

CSCO: A Backside-PDN-Aware Clock-Signal Co-Optimization Framework for Improved PPA

Zixiao Wang
CUHK
Hong Kong, China
zxwang22@cse.cuhk.edu.hk

Leilei Jin*
CUHK
Hong Kong, China
leileijin@cuhk.edu.hk

Zhen Zhuang
CUHK
Hong Kong, China
zhuang_zhen@126.com

Rongmei Chen
Peking University
Beijing, China
crm@pku.edu.cn

Bei Yu*
CUHK
Hong Kong, China
byu@cse.cuhk.edu.hk

Abstract

Backside power delivery networks (BSPDN) have emerged as a promising technology for advanced logic nodes to address IR-drop and PPA challenges. While BSPDN introduces additional routing resources on the backside, these resources are limited and must be carefully partitioned between clock and signal nets, creating a critical resource allocation tradeoff. Prior work either moves only the clock network or assumes a fixed clock tree and optimizes only signal nets, failing to explore the tradeoff space of backside resource allocation. Moreover, lacking frontside power-ground shielding, BSPDN introduces severe signal integrity (SI) degradation. We propose CSCO, a data-driven BSPDN-aware co-optimization framework that jointly allocates limited backside resources between clock and signal nets across frontside/backside layers. CSCO employs efficient search strategies to identify critical nets for backside routing without repeated evaluation, navigating the clock-signal allocation tradeoff to balance IR-drop, routing congestion, and PPA. The framework also leverages backside routing to mitigate coupling noise and crosstalk-induced SI issues. Experiments demonstrate improved WNS/TNS, frequency, and SI robustness without additional shielding overhead.

CCS Concepts

• **Hardware** → **Physical design (EDA)**; *Power and energy*; *Timing analysis*.

*Corresponding authors.



This work is licensed under a Creative Commons Attribution 4.0 International License.

ICCAD '26, San Jose, CA, USA

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2873-0/2026/11

<https://doi.org/10.1145/3831252.3833949>

Keywords

backside power delivery network, backside routing, clock-signal co-optimization, physical design, signal integrity

ACM Reference Format:

Zixiao Wang, Leilei Jin, Zhen Zhuang, Rongmei Chen, and Bei Yu. 2026. CSCO: A Backside-PDN-Aware Clock-Signal Co-Optimization Framework for Improved PPA. In *IEEE/ACM International Conference on Computer-Aided Design (ICCAD '26)*, November 08–12, 2026, San Jose, CA, USA. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3831252.3833949>

1 Introduction

As technology scaling advances into the sub-3nm era, physical design faces increasingly severe back-end-of-line (BEOL) challenges including routing congestion, IR-drop, and reliability degradation [8, 32, 35, 38]. Metal resources traditionally allocated for signal routing and power delivery have become insufficient to sustain aggressive power, performance, and area (PPA) targets, making BEOL architecture and power delivery the dominant bottlenecks in system-level scaling. A promising solution has emerged through backside power delivery networks (BSPDN) [9, 20, 40]. By relocating the power grid to the wafer backside via nano-through-silicon-via (nTSV) connections, BSPDN alleviates IR-drop while liberating frontside metal layers for denser signal routing and higher operating frequencies as a critical enabler for next-generation design-technology co-optimization (DTCO) and 3D IC integration [12, 18, 19, 26, 28, 36].

However, this architectural shift breaks the conventional separation between signal, clock, and power networks, demanding new cross-layer co-design methodologies [7, 11, 15, 27]. Current research has explored backside routing in progressively broader yet still isolated ways, as illustrated in Fig. 1. Early BSPDN implementations focus exclusively

This work is supported by The Research Grants Council of Hong Kong SAR (No. CUHK14211824, CUHK14201624, and T46-415/25-R).

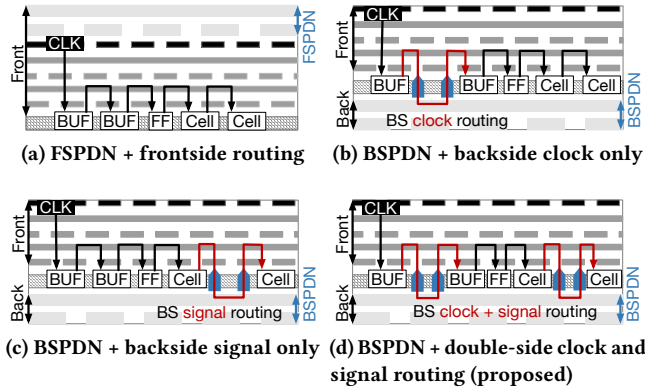


Figure 1: Backside design for clock-signal co-optimization with limited backside resources.

on power delivery without exploiting residual backside routing resources [27]. Subsequent clock-specific approaches (Fig. 1(b)) relocate clock networks using specialized backside buffer cells, demonstrating improvements in clock skew and latency [7, 17, 35]. Other methods [16, 20, 22, 34] have explored the relocation of signal nets according to their geometry features (Fig. 1(c)), with an assumption that the clock net is pre-constructed.

Collectively, these BSPDN approaches suffer from three critical limitations that fundamentally constrain achievable PPA gains. First, clock and signal optimization remain decoupled, so methods targeting only clocks neglect timing-critical signal paths, while geometric heuristics for signals often miss compact yet performance-critical nets due to weak correlation between timing criticality and geometric properties. Second, existing approaches usually presuppose fixed power grid configurations (e.g., predetermined metal stripe widths and pitches) prior to net selection. However, the backside net routing and PDN configuration are inherently dependent. PDN synthesis proceeds without knowledge of routing demands, while routing engines ignore PDN-induced blockages, fragmenting the design space and preventing discovery of design-specific optimal solutions that balance power integrity, routing congestion, and timing closure. Third, existing works overlook a newly emergent physical effect: removing frontside power grids eliminates electromagnetic shielding between signal nets, creating a "gravity-less" coupling environment where crosstalk increases 2–4× [13]. Without the confinement of frontside power straps, signal metals experience amplified noise propagation that can offset or even negate BSPDN’s timing benefits [10, 21, 23]. This insight motivates a co-optimization framework that exploits backside power layers as shields for SI-critical paths while jointly allocating backside resources between signal and clock nets.

