Peak Power and Dynamic IR-drop Assessment via Waveform Augmenting

Yihan Wen
Beijing University of Technology
yhwen@emails.bjut.edu.cn

Juan Li
Beijing University of Technology
lijuan@bjut.edu.cn

Bei Yu
The Chinese University of Hong Kong
byu@cse.cuhk.edu.hk

Xiaoyi Wang
Unaffiliated Scholar
wxy001477@qq.com

ABSTRACT

Pre-silicon power and IR-drop estimation are crucial parts of the chip design process. Vector-based and vectorless assessments are commonly employed to estimate the worst peak power and dynamic IR-drop of the design. However, with rapid growth in chip scale and complexity, the waveform-driven vector-based assessment encounters the coverage challenge due to the difficulty in generating test waveforms that encompass all potential worst-case scenarios. Additionally, Vectorless assessments consistently yield overly pessimistic estimations which may lead to significant overdesign. This paper proposes a semi-vector-based assessment flow aimed at offering a more reasonable estimation of worst-case peak power and IR-drop. In the proposed assessment, functionally independent modules are identified through the analysis of module toggle activity correlation (MTAC) on the existing waveform. By making these functionally independent modules toggle simultaneously, an augmented waveform approximating the worst peak power and dynamic IR-drop scenario is actively generated while preserving a similar MTAC to the existing waveform. By applying dynamic power and IR-drop analysis to the augmented waveform, previously unaddressed weaknesses are identified. Experimental results on an industrial design indicate that the proposed worst-case assessment result is 3× and 2.5× more accurate than the vector-based result for worst peak power and dynamic IR-drop. Similarly, it is 33× and 6× more accurate than the vectorless result.

1 INTRODUCTION

As chip performance increases and process nodes shrink, the challenges of power consumption and power integrity in the design process of very large-scale integrated circuits (VLSI) have become increasingly prominent. The high power density resulting from peak power dissipation poses challenges for both packaging and cooling solutions, as well as for system reliability [1]. In addition, larger power consumption on instances imposes excessive transient current on the power delivery network, which causes severe IR-drop. The drop not only introduces timing uncertainty but also slows

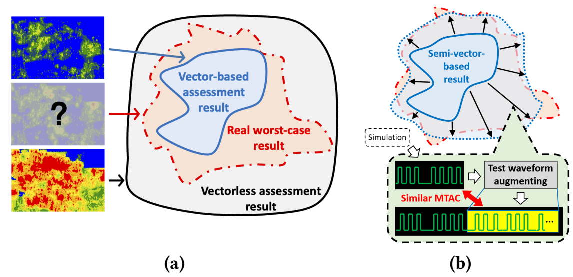


Figure 1: The coverage weakness of the traditional vector-based and vectorless assessment (a) and the improvement of proposed semi-vector-based assessment by the waveform augmenting (b).

down the cell slew rate, ultimately degrading the overall performance of the chip [2]. Hence, accurate estimation of the worst-case peak power and dynamic IR-drop during the design process becomes essential [3].

However, the traditional vector-based and vectorless power and IR-drop assessments face challenges in covering the true worst-case scenarios in the complex and large-scale modern design, as demonstrated in Fig. 1(a). In industrial design flows, vector-based assessment is employed to address the power and IR-drop of certain real test programs. Dynamic transient analyses are conducted on waveform files generated from gate-level simulations [4]. However, the analysis is constrained by the runtime and resource limitations, restricting the length of waveform files that can be solved. In this case, anticipating the waveform segment that activates the true worst-case power and IR-drop scenario is challenging [5]. Consequently, vector-based estimates only reflect the lower bound of the actual worst-case peak power and dynamic IR-drop. Unexpected power and IR-drop weaknesses that are not addressed by the waveform can significantly impact the reliability and performance of chip products. On the other hand, vectorless assessment [6–11] estimate the worst-case without relying on an accurate activity waveform. However, the results of vectorless assessments are heavily influenced by user-defined activity settings, such as the toggle rate on the primary inputs. In vectorless analysis, the switching activity of instances is determined by switching probability propagation on the netlist, which fails to accurately reflect real working scenarios and consistently yields overly pessimistic estimations. As a result, a large gap exists between the vector-based and the vectorless assessments, leading to estimations of the true worst-case that are either overly optimistic or overly pessimistic.

In this paper, a semi-vector-based assessment is proposed to estimate the worst-case power and dynamic IR-drop. As illustrated in Fig. 1(b), different from the previous works [12–14] which find

This work was supported in part by grants from the Research Grants Council of Hong Kong SAR (No. CUHK14210723 and No. CUHK14211824).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICCAD '24, October 27–31, 2024, New York, NY, USA

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-1077-3/24/10…$15.00
https://doi.org/10.1145/3676536.3676767

the worst-case input vector focusing on the netlist structure, the statistical information of the existing waveform is investigated in the proposed approach. First, the activity dependence relationships of modules in the waveform are quantified by module toggle activity correlation (MTAC) analysis. Then, functionally independent modules are identified by a graph clustering algorithm based on the distance measured by the MTAC. Finally, the existing waveform is actively augmented to approximate the worst power and IR-drop scenario while keeping a similar MTAC. The main contributions of the proposed semi-vector-based assessment method are as follows:

- The proposed assessment focuses on the toggle activity within the existing waveforms, eliminating the need for a netlist structure. This not only improves the efficiency of large-scale designs but also makes it applicable to various design types.
- The proposed assessment identifies the functionally independent modules resulting from both the netlist structure and the system-level scheduling of working scenarios within the existing waveform. This aids designers in enhancing test cases to more effectively cover real worst-case scenarios.
- The proposed assessment approximates the worst-case scenario by aligning the highest toggle waveform segment of functionally independent modules. This augmenting strategy maintains the waveform within each module unchanged, thus preserving their correspondence to real working scenarios, which alleviates unreasonable pessimistic estimations.

Experiments conducted on a paralleled hash processor verified that the proposed method accurately identified the functionally independent modules. Additionally, the proposed method was applied to an industrial multi-core ARM CPU design to validate the improvement in worst-case estimation. On average, the proposed method estimated the peak power with a relative error of 2.93% compared to the golden value and estimated the worst dynamic IR-drop with a relative error of 7.62%, which is 3× and 2.5× smaller than the vector-based result respectively. Similarly, the relative error of the proposed result is 33× and 6× smaller than the vectorless result. This indicates that the proposed method avoids an overly pessimistic estimation of the worst-case.

The rest of this paper is organized as follows: Section 2 reviews the previous work on worst-case peak power and dynamic IR-drop estimation. Section 3 discusses the proposed semi-vector-based assessment. Section 4 gives out the experimental. Section 5 concludes this paper.

