Enhancing System-Wide Power Integrity in 3D ICs with Power Gating

Hailang Wang and Emre Salman
Department of Electrical and Computer Engineering, Stony Brook University,
Stony Brook, NY 11794 USA
E-mail: {hailang.wang, emre.salman}@stonybrook.edu

Abstract— Power gating is a commonly used method to reduce subthreshold leakage current in nanoscale technologies. In through silicon via (TSV) based three-dimensional (3D) integrated circuits (ICs), power gating can significantly degrade system-wide power integrity since the decoupling capacitance associated with the power gated block/plane becomes ineffective for neighboring active planes, as demonstrated in this paper. A reconfigurable decoupling capacitor topology is investigated to alleviate this issue by exploiting the ability of via-last TSVs to bypass plane-level power networks when delivering the supply voltage. Reconfigurable decoupling capacitors placed within a plane can provide charge to neighboring planes even when the plane is power gated, thereby significantly reducing both RMS power supply noise (by up to 46%) and RMS power gating (in-rush current) noise (by up to 85%) at the expense of a slight increase in area (by 1.55%) and peak power consumption (by 1.36%).

Keywords— Power delivery, 3D IC, power gating

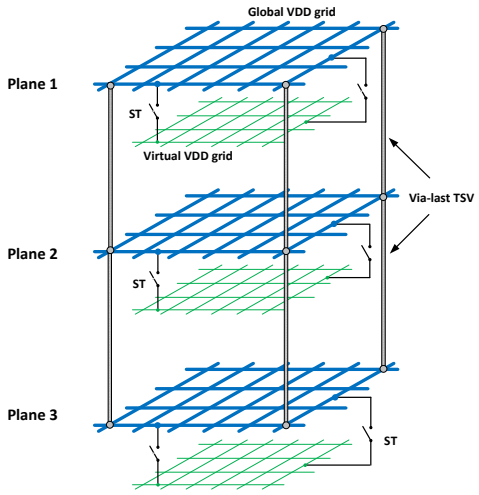

Fig. 1. Conceptual representation of a power distribution network for a three-plane 3D IC with via-last TSVs. Power gating is illustrated with global and virtual power grids and sleep transistors (ST).

I. Introduction

Through silicon via (TSV) based three-dimensional (3D) integrated circuits have emerged as a promising technology for both high performance and low power integrated circuits (ICs) [1]. 3D ICs alleviate the adverse effects of global interconnects, thereby enabling higher performance at lower power consumption while potentially lowering cost [1, 2]. A significant challenge in vertical 3D integration is the design of a reliable power distribution network. Existing research efforts have investigated TSV types, placement, optimization, and power grid architectures [3–5]. Decoupling capacitors in 3D ICs have also been considered [6,7]. However, the effect of *power gating* on the power integrity of 3D ICs has not received much attention. Power gating effectively reduces subthreshold leakage current by turning off the power delivery path of a circuit module when the module is not active [8]. For 3D systems, due to the characteristics of heterogeneous integration and high parallelism, the amount of nonswitching circuits can be significantly high. Thus, 3D ICs are expected to be heavily power gated to sufficiently reduce leakage power.

In [9], Todri *et al.* have investigated the effect of plane-level power gating on power integrity in 3D ICs. It has been demonstrated that the decoupling capacitance placed within a plane is highly effective in reducing the power supply noise of neighboring planes. The interdependence between decoupling capacitance and power gating, however, has not been considered.

In traditional decoupling capacitor topologies, the decoupling capacitance within a circuit block is connected to the local power grid (closer to the switching circuits) to effectively reduce supply noise [10, 11]. However, if the block (or an entire plane) is power gated, those capacitors are disconnected from the global power network and therefore cannot provide charge to the neighboring, potentially active blocks or planes.