To address these issues, we propose CSCO, a BSPDN-aware co-optimization framework that jointly determines the BSPDN topology and the allocation of signal and clock nets across frontside and backside layers (Fig. 1(d)). Unlike prior approaches that fix PDN configurations before net selection, CSCO recognizes that optimal power grid structures depend on design-specific routing demands and explores both dimensions simultaneously. It employs fast surrogate modeling to navigate the coupled design space without costly full-routing iterations and exploits BSPDN shielding to protect SI-critical nets. This approach enables discovery of the better solutions that balance power integrity, signal integrity, and PPA metrics. Our contributions are:

- **Unified Co-Optimization Framework:** The first methodology that concurrently synthesizes dual-side signal and clock networks under BSPDN constraints, achieving cross-layer integration of routing and power delivery planning.
- **Data-Driven Surrogate Modeling:** A low-rank surrogate models timing–resource coupling without routing.
- **Signal Integrity Mitigation via BSPDN Shielding:** Uses BSPDN’s electromagnetic properties to improve SI.
- **Demonstrated Effectiveness:** On realistic CPU benchmarks, CSCO achieves significant improvement in PPA metrics while satisfying power and signal integrity constraints.

2 Preliminaries

We first introduce the necessary background.

2.1 Backside PDN

BSPDNs use thick, low-resistivity backside metals connected to frontside buried power rails by sub-100-nm nTSVs [24, 29, 37, 40, 41]. This separation reduces power-distribution resistance by up to 30% and frees frontside layers for signal routing [11, 12, 18, 36], as shown in Fig. 2(a).

Clock-Centric Backside Routing. Existing approaches [17, 27, 32, 35] relocate entire clock networks using backside buffer cells and nTSVs, improving skew and latency through GNN-based endpoint classification [7] or analytical optimization [17]. However, they ignore timing-critical signal routes and may migrate complete trees when only critical branches need optimization, wasting backside resources and nTSVs.

Signal-Centric Backside Routing. Signal-centric methods [16, 20, 22, 34] select nets using bounding-box size, pin count, or routing congestion. These geometric proxies can miss compact critical nets and favor large noncritical nets, producing suboptimal allocations. Net-level migration can also conflict with backside PDN structures, causing DRC violations that require post-routing fixes [27].

Lack of Clock-Signal Co-optimization. Existing works optimize clock and signal routing sequentially [7, 17, 22, 34,

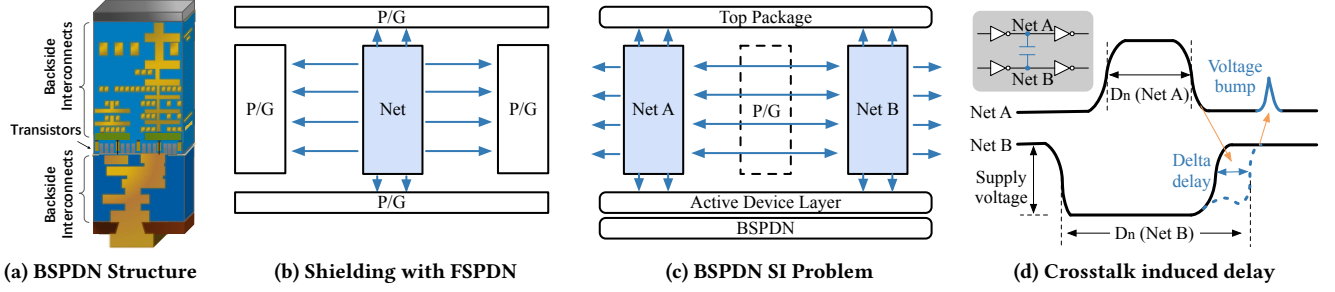


Figure 2: SI problem on BSPDN design.

35], despite their interdependence and competition for limited backside resources. Allocating more resources to clocks necessarily leaves fewer for signals, and vice versa. Clock tree synthesis determines signal-path timing budgets, whereas subsequent signal migration cannot influence clock routing. This separation precludes joint clock-signal optimization.

SI Degradation Induced by BSPDN. Relocating the PDN to the backside removes frontside P/G grids that normally shield signal nets [13]. As shown in Figs. 2(b) and 2(c), FSPDN strips maintain favorable self-to-coupling capacitance ratios, whereas BSPDN eliminates this shielding and increases direct coupling between adjacent nets. The resulting aggressor-victim interference causes delta delay ΔD and voltage bumps on timing-critical nets (Fig. 2(d)). Backside signal routing can therefore exacerbate SI issues, requiring joint timing, routability, and SI optimization.

2.2 Efficient Surrogate Model

Machine-learning surrogates replace costly physical-design evaluations in EDA. Neural models predict routability [39], pre-routing timing [6], and slack [14], but typically require thousands of full-routing and signoff labels. Gaussian-process Bayesian optimization [31] is more data-efficient and used in analog synthesis [25], but cubic scaling and continuous, low-dimensional assumptions limit high-dimensional binary allocation. Factorization machines (FMs) [30], originally developed for recommendation systems, capture pairwise interactions through low-rank embeddings with $\mathcal{O}(nl)$ complexity, require little data, and naturally handle binary inputs. This compact parameterization suits costly signoff labels. Their success in materials and drug-discovery optimization [33], together with limited use in EDA, motivates our adoption.

2.3 Evaluation Metrics

We evaluate the practical impact of backside allocation on overall design performance.

WNS and TNS. Worst Negative Slack (T_{wns}) and Total Negative Slack (T_{tns}) are standard timing indicators; larger values indicate better performance and higher effective frequency.

Delta-Delay and Bump Ratios. Signal integrity is characterized using the delta-delay ratio and bump ratio:

$$r_{delay} = \frac{\Delta D}{D_d + D_n + \Delta D}, \quad r_{bump} = \frac{V_{bump}}{V_{DD}}. \quad (1)$$

Here, D_d and D_n are the intrinsic driver and net delays, and V_{DD} is the supply voltage. Higher ratios indicate stronger SI-induced delay or voltage disturbance; we also count nets exceeding preset thresholds.

3 Methods

We now introduce the proposed BSPDN-aware design framework, CSCO, which operates within the backside-aware design flow detailed in Section 4.1. As illustrated in Fig. 3, CSCO augments the conventional 3D-IC implementation process with data-driven stages that jointly optimize net partitioning, BSPDN planning, and timing.

Framework Overview. CSCO operates in two major phases, training and inference, to transfer knowledge from empirical exploration to efficient design-space prediction. The training phase focuses on data collection and surrogate model learning, while the inference phase exploits the trained model for BSPDN-aware optimization and timing estimation.

CSCO's *key novelty* is a surrogate-driven BSPDN formulation unifying PDN planning, net partitioning, and timing.