2 PRELIMINARIES

The instance power dissipation of CMOS circuits is calculated in Eq. (1) by the summation of internal power P_{internal}, switching power $P_{\text{switching}}$, and leakage power P_{leakage} where C_L is the switching capacitance, f is the frequency of clock, V_{DD} is the voltage of power supply, t_v is the voltage settling time, I_{short} is the short current and α is the switching activity [15].

$$
\begin{aligned}
P_{\text{t}} &= P_{\text{switching}} + P_{\text{internal}} + P_{\text{leakage}} \\
P_{\text{switching}} &= \alpha C_L V_{\text{DD}}^2 f \\
P_{\text{internal}} &= \alpha t_v V_{\text{DD}} I_{\text{short}} \\
P_{\text{leakage}} &= V I_{\text{leakage}}
\end{aligned}
\tag{1}
$$

Once the instance power is solved, the instance current profile is determined, and the dynamic IR-drop is described by a linear ordinary differential equation (ODE) [3]. To obtain accurate estimates of dynamic power and IR-drop, the toggle activity α is crucial as both P_{internal} and $P_{\text{switching}}$ are proportional to α. Depending on how toggle activity is determined, assessment methods can be classified into two categories including vector-based and vectorless assessments.

In the vector-based assessments [4], gate-level simulations are utilized to find which instants can switched and to determine when the switching occurs based on some real test vector. The dynamic power and IR-drop analysis is driven by waveform dumped in the voltage change dump (VCD) format, which records the changes in signal values over time during the simulation. The vector-based assessment provides a ground truth for peak power and IR-drop under the preset working scenarios of the waveform. However, the vector-based assessment encounters challenges for the following reasons. Firstly, the waveform may not be accessible in the early stages of the design flow [16]. Secondly, simulating the design with all possible inputs is inefficient due to the required time and limited machine resources [15]. Last but not least, since it is difficult to prove that the waveform indeed covers the worst-case, the vector-based assessment is hard to reflect the worst-case dynamic power and IR-drop [5].

In contrast, vectorless assessment does not rely on accurate switching activity information for instances. The current constraints-based approaches [17–21] initiated in [17] use current constraints to restrict the range of instances current, then solve a linear equation to evaluate the worst-case static IR-drop. [18, 21] extend the vectorless verification with transient current constraints. [19, 20] improve the efficiency of vectorless analysis to get a scalable performance on modern large designs. However, these methods ignore the exact circuitry under the grid [22], which leaves a challenge on how to set appropriate current constraints that meet the real working scenarios. Furthermore, knowing the power consumption in advance is necessary to profile the current constraints, which is also challenging. On the other hand, the statistical and probabilistic-based methods [6–11] consider what instances switched and how often they switched in a statistical perspective, without relying on actual waveform-based simulations. The transition density first proposed in [6], also equivalent to the toggle rate, represents the switching probabilities of the signal in the netlist. The toggle rate of the netlist input is specified and then propagated throughout the entire netlist using the statistical method, such as the sample-based method [11], the stochastic sequential machine [8], the Bayesian network [9], and the correlation-based methods [7, 10]. Once the switching probability of all instances is determined, a virtual switching vector is generated. The power consumption is then computed and the dynamic current profile for the IR-drop analysis is also generated.

However, probabilistic-based methods still require manual specification of the toggle rate on primary inputs of the netlist, leading to the challenge of input dependence [23]. Essentially, the average power of the chip is directly determined by the toggle rate specified on the primary inputs [24]. The higher the toggle rate is set, the greater the power applied and the larger the area of IR-drop hotspots generated. A commonly used trick is tuning the toggle rate to match the total switching percentage to a known scenario in some real waveform. However, real waveforms typically only

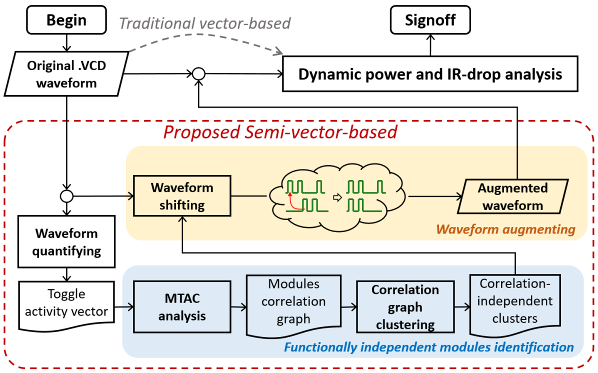


Figure 2: Semi-vector-based peak power and dynamic IR-drop assessment flow.

activate logic in a small region, while vectorless assessment propagates toggles throughout the entire netlist, which always leads to disagreement [25]. Tuning the toggle rate to match the waveform in a local region often leads to significant pessimistic estimations. Therefore, in industrial design flows, vectorless assessment is commonly used to estimate the early power and IR-drop. Once the real waveform is simulated, power and IR-drop signoff primarily rely on the vector-based results.

To address the coverage problem of vector-based analysis mentioned above, one effective approach is to identify the functionally independent parts of the circuit and allow them to switch at the same time to capture the worst-case scenario. The maximum instantaneous current (MIC) estimation [5] identifies the mutually exclusive gates that may or may not toggle at the same time by the logic structure of the netlist. [26] transfers this question to the boolean satisfiability (SAT) problem and [27] considers the variance in the number of switching events. [12–14] intend to find the input vector that can activate the worst-case. However, these methods focus on the gate-level maximum instantaneous current in a single clock cycle, which ignores the implicit constraints from system behaviors in complex designs. Although some functional paths are mutually exclusive in the netlist structure, they will never toggle at the same time due to the system scheduling. Considering those instances working at the same time can lead to pessimistic estimations. Due to the design complexity, these implicit constraints are hard to foresee based on the netlist logical structure, especially at the instance or in-depth module level. Fortunately, these constraints are always reflected in the existing waveform since it is simulated from a real test program. Essentially, the toggle activity of modules in waveform can be viewed as a time series. Therefore, statistical methods, such as correlation analysis method [28] can be utilized to model the inherent dependence characteristic in the waveform. The recent research [29] uses the correlation coefficient to identify the reliability of high-correlated gates. However, how to analyze the waveform to identify the potential worst power and IR-drop scenario is still an open question.

3 SEMI-VECTOR-BASED ASSESSMENT VIA MODULE TOGGLE ACTIVITY CORRELATION (MTAC) BASED WAVEFORM AUGMENTING

The flowchart of the proposed semi-vector-based peak power and IR-drop assessment is shown in Fig. 2. Different from the traditional vector-based assessment which applies power and IR-drop analysis only to the given waveform, the proposed semi-vector-based assessment augments the waveform to approximate the potential worst-case. Firstly, the MTAC of the original waveform is analyzed and the functionally independent modules are identified in section 3.1. Then, the waveform is augmented in section 3.2 while maintaining a similar MTAC. Finally, the dynamic analysis is applied to both the original waveform and the augmented waveform.