In [12], Tong *et al.* have proposed reroutable decoupling capacitors in 2D ICs. The primary objective has been to relax the tight tradeoff between power gating noise and leakage power by connecting the decoupling capacitors to the global power grid when a block is power gated. Thus, these decoupling capacitors remain charged, significantly reducing the power gating noise at the expense of capacitor leakage current. However, in [12], the ability of reroutable capacitors in reducing the power supply noise of neighboring blocks has not been explored. This ability is facilitated particularly in 3D ICs due to shorter global interconnects. In this paper, reconfigurable capacitors are proposed to alleviate power supply noise of the neighboring active planes in 3D ICs. Two characteristics of via-last TSVs are exploited to increase the effective range of a de-

This research is supported in part by the National Science Foundation CAREER grant under No. CCF-1253715 and the Office of the Vice President for Research at Stony Brook University.

coupling capacitance: (1) low resistivity and (2) ability to bypass plane-level power network when delivering the power supply voltage.

The rest of the paper is organized as follows. The reconfigurable decoupling capacitor placement topology for 3D ICs with power gating is introduced in Section II. Simulation results are presented in Section III. Finally, the paper is concluded in Section IV.

II. Reconfigurable Decoupling capacitors in 3D ICs with Power Gating

Typical methods of implementing power gating in 3D ICs are introduced in Section II-A. The reconfigurable decoupling capacitor topology is presented in Section II-B.

A. Background

Similar to 2D planar technologies, sleep transistors with high threshold voltage are utilized to achieve power gating in 3D ICs. Depending upon the type of TSVs (via-first/middle or via-last), distributed or lumped power gating topologies have been proposed [13]. In the distributed topology (shown to be more appropriate for via-last TSVs), sleep transistors are distributed throughout the entire 3D stack whereas in lumped topology (shown to be more appropriate for via-first/middle TSVs), all of the sleep transistors are located at the topmost plane [13]. An example of a via-last TSV based 3D power network with distributed power gating topology is illustrated in Fig. 1. Note that via-last TSVs pass through the metal layers and connect the topmost metal layer on each plane. Also, note that in via-last TSV technology, the TSV resistance is significantly smaller as compared to via-first/middle technologies. These two characteristics of via-last TSVs facilitate the utilization of a decoupling capacitor within a plane to reduce the power supply noise of the neighboring planes (due to greater effective range).

B. Reconfiguration of the Decoupling Capacitors

Since system-wide power integrity is a critical challenge, effective utilization of intentional decoupling capacitance is crucial, even when power gating is adopted. This issue is exacerbated in 3D ICs due to the power grid impedances of the multiple planes and higher integration levels. In 2D ICs, it is relatively more difficult to utilize the capacitance within a power gated domain for the remaining, active regions due to longer global interconnects. In 3D ICs, however, due to reduced interconnect length and relatively low resistance of via-last TSVs, decoupling capacitance within a power gated plane can still be effective for the neighboring active planes. According to [9], the effective range of a decoupling capacitor exceeds single plane in 3D ICs with via-last TSVs, as also observed in this work. Thus, reconfigurable decoupling capacitor topology is an effective method to enhance power integrity in 3D ICs with power gating.

In the reconfigurable topology, two switches are introduced, as conceptually illustrated in Fig. 2. If a certain

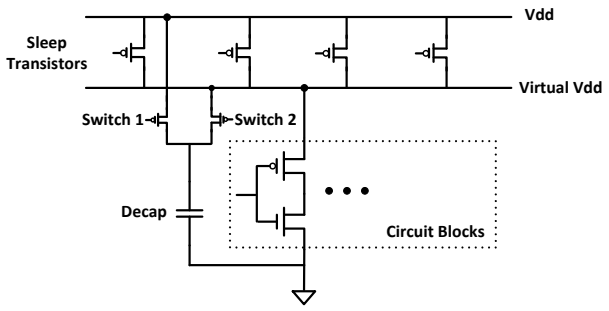

Fig. 2. Conceptual representation of the reconfigurable decoupling capacitor topology with power gating.