3.1 Problem Formulation

The goal of this stage is to determine the optimal allocation of all nets $\mathcal{N}$ in a given design under a specific BSPDN configuration. Let $\mathcal{N} = \{s_1, \dots, s_m, c_1, \dots, c_n\}$, where s_i and c_j denote signal and clock nets, respectively. With above symbols, we formulate backside resource allocation as an optimization problem under BSPDN-dependent constraints:

$$\text{Eval}(\mathbf{x}) = T_{wns}(\mathbf{x}) + \lambda T_{tns}(\mathbf{x}), \quad (2)$$

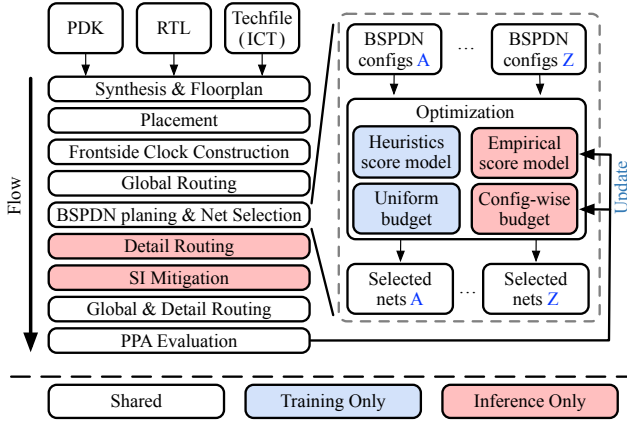


Figure 3: Overall design flow of CSCO.

where T_{wns} and T_{tns} denote the worst and total negative slack, respectively, and λ controls their relative importance. The design variable is a binary vector

$$x_i \in \{0, 1\}, \quad \forall i \in [1, m + n], \quad (3)$$

where $x_i=1$ indicates that net i is assigned to the backside, otherwise the net is assigned to the frontside.

This optimization problem is challenging for the following three reasons. First, evaluating any candidate $\mathbf{x}$ requires full routing and timing analysis, making each trial extremely expensive. Second, the search space is a high-dimensional 0–1 domain with hundreds of thousands of variables, rendering brute-force or classical combinatorial methods infeasible. Third, extending this basic formulation to be explicitly BSPDN-aware is nontrivial: the BSPDN configuration simultaneously impacts IR-drop, routing resources, and timing, and must be incorporated through a compact yet physically meaningful constraint, which we introduce next via a configuration-dependent nTSV budget model.

BSPDN-Dependent Constraints and Budget Determination. Unlike prior works that treat nTSV resources as fixed, we explicitly model BSPDN dependence, allowing power-network density to directly affect routing capacity and timing. A denser BSPDN enhances IR-drop robustness but limits routing resources, so we capture this trade-off through a global nTSV budget constraint:

$$\sum_{i=1}^{m+n} t_i x_i \leq B, \quad (4)$$

where t_i is the number of nTSVs required by net i (equal to its pin count). Timing improves as B increases until a configuration-dependent threshold, beyond which excessive nTSVs degrade routability, as shown in Fig. 4.

The budget B is determined empirically by sweeping nTSV allocations and recording corresponding PPA outcomes. The

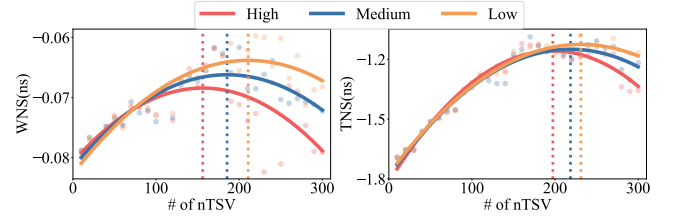


Figure 4: Timing performance (WNS and TNS) as a function of nTSVs count under different BSPDN densities.

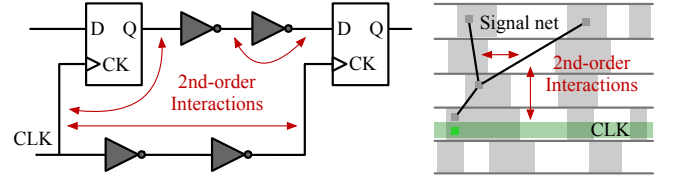


Figure 5: The second-order relationship between nets.

inflection point of the performance–nTSV curve is selected as the final budget, which remains fixed during inference for consistent evaluation across configurations. This BSPDN-aware formulation links power-delivery planning with timing optimization, bridging two stages traditionally treated independently in 3D-IC design flows.

Directly optimizing Eqs. (2) to (4) is prohibitive because routing and timing analysis are expensive. We therefore approximate $\text{Eval}(\mathbf{x})$ with a data-efficient surrogate.

3.2 Surrogate Model

The true evaluation function $\text{Eval}(\mathbf{x})$ requires full routing and timing analysis, making data collection extremely expensive. Such scarcity renders deep-learning surrogates impractical, motivating the use of a lightweight and interpretable model that can learn from limited samples.

Low-Rank Interaction Modeling. To efficiently approximate the objective function Eq. (2) with limited data, we employ a second-order FM model [30]. Given a binary partition vector $\mathbf{x} \in \{0, 1\}^{m+n}$ indicating backside assignment, the k -order FM prediction is defined as:

$$\hat{y}(\mathbf{x}) = w_0 + \mathbf{w}^\top \mathbf{x} + \sum_{d=2}^k \sum_{i_1 < \dots < i_d} \langle \mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_d} \rangle \prod_{j=1}^d x_{i_j}, \quad (5)$$

where $w_0 \in \mathbb{R}$ is a global bias, $\mathbf{w} \in \mathbb{R}^{m+n}$ models first-order (individual-net) effects, $\mathbf{v}_i \in \mathbb{R}^l$ is a latent embedding vector representing the hidden timing characteristics of net i and the operator $\langle \mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_d} \rangle = \sum_{f=1}^l \prod_{j=1}^d v_{i_j, f}$ captures d -way interactions. All $\mathbf{v}_i$ are assembled into the latent matrix $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{m+n}]^\top$, so that $\{w_0, \mathbf{w}, \mathbf{V}\}$ constitute all learnable parameters of the model.