3.1 Functionally independent modules identification through MTAC analysis

As illustrated in the blue block in Fig. 2, the functionally independent modules are identified in the following steps. First, the waveform is quantified into modules toggle activity vectors in both the time domain and the hierarchical domain of instances. Then, correlation coefficients are calculated between different modules. Finally, the modules are divided into functionally independent clusters based on their MTAC distance.

3.1.1 *Quantifying waveform to module toggle vectors.*

The number of instances in a modern design can be vast. Analyzing waveform toggle activity at the instance level would result in unacceptable resource and runtime costs. Fortunately, instances are typically organized in a hierarchical tree structure at different levels of abstraction to break down complex systems. Therefore, activity analysis could be conducted on the hierarchical instances (H-insts) rather than the flattened instances to reduce resource and runtime consumption. As shown in Fig. 3, the modules are specified as the deepest H-insts that include a total number of instances larger than the given granularity G_m.

On the other hand, toggle events of each instance may occur on every clock cycle. A typical waveform in the industrial design flow often spans over tens of thousands of clock cycles, which is quantitatively large. Therefore, toggle events are also quantified into time slots, resulting in the module toggle activity vector $\mathbf{T}$ in Eq. (2). T_i^t denotes the number of toggle events on all instances in module i in the t-th time slot corresponding to the time period $(t-1)\cdot\tau \sim t\cdot\tau$, where τ denotes the period length of the slot.

$$\mathbf{T}_i = \begin{bmatrix} T_i^1 & T_i^2 & \cdots & T_i^t & \cdots & T_i^{\mathcal{T}} \end{bmatrix} \tag{2}$$

3.1.2 *Calculating MTAC coefficient.*

To measure the dependence relationship of toggle activity between modules, the Pearson Product-Moment Correlation Coefficient [28] (PPMCC) c_{ij} is used to represent the MTAC coefficient between two modules i and j in Eq. (3). T_i and T_j denote the toggle

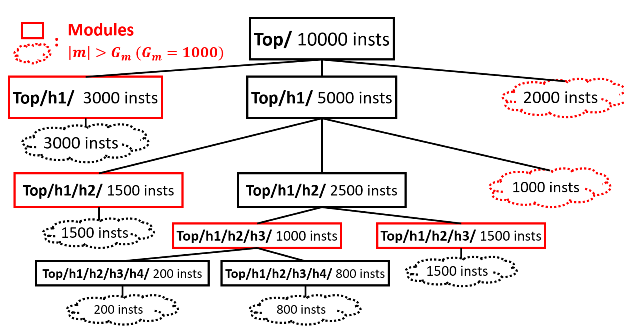


Figure 3: Specifying modules from the instances hierarchical structure.

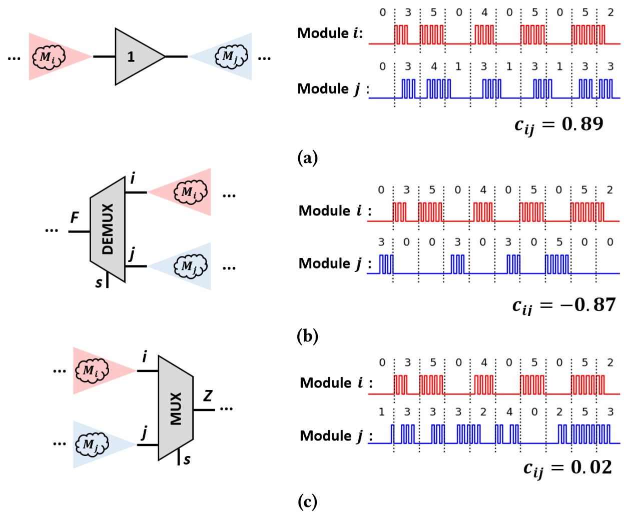


Figure 4: Example of (a) positive-correlated modules, (b) negative-correlated modules, and (c) correlation-independent modules.

vectors of two modules; cov denotes the covariance between two vectors and σ denotes the standard deviation of vector.

$$c_{ij} = \frac{\mathrm{cov}(\mathbf{T}_i, \mathbf{T}_j)}{\sigma_{\mathbf{T}_i}\sigma_{\mathbf{T}_j}} \tag{3}$$

c_{ij} measures the linear correlation of module toggle activity over time and takes values in the range $[-1, 1]$. Given a threshold c_ϵ, the modules toggle activity exhibits three typical correlation cases.

For the case $c_{ij} \geq c_\epsilon$, modules i and j exhibit positive correlation, indicating that the two modules have consistently worked together. An example is provided in Fig. 4(a), where modules i and j are located on the same path. The toggle activity propagates from module i to j, resulting in both modules exhibiting a similar tendency of toggle activity variation, demonstrating a strong positive correlation with MTAC coefficient $c = 0.89$. For the case $c_{ij} \leq -c_\epsilon$, modules manifest a negative correlation, indicating that the two modules have not worked simultaneously. An example is provided in Fig. 4(b), where modules i and j are located on two output paths of a demultiplexer. The selection of two paths are mutually exclusive that if one path is hit, the other will be idle. In this case, the toggle activity vector of a and b exhibit a strong negative correlation with $c = -0.87$. For the case $-c_\epsilon < c_{ij} < c_\epsilon$, modules exhibit little correlation, indicating that the toggle activity of two modules is independent. An example is provided in Fig. 4(c), where modules i and j are located on two input paths of a multiplexer. Since the two modules are on different functional paths, the toggle activity vectors exhibit independence with $c = 0.02$.

Considering all modules in the entire design, the overall MTAC is denoted by a complete graph $G\{V, E\} = K_n$, where n equals the number of modules. Vertices in V correspond to each module, and the weights of edges in E represent the MTAC coefficient between pairs of modules. Since the MTAC coefficients are calculated between any two modules, G is fully connected. Thus, the MTAC graph adjacency matrix $\mathcal{A}_G$, also known as the modules correlation matrix, is sufficient to represent all the correlations among modules where $\mathcal{A}_G[i,j]_{i=j} = 1$ on the leading diagonal and $\mathcal{A}_G[i,j]_{i\neq j} = c_{ij} = c_{ji}$ representing the MTAC coefficient between module i and j.

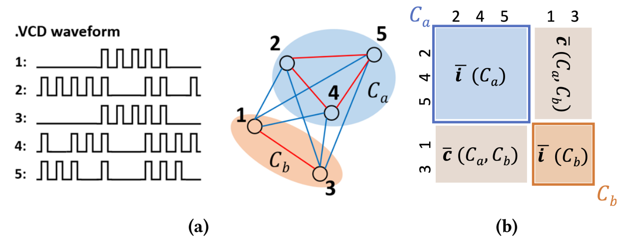


Figure 5: The correlation-independent clusters (a) on the MTAC graph and the clusters correlation matrix (b).

