 2003, Seoul, Korea, August 25–27, 2003; 2003. p. 26–9.
6. Ishai Y, Sahai A, Wagner DA. Private circuits: Securing hardware against probing attacks. In: Advances in cryptology—CRYPTO 2003, 23rd annual international cryptology conference, Santa Barbara, California, USA, August 17–21, 2003; 2003. p. 463–81.
7. Nikova S, Rechberger C, Rijmen V. Threshold implementations against side-channel attacks and glitches. In: Information and communications security, 8th international conference, ICICS 2006, Raleigh, NC, USA, December 4–7, 2006, Proceedings; 2006. p. 529–45.
8. Selvam R, Tyagi A. Power side-channel resistance of RNS secure logic. In: 31st international conference on VLSI design and 17th international conference on embedded systems, VLSID 2018, Pune, India, January 6–10, 2018; 2018. p. 143–8.
9. Bucci M, Giancane L, Luzzi R, Trifiletti A. Three-phase dual-rail pre-charge logic. In: Cryptographic hardware and embedded systems—CHES 2006, 8th international workshop, Yokohama, Japan, October 10–13, 2006, Proceedings; 2006. p. 232–41.
10. Tiri K, Verbauwhede I. A logic level design methodology for a secure DPA resistant ASIC or FPGA implementation. In: 2004 design, automation and test in Europe conference and exposition (DATE 2004), 16–20 February 2004, Paris, France; 2004. p. 246–51.
11. Tiri K, Verbauwhede I. Design method for constant power consumption of differential logic circuits. CoRR, abs/0710.4756; 2007.
12. Selvam R, Tyagi A. Power distribution network capacitive decoupling for side-channel resistance. In: 2021 IEEE international symposium on smart electronic systems (iSES); 2021. p. 183–8.
13. Tang KT, Friedman EG. Transient IR voltage drops in CMOS-based power distribution networks. In: Proceedings of the 43rd IEEE Midwest symposium on circuits and systems (Cat. No.CH37144), vol. 3; 2000. p. 1396–9.
14. Mao J, Kim W, Choi S, Swaminathan M, Libous J, O'connor D. Electromagnetic modelling of switching noise in on-chip power distribution networks. In: 8th international conference on electromagnetic interference and compatibility; 2003. p. 47–52.
15. Tanaka H, Matsushima T, Yano Y, Wada O. Compensating method of equivalent current sources of LSI-core macromodel considering voltage fluctuations in on-chip power distribution network. IEEE Trans Electromagn Compat. 2022;64(4):1250–6.
16. Tang KT, Friedman EG. Simultaneous switching noise in on-chip CMOS power distribution networks. IEEE Trans Very Large Scale Integr (VLSI) Syst. 2002;10(4):487–93.
17. Joo J, Sun Y, Lee J, Kong S, Kang S, Song I, Hwang C. Modeling of power supply noise associated with package parasitics in an on-chip ldo regulator. In: 2021 IEEE international joint EMC/SI/PI and EMC Europe symposium; 2021. p. 395–9.
18. Mayhew M, Muresan R. Modeling the effect of nmos gate capacitance in an on-chip decoupling capacitor paa countermeasure. In: 2014 IEEE 57th international Midwest symposium on circuits and systems (MWSCAS); 2014. p. 121–4.
19. Mayhew M, Muresan R. On-chip nanoscale capacitor decoupling architectures for hardware security. IEEE Trans Emerg Top Comput. 2014;2(1):4–15.
20. Dofe J, Yu Q. Exploiting PDN noise to thwart correlation power analysis attacks in 3d ics. In: Proceedings of the 20th system level interconnect prediction workshop, SLIP@DAC 2018, San Francisco, CA, USA, June 23, 2018; 2018. p. 6:1–6.
21. Kenarangi F, Partin-Vaisband I. Exploiting machine learning against on-chip power analysis attacks: tradeoffs and design considerations. IEEE Trans Circuits Syst I Regul Pap. 2019;66(2):769–81.
22. Muresan R. On-chip CMOS self-decoupling battery cell system for security protection. Can J Electr Comput Eng. 2020;43(2):83–91.