In this work, we adopt the second-order case ($k=2$) for several circuit-level considerations: 1) most timing dependencies arise from *pairwise* interactions, including between signal nets connected to the same cell, and between clock and signal nets along critical paths, as illustrated in Fig. 5. Higher-order correlations contribute marginally yet increase data complexity. Thus, a second-order FM effectively captures the dominant physical dependencies while remaining data efficient. 2) Compared with neural surrogates, FM offers greater interpretability and stability under limited training data, making it particularly suitable for BSPDN-aware optimization. 3) Furthermore, once trained, the FM model generalizes across different BSPDN configurations without retraining, since it encodes intrinsic inter-net relationships rather than configuration-specific parameters. This cross-configuration consistency allows the same surrogate to guide optimization under varying PDN densities.

Additionally, a standard simplification [30] for $k=2$ reduces computational complexity from $\mathcal{O}((m+n)^2l)$ in Eq. (5) to $\mathcal{O}((m+n)l)$:

$$\hat{y}(x) = w_0 + w^\top x + \frac{1}{2}(\|V^\top x\|_2^2 - \|\text{Diag}(x)V\|_F^2), \quad (6)$$

which significantly improves efficiency and is well-suited for GPU parallelization.

Cold-Start Data Construction. Before training the surrogate model, we require an initial dataset for cold-start learning. We propose an empirical selection criterion to identify nets most likely to impact timing when migrated to the backside. For each net type (signal or clock), we extract five lightweight post-global-routing metrics available without detailed routing (Table 1).

After normalizing each metric to $[0, 1]$, we compute a composite score quantifying the timing improvement potential of each net:

$$\text{Score}_{\text{sig}} = \alpha_1 \tilde{t}_{\text{tran}} + \alpha_2 \tilde{n}_{\text{load}} + \alpha_3 \tilde{C}_{\text{est}} + \alpha_4 \tilde{N}_{\text{path}} - \alpha_5 \tilde{r}_{\text{first}}, \quad (7)$$

$$\text{Score}_{\text{clk}} = \beta_1 \tilde{t}_{\text{tran}} + \beta_2 \tilde{n}_{\text{load}} + \beta_3 \tilde{N}_{\text{crit}} - \beta_4 \tilde{t}_{\text{arr}} - \beta_5 \tilde{r}_{\text{first}}, \quad (8)$$

where $(\tilde{\cdot})$ denote normalized values and α_i, β_i are weighting coefficients (set equally in this work). Higher scores identify nets likely to improve timing after backside migration.

This criterion efficiently constructs the training dataset by focusing computation on timing-critical nets, thereby reducing routing iterations required for model initialization.

3.3 Optimization and Inference

With the surrogate established, optimization uses training to generate diverse samples and inference to obtain the BSPDN-aware net partition.

Training Phase: Greedy + Exploration. During training, we aim to collect high-quality yet diverse partitioning samples. Nets are first ranked by their normalized utility, defined

Table 1: Normalized post-global-routing metrics used for cold-start net selection.

Symbol (Unit)	Definition	Range	Net Type
$\tilde{t}_{\text{tran}}$ (ps)	Transition delay	≥ 0	Both
$\tilde{r}_{\text{first}}$ (N/A)	Rank of first critical path	$\{1, \dots, m+n\}$	Both
$\tilde{n}_{\text{load}}$ (N/A)	Number of loads	$\mathbb{Z}^+$	Both
$\tilde{C}_{\text{est}}$ (fF)	Estimated net capacitance	≥ 0	Signal
$\tilde{N}_{\text{path}}$ (N/A)	Occurrence count in all paths	$\{1, \dots, m+n\}$	Signal
$\tilde{t}_{\text{arr}}$ (ns)	Arrival time	≥ 0	Clock
$\tilde{N}_{\text{crit}}$ (N/A)	Occurrences in top-200 critical paths	$\{0, \dots, 200\}$	Clock

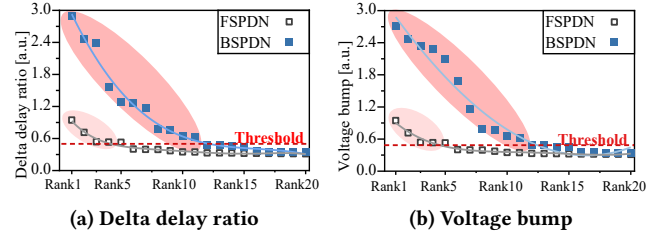


Figure 6: Crosstalk impacts on nets: voltage bump and delta-delay at 4GHz operating frequency.

as the empirical score (Eqs. (7) and (8)) divided by the required number of nTSVs t_i . We then perform greedy selection under a preset clock/signal ratio until the total nTSV budget is reached. To enhance diversity, a fixed fraction of selected nets is randomly replaced in each run, enabling broader exploration and improving surrogate coverage.

Inference Phase: FM-Guided Greedy + Simulated Annealing. In inference, the trained FM model provides a fast estimation of timing performance. For each net, we compute its importance by combining the first-order term and its pairwise interactions in the FM output, then divide by t_i to reflect its benefit per backside cost:

$$I_i = \frac{w_i + \sum_{j \neq i} \langle v_i, v_j \rangle x_j}{t_i}, \quad (9)$$

where I_i denotes the importance of net i , w_i is its first-order coefficient, $\langle v_i, v_j \rangle$ represents the pairwise interaction from the FM model, and t_i is the number of nTSVs required by the net. We again perform greedy selection under the clock/signal ratio until the total budget is filled, yielding an initial feasible solution. To further refine this solution, we adopt simulated annealing (SA) directly on the surrogate. At each iteration, we swap a few nets between dies and accept the resulting candidate x' with probability:

$$P_{\text{accept}} = \min\left(1, \exp\left(\frac{\hat{y}(x') - \hat{y}(x)}{T}\right)\right), \quad (10)$$

Table 2: Performance metrics comparison across all benchmarks at 2GHz target frequency. Bold values indicate the best results. FS denotes frontside PDN with frontside routing. BSPDN+FR denotes BSPDN with frontside routing.