3.1.3 *Identifying functionally independent module clusters.*

After calculating the MTAC, the functionally independent modules are identified through the correlation-independent module clusters, as shown in Fig. 5(a). The clusters are represented by the disjoint set $\boldsymbol{C}$ which includes N cluster sets $\{C_1, \cdots, C_N\}$. Each cluster set includes modules $\{m_1, m_2, \cdots\}$. Eq. (4) defines the clusters correlation weight $\bar{\mathbf{c}}$ by averaging all edges c_{ij} connecting modules in two different clusters C_a and C_b on the MTAC graph, which measures the overall correlation between two clusters. $\bar{\mathbf{c}}$ is defined similarly to the graph *cut* in [30], but solving the average of all edges rather than the summation to eliminate the effect of the number of modules in clusters.

$$\bar{\mathbf{c}}(C_a, C_b) = \frac{1}{|C_a||C_b|} \sum_{i \in C_a, j \in C_b} |c_{ij}| \tag{4}$$

$$\bar{\mathbf{i}}(C_a) = \frac{2}{|C_a|^2 - |C_a|} \sum_{p,q \in C_a, p \neq q} |c_{pq}| \tag{5}$$

Eq. (5) defines the internal correlation weight $\bar{\mathbf{i}}$ of one cluster by averaging all edges inside the cluster, which measures the MTAC compactness of the cluster.

The correlation-independent clusters set $\mathbf{C}$ satisfies the constraints in Eq.(6): $\bar{\mathbf{c}}$ between any two clusters should be no greater than the independent threshold c_ϵ (6a) and $\bar{\mathbf{i}}$ of any cluster should be no less than c_ϵ (6b). This constraint ensures the MTAC between different clusters should be small compared to the internal MTAC inside one cluster, which reflects the clusters independence.

$$\bar{\mathbf{c}}(C_a, C_b) \leq c_\epsilon \tag{6a}$$

$$\bar{\mathbf{i}}(C_a) \geq c_\epsilon, \bar{\mathbf{i}}(C_b) \geq c_\epsilon \tag{6b}$$

$$\forall C_a \in \mathbf{C}, \forall C_b \in \mathbf{C}, a \neq b$$

Finding $\mathbf{C}$ is well studied in graphic spectral clustering [31]. The Shi–Malik (S-M) algorithm [32] which is widely used in community detection in social networks is an effective method. The S-M algorithm computes the eigenvalues and eigenvectors of the graph Laplacian matrix and selects the eigenvector with the second-smallest eigenvalue to bi-partition the graph where the normalized cut between two clusters is minimized.

Applying the S-M algorithm to the MTAC graph, the identification of correlation-independent clusters is summarized in algorithm 1. Different from the split-only process in [32], the proposed clustering algorithm employs a two-stage 'V' structure flow. Firstly, it recursively breaks the modules into disjoint clusters. Then, it

Algorithm 1: Identify the correlation-independent module clusters on the MTAC graph.

input : Adjacent matrix of the MTAC graph $\mathcal{A}_G$; Independence threshold c_ϵ

output: Correlation-independent module clusters disjoint set $\boldsymbol{C} = \{C_1, \cdots, C_N\}$ includes N clusters, which satisfied Eq. (6)

SPLIT ($\boldsymbol{C}, N, \mathcal{A}_{Gi}, C_i$):

1* *Split* C_i *to* C_a *and* C_b *by the S–M algorithm in [32];*
2* *Calculate* $\bar{\mathbf{c}}(C_a, C_b)$ *by Eq. (4);*
3* **if** $\bar{\mathbf{c}}(C_a, C_b) < c_\epsilon$ **then**
4*   *Update* $\boldsymbol{C}$ *by replacing* C_i *to* C_a *and* C_b*;*
5*   *Update* $N = N + 1$*;*
6*   **if** $|C_a| \geq 2$ **then**
7* $\mathcal{A}_{Ga} = \mathcal{A}_{Gi}[\forall M_a \in C_a, \forall M_a \in C_a]$;
8*   SPLIT ($\boldsymbol{C}, N, \mathcal{A}_{Ga}, C_a$)
9*   **end**
10*   **if** $|C_b| \geq 2$ **then**
11* $\mathcal{A}_{Gb} = \mathcal{A}_{Gi}[\forall M_b \in C_b, \forall M_b \in C_b]$;
12*   SPLIT ($\boldsymbol{C}, N, \mathcal{A}_{Gb}, C_b$)
13*   **end**
14* **end**

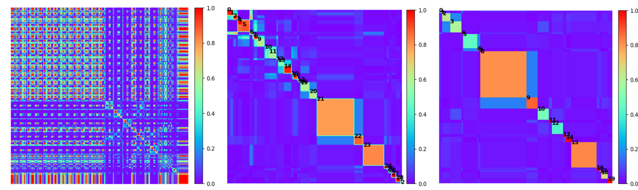
MERGE ($\boldsymbol{C}, N$):

1· **for** C_a *in* $\boldsymbol{C}$ **do**
2·   **for** C_b *in* $\boldsymbol{C} - \{C_a\}$ **do**
3·   *Calculate* $\bar{\mathbf{c}}(C_a, C_b)$ *by Eq. (4);*
4·   **if** $\bar{\mathbf{c}}(C_a, C_b) \geq c_\epsilon$ **then**
5·   *Update* $\boldsymbol{C}$ *by merging* C_b *into* C_a*;*
6·   *Update* $N = N - 1$*;*
7·   MERGE ($\boldsymbol{C}, N$)
8·   **end**
9·   **end**
10· **end**

1 *Set* $\boldsymbol{C} = \{C_0\}, C_0 = \{1, 2, \cdots, n\}, N = 1$*;*
2 SPLIT ($\boldsymbol{C}, N, \mathcal{A}_G, C_0$)
3 MERGE ($\boldsymbol{C}, N$)
4 *Return* $\boldsymbol{C}$*;*

continually merges pairs of clusters until Eq. (6) is satisfied. In the "SPLIT" process from line 1* to line 14*, the S-M algorithm is utilized to bi-partition the MTAC graph. Then, this bi-partition is recursively conducted on the sub-graph corresponding to the two clusters until the S-M algorithm fails to achieve a well split which $\bar{\mathbf{c}} < c_\epsilon$, indicating that the graph is completely partitioned. This ensures all clusters meet the constraint (6b). However, only applying the "SPLIT" process is not enough to satisfy the constraint (6a). Thus, the "MERGE" process from line 1· to line 10· continuously merges pairs of clusters which $\bar{c} \geq c_\epsilon$, until all clusters satisfy the constraint (6a). It is easy to prove that merging clusters will not violate the constraint (6b).