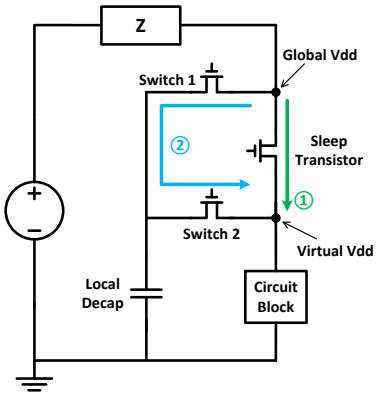

Fig. 3. Illustration of the additional resistive path between the global and virtual power networks formed by the reconfigurable switches.

plane is active, the decoupling capacitors on that plane are connected to the virtual V_{DD} grid through switch 2, thereby reducing the power supply noise on that plane. Alternatively, if the plane is power gated (sleep transistors are turned off), the decoupling capacitors are connected to the global V_{DD} grid, bypassing the sleep transistors. Thus, even if the plane is power gated, the decoupling capacitors remain effective for the active planes. The overhead of this topology includes the reconfigurable switches, metal resources to route the control signals, and higher leakage current if the capacitors are implemented as MOS capacitors, as quantified in this work.

Similar to sleep transistors [14], high-V_{th} MOS switches are used to minimize the voltage at the virtual V_{DD} grid when the plane is power gated (switch 1 is on and switch 2 is off). Note that the two reconfigurable switches form an additional path from global V_{DD} grid to virtual V_{DD} grid, as depicted in Fig. 3. Thus, the effective resistance of the sleep transistors and the effective resistance of the reconfigurable switches are in parallel, partially reducing the off-resistance between global and virtual V_{DD} grids. High-V_{th} switches are therefore required to maintain significant savings in the leakage current when the plane is power gated.

III. Simulation Results

A comprehensive case study is developed to investigate the benefits and tradeoffs of the proposed reconfigurable decoupling capacitor topology. The analysis setup is de-

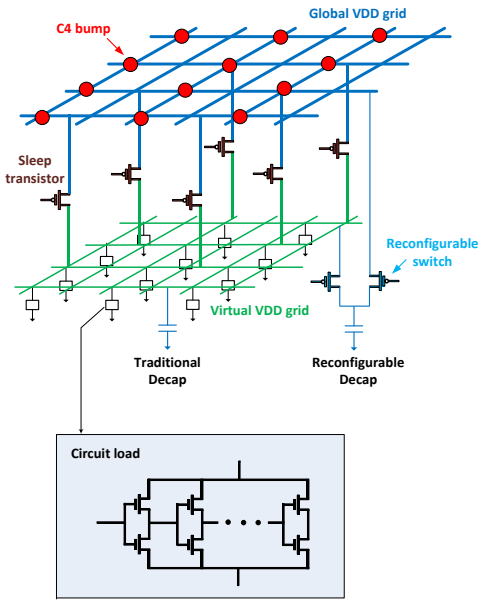

Fig. 4. Plane-level power network illustrating distributed sleep transistors, decoupling capacitors (traditional and proposed topologies), switching load circuits (gates with active devices), and the C4 bumps (for the top plane only).

scribed in Section III-A. Simulation results are presented in Section III-B where the proposed topology is compared with the traditional topology in terms of area overhead, power supply noise, power gating noise, and leakage current.

A. Simulation Setup

A highly distributed model of a power distribution network is developed for a three-plane 3D IC with via-last TSVs. A 45 nm CMOS technology with 10 metal layers in each plane is assumed [15]. A portion of the power network with an area of 1 mm by 1 mm is analyzed using HSPICE. Each plane consists of a global power network, virtual power network, distributed PMOS sleep transistors, distributed decoupling capacitance, and distributed switching load circuits, as depicted in Fig. 4.

Top two metal layers (9 and 10) on each plane are dedicated to global power distribution network with an interdigitated grid of 11×11.Metal layers 8 and 7 are used as the virtual power network represented by an interdigitated grid of 21×21. Power gating is achieved using a distributed method where the sleep transistors that control a plane are placed within that plane [13]. Primary physical characteristics of the 3D power grid are listed in Table I. The pitch and width of the metal lines are determined based upon the technology design rules [15] while also considering routing constraints.