	SHA256 (#Cell = 7.9k)					JPEG (#Cell = 182k)					ARM Cortex-A7 (#Cell = 108k)				
	FS	BSPDN+FR	DAC24[7]	DAC25[17]	CSCO	FS	BSPDN+FR	DAC24[7]	DAC25[17]	CSCO	FS	BSPDN+FR	DAC24[7]	DAC25[17]	CSCO
Utilization (%)	82.91	82.36	82.36	82.36	82.36	78.24	75.50	75.50	75.50	75.50	78.16	74.15	74.15	74.15	74.15
Area (μm^2)	1319	1311	1311	1311	1311	31285	32280	32280	32280	32280	14267	13491	13491	13491	13491
#Metal (BS+FS)	0+6	3+6	3+6	3+6	3+6	0+7	3+7	3+7	3+7	3+7	0+8	3+8	3+8	3+8	3+8
#nTSV	N/A	N/A	270	194	230	N/A	N/A	396	768	800	N/A	N/A	316	634	600
Clock Skew (ps)	24.17	27.94	27.02	19.72	32.09	112.16	35.37	31.93	24.86	34.17	217.67	95.06	66.79	52.97	43.49
Clock Latency (ps)	83.29	75.31	64.63	51.47	75.13	226.26	149.31	112.11	79.31	124.46	456.98	206.51	147.67	123.87	135.80
Wirelength (m)	0.05	0.04	0.04	0.04	0.04	0.81	0.70	0.69	0.68	0.69	1.47	1.37	1.38	1.38	1.39
Total Power (mW)	7.41	7.17	6.77	7.09	7.05	152.92	143.42	149.01	137.44	140.21	140.97	131.93	122.67	120.01	142.68
IR-drop (mV)	62.54	14.13	14.32	14.25	14.35	55.95	23.79	24.09	24.02	25.99	170.02	17.88	18.93	18.03	18.58
T_{wns} (ns)	-0.196	-0.111	-0.094	-0.106	-0.061	-0.150	-0.126	-0.122	-0.125	-0.077	-0.194	-0.081	-0.092	-0.106	-0.011
T_{ins} (ns)	-4.06	-1.96	-2.04	-2.14	-1.09	-28.53	-25.25	-24.44	-25.06	-7.66	-1.756	-1.637	-1.579	-1.499	-1.139
Eff. Freq (GHz)	1.50	1.64	1.70	1.65	1.78	1.54	1.61	1.63	1.60	1.73	1.44	1.72	1.69	1.65	1.96
# $r_{\text{delta}} > 40\%$	3	8	8	8	2	3	16	16	16	4	7	17	17	17	2
# $r_{\text{bump}} > 30\%$	1	3	3	3	0	2	12	12	12	2	5	27	26	27	3

Table 3: Technology specifications and design rules.

BS	Width/Spacing(μm)	Pitch(μm)	FS	Width/Spacing(μm)	Pitch(μm)
BSM0	0.053 / 0.028	0.04	M0	0.020 / 0.020	0.04
BSM1	0.125 / 0.375	0.5	M1	0.037 / 0.020	0.06
BSM2	0.5 / 1.5	2	M2/3	0.020 / 0.020	0.04
BSM3	0.5 / 1.5	2	M4-M8	0.038 / 0.038	0.08

where $\hat{y}(\cdot)$ is the FM-predicted performance and T is the temperature that gradually cools down (e.g., $T_{t+1} = \eta T_t$, $\eta \in (0, 1)$). The optimization stops when no further improvement is observed or a preset iteration limit is reached. The final partitioning $\mathbf{x}^*$ is used for full routing and PPA evaluation.

This two-phase design balances exploration and exploitation: the training phase diversifies data for surrogate learning, while the inference phase leverages the surrogate to efficiently refine the final BSPDN-aware allocation.

3.4 Signal Integrity Mitigation

Based on post-detailed-routing analysis of our initial BSPDN and FSPDN designs, we observe notable differences in signal integrity behavior. As shown in Fig. 6, BSPDN exhibits 2–4 $\times$ higher voltage bumps and delay variations than FSPDN. The exponential decay trend in both metrics indicates that only a subset of nets is inherently more vulnerable to SI degradation. The red-shaded regions in Fig. 6 mark the performance gap between the two PDN schemes, with the red dashed lines denoting the 0.5 ratio safety threshold. In our methodology, nets with voltage bump or delta-delay ratios exceeding 40% are flagged as SI-critical. These nets are ranked by severity and explicitly reassigned to backside routing tracks, where stronger electromagnetic shielding can effectively suppress excessive interference, while preserving global routing balance and timing closure.

This SI optimization stage is designed to operate independently from the primary net selection in our flow, since SI degradation is driven more by routing congestion and net density than by timing metrics like WNS or TNS. Accordingly, we reserve backside routing resources for SI-critical nets, which proves more resource-efficient than extending BSPDN power rails into frontside metal layers.

4 Experimental Results

This section presents experiments on representative designs to validate the effectiveness of the proposed CSCO method. Both quantitative and qualitative results are reported.

4.1 Experiment Settings

Technology Specifications. nTSVs connect buried power rails (BPRs) to backside metal layers in our BSPDN [37]. BPRs reduce cell height by eliminating conventional VDD/VSS BEOL layers above the FEOL. Table 3 lists both metal stacks. **Backside-Aware Design Flow.** Commercial implementation and signoff tools do not natively support explicit backside metal layers. We therefore developed a custom Interconnect Technology file (ICT) defining a backside-optimized BSM1–BSM3 BEOL stack and corresponding QRC files for parasitic extraction. TCAD simulations yielded SPICE-compatible nTSV models with $C_{\text{nTSV}} = 0.444$ fF and $R_{\text{nTSV}} = 23 \Omega$.

The flow iterates frontside/backside implementation, PDN-aware backside routing, parasitic extraction and merge, and signoff timing. It feeds parasitics and timing back to backside routing without major changes to established flows.

Benchmark Designs. We evaluate SHA256 (172 clock nets and 7.8k signal nets), JPEG (3.5k clock nets and 180k signal nets) [1], and ARM Cortex-A7 (1.5k clock nets and 116k signal nets) [2], a power-efficient CPU for mobile and embedded systems. All designs are implemented in 7 nm technology

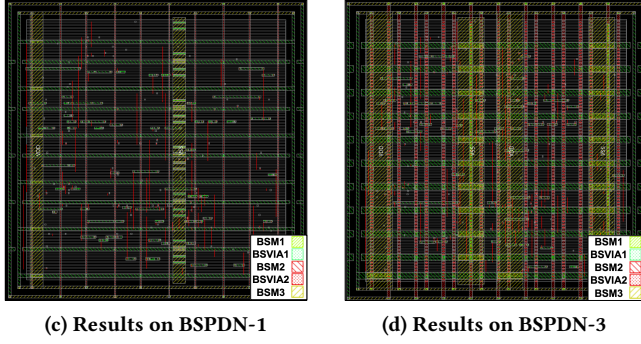
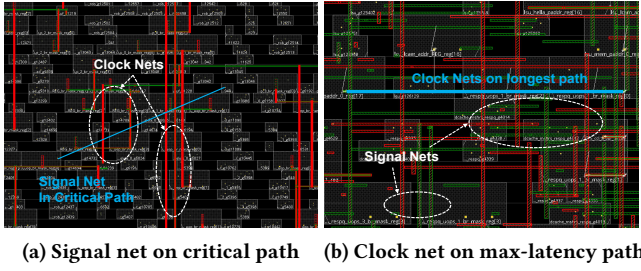


Figure 7: Routing critical nets through backside and CSCO results under various BSPDN configurations.