To more intuitively identify the correlation-independent module clusters, the clusters correlation matrix $\mathcal{A}_C$ in Eq. (7) spy the clusters correlation in partitioned matrix format. The diagonal blocks $\boldsymbol{c}_a \cdots \boldsymbol{c}_b$ represent the clusters $C_a \cdots C_b$ where $J_{|C_a|}$ denotes the



(a) (b) (c)

Figure 6: The modules correlation matrix (a), the clusters correlation matrix obtained by the S–M algorithm (b) and by the proposed algorithm (c).

all-one matrix shape with modules number in the cluster; The off-diagonal blocks $\boldsymbol{c}_{ab}$ represent the average MTAC between different modules a and b.

$$
\begin{aligned}
\mathcal{A}_C &= \begin{pmatrix} \mathbf{c}_a & \cdots & \mathbf{c}_{ab} \\ \vdots & \ddots & \vdots \\ \mathbf{c}_{ba} & \cdots & \mathbf{c}_b \end{pmatrix} \\
\mathbf{c}_a &= \bar{\mathbf{i}}(C_a) \odot J_{|C_a|} \\
\mathbf{c}_b &= \bar{\mathbf{i}}(C_b) \odot J_{|C_b|} \\
\mathbf{c}_{ab} &= \bar{\mathbf{c}}(C_a, C_b) \odot J_{|C_b||C_a|} = \mathbf{c}_{ba}^T
\end{aligned}
\tag{7}
$$

An example of the correlation-independent modules identification on a real design waveform is shown in Fig. 6. The modules correlation matrix is shown in Fig. 6(a); The clusters correlation matrix obtained by the original S-M algorithm in [32] and the proposed algorithm 1 are shown in Fig. 6(b) and Fig. 6(c) correspondingly. The independent threshold c_ϵ is set to 0.2 and absolute values were taken to better illustrate the strength of MTAC between clusters. The result indicates that clusters from the original algorithm in [32] did not satisfy the constraint (6a), as the correlation weight between clusters in the top-left corner remains larger than the threshold. By merging those violation clusters, the proposed algorithm 1 ensures all the constraints in Eq. (6).

3.2 Worst-case approximation by waveform augmenting

As depicted in the yellow block in Fig. 2, the worst scenario waveform is approximated relying on the MTAC features of the original waveform. The basic idea of augmenting the waveform to the worst-case is aligning the highest toggle slots to assume those correlation-independent modules switching simultaneously, which is illustrated

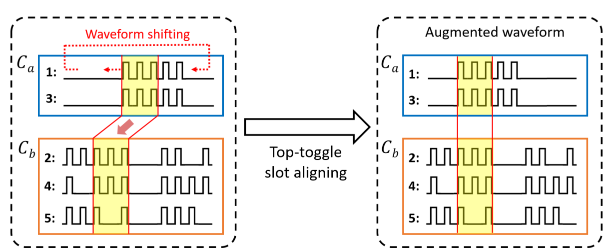


Figure 7: Augment the waveform by aligning the highest switching slot of correlation-independent modules clusters.

Algorithm 2: Approximate the worst peak power and dynamic IR-drop waveform by aligning top-toggle slots.

input : Waveform $\mathbf{V}$ with start time t_0 and end time t_e;
Module clusters disjoint set C;
Modules toggle vector $\mathbf{T}$ with slot length τ

output: Augmented waveform $\hat{\mathbf{V}}$

1 **for** C *in* C **do**
2   *Set* $\mathbf{T}_C = \sum_{m \in C} \mathbf{T}_m$;
3   *Set* $t_s = (\operatorname{argmax}(\mathbf{T}_C) - 1) \cdot \tau$;
4   **for** m *in* C **do**
5   *Set* $\hat{\mathbf{V}}_m = \text{concat}(\hat{\mathbf{V}}_m[t_s : t_e], \hat{\mathbf{V}}_m[t_0 : t_s])$;
6   **end**
7 **end**
8 *Return* $\hat{\mathbf{V}}$;

Table 1: The functionally independent module clusters identification result on multi-channel hash processor

Indepedent channels	2	4	6	8
Identified indepedent module clusters number	6	12	18	24
ARI	0.73	0.80	0.83	0.82

by the waveform shifting process in Fig. 7. The waveforms of modules within the same correlation-independent cluster are considered collectively as their toggle activity are highly interrelated. Altering those could disrupt the MTAC of waveform, leading to an unreasonable augmenting. However, modules in different clusters are behaviorally independent, implying they may switch simultaneously. If such simultaneous switching does not occur in the existing waveform, the worst power and IR-drop scenarios are not covered. By shifting the original waveform to align the highest switching slot, the augmented waveform approximates the worst-case scenario.

The details of generating the augmented waveform are outlined in algorithm 2. The total toggle vector $\mathbf{T}_C$ of all modules in cluster is calculated in line 2. Then, the start time of the highest switching slot t_s is calculated in line 3. Finally, the waveform is shifted so that the highest switching slot is aligned at the beginning of waveform, as indicated in line 5.

4 EXPERIMENTAL RESULT AND DISCUSSION

In this section, the proposed semi-vector-based dynamic power and IR-drop assessment is validated on several test designs. First, the accuracy of identifying functionally independent module clusters is tested in Section 4.1. Then, the proposed semi-vector-based worst peak power and dynamic IR-drop estimation are validated on a modern multi-core ARM CPU design and compared with traditional vector-based and vectorless methods in Section 4.2. The waveforms in all experiments were simulated by a commercial functional verification tool. For module toggle vector construction, the module granularity was set to $G_m = 100$ instances, and the time slot length was set to $\tau = 10$ ns (30+ clock cycles). The independent threshold in algorithm 1 was set to $c_\epsilon = 0.2$.

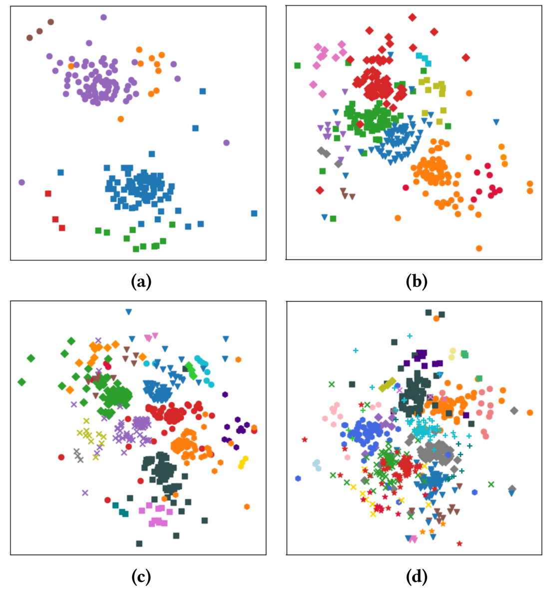


Figure 8: The comparison between functionally-independent module clusters identified in algorithm 1 and the ground truth in parallel hash processor implemented with (a) 2 channels, (b) 4 channels, (c) 6 channels, and (d) 8 channels.