Physical characteristics of the via-last TSVs, C4 bumps, and the package impedances are listed in Table II. A flip-chip package is assumed and modeled with a lumped resistance of 1 mΩ and inductance of 120 pH [16]. C4 bumps are regularly placed with a pitch of 200 μm over the 1 mm × 1 mm area [5]. Each C4 bump has a resistance of 5 mΩ and inductance of 200 pH. Clustered via-last TSVs are distributed throughout the area as a 10×10 array and connect the global power grid on each plane. Each TSV cluster consists of nine TSVs. A single via-last TSV is modeled as an *RLC* circuit [17] with the impedance values listed in Table II.

TABLE I

PRIMARY PHYSICAL CHARACTERISTICS OF THE GLOBAL AND LOCAL POWER GRIDS [15].

Parameters		Values
Global grid (Metal 10 & 9)	Pitch	45 μm
	Width	40 μm
	Resistivity (ohm/sq)	0.03
Local grid (Metal 8 & 7)	Pitch	23.5 μm
	Width	20 μm
	Resistivity (ohm/sq)	0.075

TABLE II

PACKAGE, TSV, AND C4 BUMP PARASITIC IMPEDANCES [5, 16].

Parameters	Values
Lumped package resistance R_{package}	1 mΩ
Lumped package inductance L_{package}	120 pH
Single C4 bump resistance R_{C4}	5 mΩ
Single C4 bump inductance L_{C4}	200 pH
Single via-last TSV resistance R_{tsv}	20 mΩ
Single via-last TSV capacitance C_{tsv}	283 fF
Single via-last TSV inductance L_{tsv}	35 pH

As opposed to using piecewise linear (PWL) current sources to model the switching load circuit (typical practice in existing work [8]), gates with active devices are used since power gating is considered. Similar to [18], inverter pairs with varying size are used to model the switching load circuit. The overall area is divided into 30 segments and a switching circuit is connected to each segment to consider the spatial heterogeneity of the current loads. The spatial load current distribution is determined based on [19], which produces a peak power density of approximately 40 W/cm^2 [20]. As an example, the load current distribution of the top plane is illustrated in Fig. 5 where the peak current for each block is indicated. The decoupling capacitors are implemented as MOS capacitors and distributed throughout the entire die area based on the spatial power supply noise distribution.

B. Results on Power Integrity and Overhead

The efficacy of the proposed decoupling capacitor placement topology is demonstrated by comparing the methodology with the traditional topology. Three power gating scenarios are considered:

- **Scenario 1**: All of the three planes are active, representing the greatest workload.
- **Scenario 2**: The top and bottom planes are active, while the middle plane is power gated.
- **Scenario 3**: Only the bottom plane is active, while the middle and top planes are power gated.

B.1 Area Overhead

Area overhead is listed in Table III. For both topologies, the decoupling capacitors are sized to ensure that the

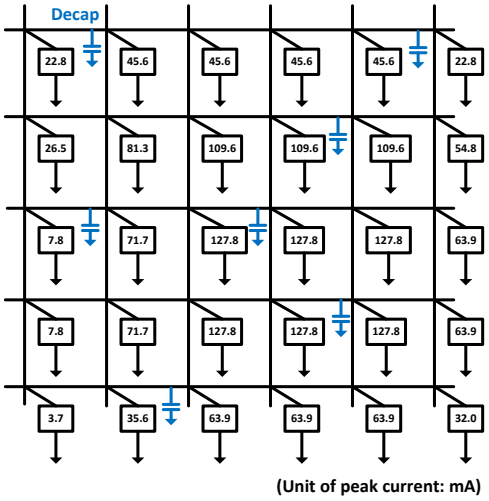

Fig. 5. Current distribution within the top plane based on [19]. The numbers refer to the peak current drawn by the digital gates at each node.

TABLE III

COMPARISON OF THE PHYSICAL AREA OVERHEAD OF THE TRADITIONAL AND RECONFIGURABLE TOPOLOGIES.