($V_{DD} = 0.75$ V) using Genus [4] and Innovus [3]. Power integrity is evaluated with an in-house IR-drop tool [8], and timing is signed off using PrimeTime SI [5].

Implementation. CSCO is implemented in Python 3.8 and evaluated on a Rocky Linux 8.10 cluster with two Xeon Platinum 8480+ CPUs (112 cores) and two NVIDIA L40S GPUs.

Hyper-parameters. We set the loss weight $\lambda = 0.1$ and latent embedding dimension $l = 12$. The FM is trained using MSE and SGD (batch size 32, learning rate 10^{-4} , and 5×10^3 update steps), with 20% of greedily selected nets randomly replaced for exploration. We collect 200 samples for SHA256 and 500 for both JPEG and ARM. During inference, simulated annealing starts at $T = 10^{-5}$, cools with $\eta = 0.995$, and stops at 10^{-6} . Clock/signal net budgets are fixed at 30%/70%.

4.2 PPA Evaluations

Beyond the metrics in Section 2.3, we report nTSV count, clock skew, wire length, total power, effective frequency, and IR-drop. Table 2 compares CSCO on three benchmarks against baselines, DAC24 [7], and DAC25 [17]. FSPDN+FR and BSPDN+FR use frontside routing under their respective PDNs; DAC24 applies GNN-based backside assignment, whereas DAC25 routes critical-path clocks on the backside.

CSCO achieves the best timing, improving WNS by up to 85% on ARM Cortex-A7 and over 60% on average, and TNS by 35–70%. These gains raise effective frequency to 1.78GHz

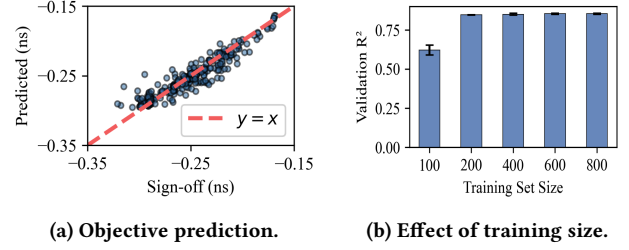


Figure 8: Surrogate model evaluation.

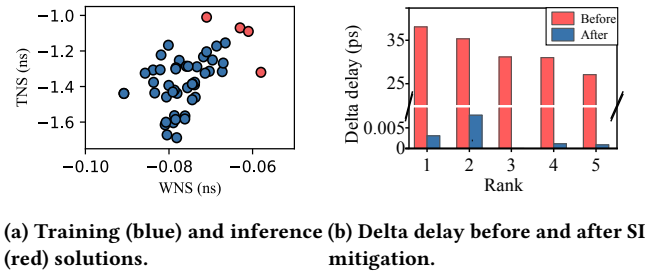


Figure 9: Solution quality and SI analysis.

for SHA256 (20% over FSPDN) and 1.96GHz for ARM (36.1% over the baseline). CSCO also uses fewer nTSVs than DAC25 with negligible area overhead (74.15% utilization for ARM), balancing performance and resource efficiency.

Visualization. Figs. 7(a) and 7(b) visualizes critical-path signal and clock nets. Frontside congestion limits their routing efficiency, whereas backside routing expands the solution space for timing optimization. Figs. 7(c) and 7(d) shows CSCO’s backside routing results on SHA256 under two BSPDN configurations, illustrating how available backside resources affect routing patterns and utilization. Fig. 9(a) compares heuristic and empirical-model solutions, showing that CSCO extends the Pareto frontier toward better solutions.

4.3 Surrogate Model

To understand how CSCO works, we evaluate the proposed surrogate FM model across multiple dimensions here.

Accuracy and Consistency. The surrogate FM in Section 3.2 is a fast proxy for sign-off timing. On 337 held-out SHA256 backside-routing allocations, it closely tracks the sign-off objective, achieving an R^2 of 0.858 and an RMSE of 0.0141 ns, small relative to the approximately 0.17 ns baseline. Strong correlation in the upper-right region of Fig. 8(a) supports search guidance toward better-timing solutions. It also preserves candidate rankings, with top- K overlaps of 100%, 90%, and 85% at $K = 5, 10$, and 20, respectively.

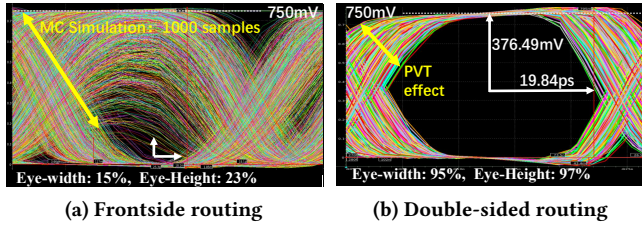


Figure 10: SI mitigation on the worst-case net: eye diagrams with (a) frontside and (b) double-sided routing.

Table 4: BSPDN power-rail configurations. Entries list BSM1/BSM2/BSM3 values; r and c denote resistance and capacitance per unit length.

BSPDN Config	BSPDN-1	BSPDN-2	BSPDN-3
Power Rail Width (μm)	0.10/0.40/1.0	0.20/0.60/2.0	0.35/0.80/3.0
Power Rail Pitch (μm)	2.1/1.8/10	2.1/1.8/10	2.1/1.8/10
Unit r ($\Omega/\mu\text{m}$)	4.20/0.21/0.016	1.86/0.14/0.008	1.10/0.10/0.005
Unit c (fF/ μm)	0.41/0.23/0.083	0.37/0.21/0.153	0.39/0.30/0.270

Training Robustness. We evaluate SHA256 validation R^2 across training-set sizes; Fig. 8(b) reports the three-run mean and standard deviation. At 200 samples, low variation confirms sampling and optimization robustness. The model converges and generalizes with a few hundred samples while reducing sign-off calls.

Efficiency. Training completes within 15 minutes on one GPU, and each backside routing assignment takes under 30 ms versus approximately 323.4 seconds for a PrimeTime sign-off iteration. The surrogate thus provides over 10,000 $\times$ faster objective evaluation. One model covers all BSPDN configurations without retraining, further reducing cost.