4.1 Accuracy of functionally independent module clusters identification

In this section, the proposed MTAC-based identification of functionally independent module clusters is validated on a hash processor design with 2, 4, 6, and 8 paralleled channels implementation cases. In this test design, the modules belonging to each channel are naturally functionally independent, making it appropriate to verify the accuracy of cluster identification in algorithm 1. For each implementation case, a 10000*ns* test waveform was simulated, including 5 to 8 random hits of each channel.

The MTAC-based clustering result obtained by algorithm 1 and the ground truth is compared in Fig. 8. Each data point in the figure represents a module in the design. The location of the data points in 2-D correlation space was arranged using the spring layout, respecting the MTAC coefficient so that modules with high correlation were plotted near each other. The different colors of data points represent the different clusters identified by the proposed algorithm 1, and the different shapes of data points represent the ground truth classification of modules belonging to different paralleled channels. For all four test designs in Fig. 8, there were a few points with the same color but different shapes. This indicates that the proposed algorithm accurately identified the functionally independent modules belonging to different channels. The numbers of identified independent clusters and the Adjusted Rand Index (ARI) [33], which objectively measures the similarity between two clusterings, were shown in Table 1. The average ARI of the four test designs was 0.79, indicating a good clustering result. The lowest ARI = 0.73 appeared in the 2-channel design. This is because the internal structure of each channel still contains functionally independent parts that are not reflected in the ground truth labels but were still captured by the proposed algorithm 1.

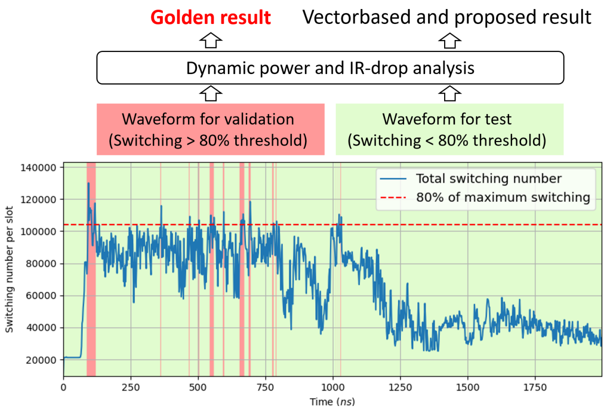


Figure 9: Selection of validation waveform and test waveform.

Table 2: Statistics of design blocks in a modern CPU

	Type	Instances number	Modules number	Clusters number	Augmented corr. error
1	CPU-core	600000+	200+	7	0.0058
2	CPU-core	1000000+	100+	6	0.0007
3	CPU-core	1500000+	1000+	3	0.0009
4	Cache	1000000+	500+	13	0.0081
5	SoC-core	3000000+	1500+	15	0.0029
6	Soc-core	1500000+	1000+	13	0.0018
7	Inter	1000000+	1000+	3	0.0030
8	Inter	1000000+	1000+	3	0.0024

4.2 Worst-case dynamic power and IR-drop estimation

The proposed semi-vector-based assessment was tested based on a cutting-edge multi-core ARM CPU design. Eight different design blocks, including the CPU core, the system-on-chip (SoC) core, the cache, and the interconnection controller, were assessed by both the traditional vector-based [4], vectorless [6], and the proposed semi-vector-based flow. The statistics of each design block are summarized in Table 2. The type of design is marked under "Type"; The instance number and module number of each design are marked under "Instances number" and "Modules number";

All design blocks had a 2000ns waveform simulated from the Dhrystone and Whetstone benchmarks and the waveforms were quantified to 1000 slots with slot length $\tau = 2ns$. However, the ground truth of the worst peak power and dynamic IR-drop may not covered by the known waveforms. To validate the accuracy of the worst-case estimation, the known waveforms were partitioned into two parts, as depicted in Fig. 9. The segments of the waveform with a switching count exceeding 80% of the global maximum values were designated as the validation waveform, while the remaining segments were designated as the test waveform. The maximum peak power and dynamic IR-drop in the validation waveform were designated as the golden value, approximating the ground truth. Subsequently, the vector-based assessment and the proposed semi-vector-based assessment were applied to the test waveform. The validation waveform remained concealed during both the vector-based and the proposed assessments. On the other hand, the vectorless assessment did not require waveform input. The toggle rate of primary input and sequential activity was set to 0.2 and 0.05 according to the design experiences. The values of peak power and dynamic IR-drop were solved by a cutting-edge commercial power integrity tool.

For each design block, the number of correlation-independent module clusters were compared in "Clusters number" in Table 2. The clusters correlation matrices of design block 1~8 are also compared in Fig. 10(a) ~ Fig. 10(h) correspondingly. The results demonstrate a strong relationship between the number of clusters and the type of design, with the number of clusters in SoC core blocks being significantly larger than that in the CPU core blocks and the interconnection controller blocks. Additionally, the clusters in SoC core blocks Fig. 10(e) and Fig. 10(f) exhibit stronger independence compared to the clusters in CPU core blocks Fig. 10(a) ~ Fig. 10(c) as indicated by the smaller average correlation coefficients in the off-diagonal positions. It is reasonable to assume that the SoC cores possess more functionally independent clusters, considering the multitude of integrated components. Generating a test case waveform covering all combinations of these components being activated is challenging. Therefore, the proposed semi-vector-based assessment estimated a potential worst-case scenario by assuming those are worked together during waveform augmenting. The "Augmented corr-error" in Table 2 measures the MTAC difference between the original waveform and the augmented waveform, calculated by the Mean Squared Error (MSE) $\frac{2}{n^2-n}\sum_{ij}(c^o_{ij} - c^a_{ij})^2$ between MTAC coefficients c^o in the original waveform and c^a in the augmented waveform. The augmented waveform of all design blocks exhibited a comparatively small augmented error which is less than 0.01, manifesting minimal disagreement with the MTAC in the original waveform.

The worst peak power and dynamic IR-drop results obtained by the proposed semi-vector-based assessment are summarized in Table 3 and compared to those of traditional vector-based and vectorless assessments, as well as the golden result. The results obtained by three assessments are compared under the "Worst peak-power" and "Worst dyn-IR-drop". The "runtime" compares the overall time consumption from loading the design database to dumping the assessment results. The experimental results of all design blocks indicate that the proposed semi-vector-based assessment flow yields more reasonable results, which are closest to the golden value for both power and IR-drop. Although the proposed assessment did not entirely cover the worst result, it yielded the smallest error with the relative error in peak power at 2.93% and the relative error in dynamic IR-drop at 7.62%. In comparison, the traditional vector-based method underestimates both the peak power and dynamic IR-drop, with relative errors of 8.98% and 19.05% respectively. On the other hand, although the vectorless assessment results covered the golden worst case for all design blocks, it made significantly pessimistic estimations with relative errors of 99.12% and 50.61% respectively, which could lead to substantial overdesign.