4.4 Signal Integrity Analysis

We count nets with $r_{\text{delay}} > 0.4$ or $r_{\text{bump}} > 0.3$ as SI violations. Table 2 shows that naive BSPDN+FR increases both violations over FSPDN across all benchmarks. On ARM Cortex-A7, the delay and bump counts increase from 7 to 17 and from 5 to 27. CSCO reduces the three-benchmark totals from 41 to 8 (80.5%) and from 42 to 5 (88.1%). On ARM, the counts fall to 2 and 3, showing that CSCO suppresses aggressor-victim coupling while preserving timing closure. Fig. 9(b) shows that the maximum delta delay among the top five critical nets decreases from approximately 30 ps to below 0.01 ps. In Figs. 10(a) and 10(b), CSCO-enabled double-sided routing improves the normalized eye width and height of the worst-case net from 15%/23% to 95%/97%, corresponding to an absolute eye width of 19.84 ps and eye height of 376.49 mV. These results demonstrate robust high-frequency operation without additional area overhead.

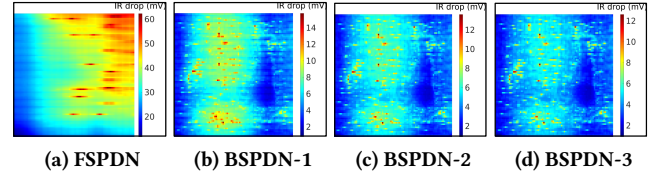


Figure 11: Static IR-drop maps for SHA256 under FSPDN and BSPDN configurations (area: $40 \times 40 \mu\text{m}^2$).

Table 5: The solutions given by CSCO on SHA256 with various BSPDN configurations.

Configs	IR-Drop (mV)	#nTSV	WNS (ns)	TNS (ns)	$r_{\text{delay}} > 0.4$	$r_{\text{bump}} > 0.3$
BSPDN-1	15.76	240	-0.060	-1.17	2	1
BSPDN-2	14.35	230	-0.061	-1.09	2	0
BSPDN-3	13.61	210	-0.067	-1.18	1	1

4.5 Joint BSPDN and Net Allocation

CSCO jointly optimizes BSPDN configurations and net allocation. Table 4 shows that progressively wider power stripes reduce line resistance at the cost of higher capacitance and less routing space. The IR-drop maps in Fig. 11 illustrate the trade-off between power integrity and routing congestion. Unlike methods requiring optimization or retraining for each PDN setting, CSCO reuses one surrogate across heterogeneous BSPDN layouts to explore the joint PDN-routing space and identify Pareto-efficient net-to-layer assignments. Table 5 shows a non-monotonic trade-off. BSPDN-3 minimizes IR-drop but slightly worsens timing from higher parasitics; BSPDN-1 achieves the best WNS with higher IR-drop. BSPDN-2 offers the best balance, combining low IR-drop, the best TNS, and stable WNS. Thus, CSCO captures configuration-dependent trade-offs and guides topology selection for design-specific PPA objectives.

5 Conclusion

This paper presents CSCO, a BSPDN-aware co-optimization framework that jointly determines PDN topology and dual-side clock/signal routing. By integrating PDN planning with backside-aware net allocation and leveraging data-driven surrogate modeling, CSCO efficiently explores the coupled design space without costly full-routing iterations. Unlike prior methods, CSCO accounts for amplified crosstalk introduced by removing frontside power shields and exploits BSPDN metals to protect SI-critical paths. Evaluations on realistic benchmark designs demonstrate that CSCO achieves significant PPA gains while satisfying power and signal integrity constraints, establishing a unified methodology for next-generation DTCO and 3D ICs.