The estimation of IR-drop hotspot regions is also illustrated in Fig. 11. Comparing to the golden result in Fig. 11(d), The vector-based result in Fig. 11(a) failed to identify the IR-drop weak regions highlighted by the purple boxes. In contrast, although the vectorless result covered these weak regions, it exhibited an overly pessimistic estimation by encompassing too many additional areas. On the other hand, the proposed result in Fig. 11(c) accurately estimated these weak regions. It is noteworthy that the proposed assessment

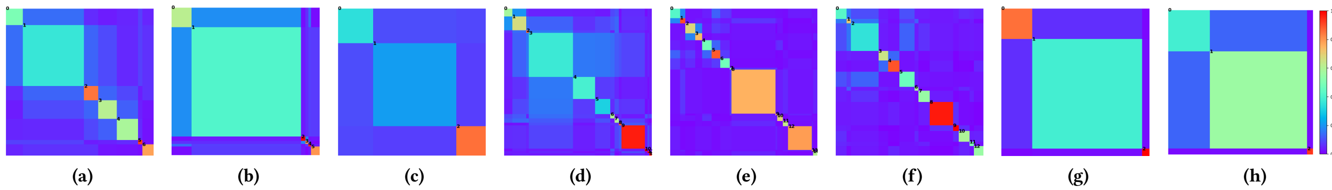


Figure 10: The clusters correlation matrices (a) ~ (h) corresponding to the design blocks 1 ~ 8 in Table 2.

Table 3: The worst peak power and dynamic IR-drop results on the industrial multi-core CPU design blocks.

	Worst peak-power (mW)				Worst dyn-IR-drop (mV)				Runtime (Minutes)		
	Golden	V-based	V-less	Proposed	Golden	V-based	V-less	Proposed	V-based	V-less	Proposed
1	1372	1308	2905	**1352**	157.8	131.9	208.9	**143.4**	128	52	136
2	3175	2984	4333	**3254**	191.8	160.6	165.8	**189.6**	201	93	242
3	4371	4113	7983	**4468**	203.0	193.0	261.8	**210.6**	258	107	283
4	1981	1704	3428	**1928**	151.5	100.6	247.6	**147.1**	111	81	124
5	2435	2199	4489	**2449**	167.2	158.3	274.7	**173.3**	436	172	488
6	455	356	1411	**407**	119.6	108.1	191.0	**127.1**	161	130	179
7	2536	2380	5020	**2589**	149.2	92.3	282.3	**174.4**	322	170	357
8	2433	2352	4809	**2464**	159.0	114.2	244.4	**187.0**	281	144	321

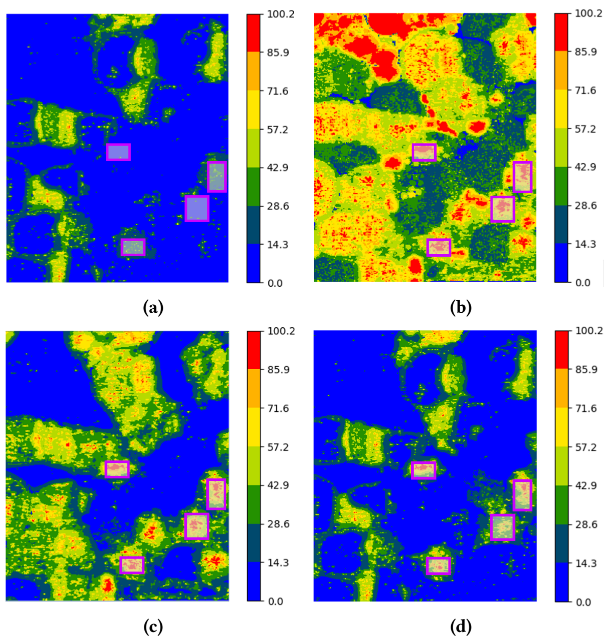


Figure 11: The dynamic IR-drop hot-spot map result of (a) the vector-based assessment, (b) the vectorless assessment, (c) the proposed assessment, and (d) the golden result.

also inevitably overestimated the dynamic IR-drop in some regions, as the augmented waveform was generated based on the behavioral correlation of instances, which may not exactly match the validation waveform. The proposed method still outperformed the traditional since it was significantly less pessimistic while ensuring that all IR-drop defects were covered.

The probability density function (PDF) of instances IR-drop distribution from the three results are also compared with the golden result in Fig. 12(a), which illustrates that the distribution from the proposed result is most similar to the golden distribution. As a reference, the correlation-independent module clusters are shown in

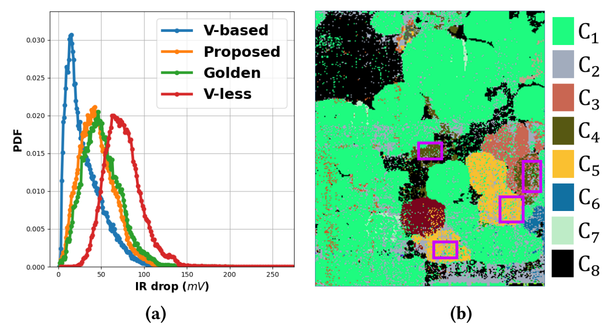


Figure 12: The PDF of instances dynamic IR-drop (a) and the correlation-independent module clusters (b).

Fig. 12(b) where the different colors of instances represent their belonging to different clusters identified by the algorithm 1. The highlighted regions appeared in the intersection regions of clusters. These clusters did not experience simultaneous high switching in the test waveform but appeared in the validation waveform. Therefore, the vector-based assessment optimistically estimated the dynamic IR-drop. On the other hand, the proposed assessment identified these clusters and approximated a simultaneous switching scenario by vector augmenting, thus covering all weak regions.

5 CONCLUSION

This paper proposes a semi-vector-based assessment method to estimate the worst-case peak power and dynamic IR-drop. The module toggle activity correlation (MTAC) of the waveform is quantified, and functionally independent modules are identified through graph clustering. The waveform is augmented by aligning the highest switching slots of functionally independent modules. By applying dynamic power and IR-drop analysis to the augmented waveform, the potential worst-case peak power and dynamic IR-drop are revealed. Experimental results demonstrate that the proposed assessment bridges the gap between vector-based and vectorless assessments and provides a better approximation of the real worst-case.

REFERENCES

[1] D. E. Lackey, P. S. Zuchowski, T. R. Bednar, D. W. Stout, S. W. Gould, and J. M. Cohn, "Managing power and performance for system-on-chip designs using voltage islands," in *Proceedings of the 2002 IEEE/ACM international conference on Computer-aided design*, 2002, pp. 195–202.

[2] S. Lin and N. Chang, "Challenges in power-ground integrity," in *IEEE/ACM International Conference on Computer Aided Design. ICCAD 2001. IEEE/ACM Digest of Technical Papers (Cat. No. 01CH37281)*. IEEE, 2001, pp. 651–654.

[3] K. Daloukas, N. Evmorfopoulos, P. Tsompanopoulou, and G. Stamoulis, "Parallel fast transform-based preconditioners for large-scale power grid analysis on graphics processing units (gpus)," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 35, no. 10, pp. 1653–1666, 2016.

[4] Rouatbi and Haroun, "Power estimation tool for sub-micron cmos vlsi circuits," in *1992 IEEE/ACM International Conference on Computer-Aided Design*. IEEE, 1992, pp. 204–209.

[5] C.-T. Hsieh, J.-C. Lin, and S.-C. Chang, "Vectorless estimation of maximum instantaneous current for sequential circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 11, pp. 2341–2352, 2006.

[6] F. N. Najm, "Transition density: A new measure of activity in digital circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 12, no. 2, pp. 310–323, 1993.

[7] P. H. Schneider and S. Krishnamoorthy, "Effects of correlations on accuracy of power analysis-an experimental study," in *Proceedings of 1996 International Symposium on Low Power Electronics and Design*. IEEE, 1996, pp. 113–116.

[8] D. Marculescu, R. Marculescu, and M. Pedram, "Stochastic sequential machine synthesis targeting constrained sequence generation," in *Proceedings of the 33rd annual Design Automation Conference*, 1996, pp. 696–701.

[9] S. Bhanja and N. Ranganathan, "Cascaded bayesian inferencing for switching activity estimation with correlated inputs," *IEEE transactions on very large scale integration (VLSI) systems*, vol. 12, no. 12, pp. 1360–1370, 2004.

[10] R. Marculescu, D. Marculescu, and M. Pedram, "Probabilistic modeling of dependencies during switching activity analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, no. 2, pp. 73–83, 1998.

[11] M. G. Xakellis and F. N. Najm, "Statistical estimation of the switching activity in digital circuits," in *Proceedings of the 31st annual Design Automation Conference*, 1994, pp. 728–733.

[12] C.-T. Hsieh, J.-C. Lin, and S.-C. Chang, "Efficient transition-mode boolean characteristic function with its application to maximum instantaneous current analysis," in *8th International Symposium on Quality Electronic Design (ISQED'07)*. IEEE, 2007, pp. 602–606.

[13] Y.-M. Jiang, K.-T. Cheng, and A. Kristic, "Estimation of maximum power and instantaneous current using a genetic algorithm," in *Proceedings of CICC 97-Custom Integrated Circuits Conference*. IEEE, 1997, pp. 135–138.

[14] H. Kriplani, F. N. Najm, and I. N. Hajj, "Pattern independent maximum current estimation in power and ground buses of cmos vlsi circuits: Algorithms, signal correlations, and their resolution," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 14, no. 8, pp. 998–1012, 1995.

[15] S. K. Varadharajan and V. Nallasamy, "Low power vlsi circuits design strategies and methodologies: A literature review," in *2017 Conference on Emerging Devices and Smart Systems (ICEDSS)*. IEEE, 2017, pp. 245–251.

[16] J. Yang, Y. Cai, Q. Zhou, and W. Zhao, "A selected inversion approach for locality driven vectorless power grid verification," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 11, pp. 2617–2628, 2014.

[17] D. Kouroussis and F. N. Najm, "A static pattern-independent technique for power grid voltage integrity verification," in *Proceedings of the 40th annual Design Automation Conference*, 2003, pp. 99–104.

[18] X. Xiong and J. Wang, "Vectorless verification of rlc power grids with transient current constraints," in *2011 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE, 2011, pp. 548–554.

[19] Z. Zhao and Z. Feng, "A spectral graph sparsification approach to scalable vectorless power grid integrity verification," in *Proceedings of the 54th Annual Design Automation Conference 2017*, 2017, pp. 1–6.

[20] H. Zhu, Y. Wang, F. Liu, X. Li, X. Zeng, and P. Feldmann, "Efficient transient analysis of power delivery network with clock/power gating by sparse approximation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 3, pp. 409–421, 2015.

[21] I. A. Ferzli, F. N. Najm, and L. Kruse, "A geometric approach for early power grid verification using current constraints," in *2007 IEEE/ACM International Conference on Computer-Aided Design*. IEEE, 2007, pp. 40–47.

[22] F. N. Najm, "Overview of vectorless/early power grid verification," in *Proceedings of the International Conference on Computer-Aided Design*, 2012, pp. 670–677.

[23] Y. A. Durrani and T. Riesgo, "Efficient power analysis approach and its application to system-on-chip design," *Microprocessors and Microsystems*, vol. 46, pp. 11–20, 2016.

[24] M. Lee, C. Choi, D. Seo, B. Bang, Y. Kang, and W. Paik, "Improving analysis coverage for dynamic ir drop sign-off in finfet soc design," in *2020 International SoC Design Conference (ISOCC)*. IEEE, 2020, pp. 332–333.

[25] S. Soman, A. Brahme, R. Venkatraman, R. Shaikh, S. Thiyagaraja, and M. Patil, "Ensuring on-die power supply robustness in high-performance designs," in *2011 24th Internatioal Conference on VLSI Design*. IEEE, 2011, pp. 220–225.

[26] J. Hao and V. Bertacco, "Powerranger: Assessing circuit vulnerability to power attacks using sat-based static analysis," in *2009 IEEE International High Level Design Validation and Test Workshop*. IEEE, 2009, pp. 54–59.

[27] L. Ren, S. Sun, M. Deo, J. Jaffari, P. Anmula, P. Boyle, J. L. Drewniak, and D. G. Beetner, "A vectorless approach for predicting switching activity in a digital circuit," *IEEE Transactions on Electromagnetic Compatibility*, vol. 58, no. 3, pp. 828–835, 2016.

[28] I. Cohen, Y. Huang, J. Chen, J. Benesty, J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," *Noise reduction in speech processing*, pp. 1–4, 2009.

[29] Z. Shi, J. Xiao, J. Jiang, Y. Zhang, and Y. Zhou, "Identifying reliability high-correlated gates of logic circuits with pearson correlation coefficient," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2023.

[30] Z. Wu and R. Leahy, "An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 15, no. 11, pp. 1101–1113, 1993.

[31] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, pp. 395–416, 2007.

[32] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[33] D. Steinley, "Properties of the hubert-arable adjusted rand index." *Psychological methods*, vol. 9, no. 3, p. 386, 2004.