References

- [1] 1999. *OpenCores: The reference community for Free and Open Source gateway IP cores*, <https://opencores.org/>.
- [2] 2011. ARM Cortex-A7 Processor, <https://www.arm.com/products/silicon-ip-cpu/cortex-a/cortex-a7>.
- [3] 2019. Cadence, "Innovus Reference Manual", https://www.cadence.com/en_US/home/tools/digital-design-and-signoff/soc-implementation-and-floorplanning/innovus-implementation-system.html.
- [4] 2021. Cadence, "Genus's user guide", https://www.cadence.com/en_US/home/tools/digital-design-and-signoff/synthesis/genus-synthesis-solution.html.
- [5] 2026. Synopsys, "Primitime User Guide", <https://www.synopsys.com/implementation-and-signoff/signoff/primitime.html>.
- [6] Erick Carvajal Barboza, Nishchal Shukla, Yiran Chen, et al. 2019. Machine learning-based pre-routing timing prediction with reduced pessimism. In *Proc. DAC*. 1–6.
- [7] Nesara Eranna Bethur, Pruek Vanna-Iampikul, Odysseas Zografos, et al. 2024. GNN-assisted Back-side Clock Routing Methodology for Advance Technologies. In *Proc. DAC*. 6 pages.
- [8] R Chen, M Lofrano, G Mirabelli, et al. 2022. Power, performance, area and thermal analysis of 2D and 3D ICs at A14 node designed with back-side power delivery network. In *Proc. IEDM*. IEEE, 23–4.
- [9] Rongmei Chen, Giuliano Sisto, Michele Stucchi, et al. 2022. Backside PDN and 2.5 D MIMCAP to double boost 2D and 3D ICs IR-drop beyond 2nm node. In *Proc. IEEE VLSI Symp*. IEEE, 429–430.
- [10] Vidya A. Chhabria, Ben Keller, Yanqing Zhang, and et al. 2022. XT-PRAGGMA: Crosstalk pessimism reduction achieved with GPU gate-level simulations and machine learning. In *Proc. MLCAD*. 63–69.
- [11] Suhyeong Choi, Carlo Gilardi, Paul Gutwin, et al. 2025. Omni 3D: BEOL-Compatible 3-D Logic With Omnipresent Power, Signal, and Clock. *IEEE TED* (2025).
- [12] Suhyeong Choi, Jinwook Jung, Andrew B Kahng, et al. 2023. PROBE3.0: a systematic framework for design-technology pathfinding with improved design enablement. *IEEE TCAD* 43, 4 (2023), 1218–1231.
- [13] Prashant Dubey, Odysseas Zografos, Nicolas Pantano, et al. 2024. Managing Crosstalk in Multi-GHz Front Side Clock for Back Side Power Enabled Sub-2nm 2D/3D ICs. In *Proc. ICICDT*. IEEE, 1–4.
- [14] Zizheng Guo, Mingjie Liu, Jiaqi Gu, et al. 2022. A timing engine inspired graph neural network model for pre-routing slack prediction. In *Proc. DAC*. 1207–1212.
- [15] Md Obaidul Hossen, Bharani Chava, Geert Van der Plas, et al. 2019. Power delivery network (PDN) modeling for backside-PDN configurations with buried power rails and *mu* TSVs. *IEEE TED* 67, 1 (2019), 11–17.
- [16] Fang-Yu Hsu, Tzu-Chuan Lin, Wai-Kei Mak, et al. 2024. A Bounding Box-based Net Partitioning Method for Double-sided Routing. In *Proc. GLSVLSI*. 397–402.
- [17] Xun Jiang, Haoran Lu, Yuxuan Zhao, et al. 2025. A Systematic Approach for Multi-objective Double-side Clock Tree Synthesis. In *Proc. DAC*. IEEE, 1–7.
- [18] Byung-Sung Kim, Subin Choi, Jung Han Lee, et al. 2024. Expanding design technology co-optimization potentials with back-side interconnect innovation. In *Proc. IEEE VLSI Symp*. IEEE, 1–2.
- [19] Joon-Seok Kim, Joonyun Kim, Dae-Jin Yang, et al. 2024. Addressing interconnect challenges for enhanced computing performance. *Science* 386, 6727 (2024), eadk6189.
- [20] J Lee, S Lee, Y Ahn, et al. 2025. A Novel Backside Signal Inter/Intra-Cell Routing Method Beyond Backside Power for Angstrom Nodes. In *Proc. IEEE VLSI Symp*. IEEE, 1–3.
- [21] Rongjian Liang, Zhiyao Xie, Jinwook Jung, and et al. 2020. Routing-Free Crosstalk Prediction. In *Proc. ICCAD*. 1–9.
- [22] Tzu-Chuan Lin, Fang-Yu Hsu, Wai-Kei Mak, et al. 2024. An effective netlist planning approach for double-sided signal routing. In *Proc. ASPDAC*. 288–293.
- [23] Fangzhou Liu, Guannan Guo, Yuyang Ye, et al. 2025. GraphCAD: Leveraging graph neural networks for accuracy prediction handling crosstalk-affected delays. In *Proc. ISPD*. 125–133.
- [24] Haoran Lu, Yandong Ge, Xun Jiang, et al. 2024. First experimental demonstration of self-aligned flip FET (FFET): A breakthrough stacked transistor technology with 2.5 T design, dual-side active and interconnects. In *Proc. IEEE VLSI Symp*. IEEE, 1–2.
- [25] Wenlong Lyu, Fan Yang, Changhao Yan, et al. 2018. Batch Bayesian optimization via multi-objective acquisition ensemble for automated analog circuit design. In *Proc. ICML*. PMLR, 3306–3314.
- [26] SS Teja Nibhanupudi, Sirish Oruganti, Rahul Mathur, et al. 2023. Buried Power Rails and Back-side Power Grids: Prospects and Challenges. In *Proc. DAC*. IEEE, 1–4.
- [27] Min Gyu Park, Amaan Rahman, and Sung Kyu Lim. 2025. BS-PDN-Last: Towards Optimal Power Delivery Network Design With Multifunctional Backside Metal Layers. In *Proc. DAC*. 1–6.
- [28] Sai Surya Kiran Pentapati, Kyungwook Chang, Vassilios Gerousis, et al. 2020. Pin-3D: A physical synthesis and post-layout optimization flow for heterogeneous monolithic 3D ICs. In *Proc. ICCAD*. 1–9.
- [29] Amaan Rahman, Hang Yang, Cong Hao, et al. 2025. Modeling and Design Methodology for Backside Integration of Voltage Converters. In *Proc. ISPD*. 242–250.
- [30] Steffen Rendle. 2010. Factorization machines. In *Proc. ICDM*. IEEE, 995–1000.
- [31] Bobak Shahriari, Kevin Swersky, Ziyu Wang, et al. 2015. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* 104, 1 (2015), 148–175.
- [32] Kumar Subramani, Rwik Sengupta, Matthew Loden, et al. 2025. Back-side Routing Enablement Considerations for Advanced Node GAA Devices. In *Proc. IEEE VLSI Symp*. IEEE, 1–3.
- [33] Ryo Tamura, Yuya Seki, Yuki Minamoto, et al. 2025. Black-box optimization using factorization and Ising machines. *arXiv:2507.18003 [cond-mat.stat-mech]* <https://arxiv.org/abs/2507.18003>
- [34] Che-Ping Tsai, Fang-Yu Hsu, Wai-Kei Mak, et al. 2024. An Effective ECO Methodology for Reducing Back-side Design Rule Violations in Double-sided Signal Routing. In *Proc. ICCAD*. 1–9.
- [35] Pruek Vanna-Iampikul, Hang Yang, Jungyoun Kwak, et al. 2024. Back-side Design Methodology for Power Delivery Network and Clock Routing. In *Proc. IEEE VLSI Symp*. 1–2.
- [36] A Veloso, B Vermeersch, R Chen, et al. 2023. Backside power delivery: game changer and key enabler of advanced logic scaling and new STCO opportunities. In *Proc. IEDM*. IEEE, 1–4.
- [37] Linqiu Wang, Feifan Xie, Jizhe Liu, et al. 2024. Power and Thermal Integrity Analysis of High Performance and Low Power CPUs at Sub-2nm Node Designed with Various Advanced Backside PDNs. In *Proc. IEDM*. 1–4.
- [38] Feifan Xie and Tiwei Wei. 2025. Exploring Efficient Thermal Management Solutions for Backside Power Delivery Network (BSPDN) Systems Using Multi-Scale Modeling. *IEEE TCPMT* (2025).
- [39] Zhiyao Xie, Yu-Hung Huang, Guan-Qi Fang, et al. 2018. RouteNet: Routability prediction for mixed-size designs using convolutional neural network. In *Proc. ICCAD*. 1–8.
- [40] S Yang, Pieter Schuddinck, M Garcia-Bardon, et al. 2023. PPA and scaling potential of backside power options in N2 and A14 nanosheet technology. In *Proc. IEEE VLSI Symp*. IEEE, 1–2.
- [41] Yun Zhou, SC Song, Halil Kükner, et al. 2024. Backside power delivery in high density and high performance context: IR-drop and block-level power-performance-area benefits. In *Proc. IEEE VLSI Symp*. IEEE, 1–2.