	Sleep Transistors	Decap	Switch	Area
Traditional	36.3 mm	2.87 nF	N/A	6.70%
Reconfigurable	36.3 mm	3.20 nF	160 mm	8.25%

worst case power supply noise is within 5% of the V_{DD} (50 mV) in scenario 1. Note that the size of the decoupling capacitors, sleep transistors, and switches (for the reconfigurable topology) in Table III refers to per plane. In the proposed topology, the area overhead increases from 6.70% to 8.25% (by only 1.55%) due to reconfigurable switches and higher decoupling capacitance required to compensate for the shielding effect of the reconfigurable switches. Note that all of the decoupling capacitors are implemented as MOS capacitors in 45 nm technology with an oxide thickness of 1 nm [15]. This slight increase in the physical area significantly enhances power integrity when one or more planes are power gated, as described in the following subsections.

B.2 Power Supply Noise

Power supply noise results are listed in Table IV for each scenario. Note that power supply noise is observed in the bottom plane. As listed in this table, when some of the planes are power gated (scenarios 2 and 3), the reconfigurable topology achieves less power supply noise by exploiting the capacitors of the power gated plane(s). For example, in scenario 3 where two planes are power gated, the reduction in peak noise is approximately 27% whereas the reduction in RMS noise is approximately 46%. Note that in the traditional topology, the peak noise in scenario 3 exceeds 50 mV despite the reduction in overall switching current due to power gating. This characteristic is due to less decoupling in the power network since the decoupling capacitors in the power gated planes cannot behave as charge reservoirs for the remaining, active planes. This

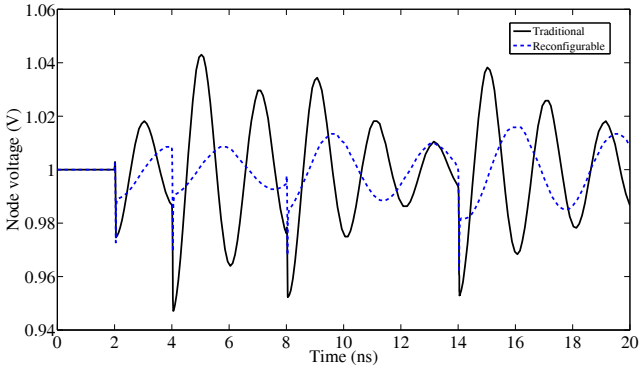

Fig. 6. Transient behavior of voltage noise at a specific node within the bottom plane for each topology for scenario 3.

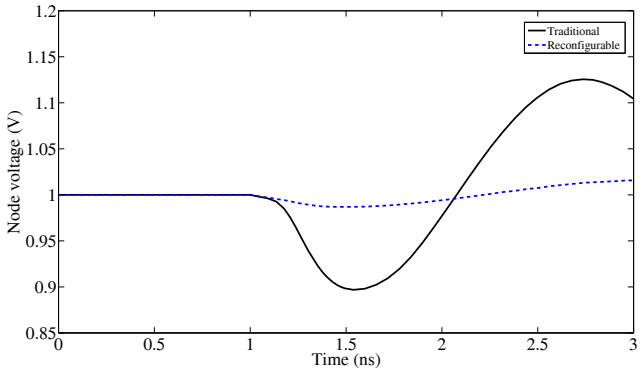

Fig. 7. Transient behavior of power gating noise at a specific node within the bottom plane for each topology for scenario 3.

observation justifies the need for reconfigurable capacitors. Transient behavior of voltage noise at a specific node within the bottom plane is also depicted in Fig. 6 for scenario 3, demonstrating the reduction in peak and RMS noise.

B.3 Power Gating Noise

To investigate power gating noise, the power gated middle plane transitions from inactive to active state in scenarios 2 and 3, and the voltage fluctuation due to in-rush current during the wake-up process is analyzed. Peak power gating noise is observed at the bottom plane. Results are listed in Table IV. The reconfigurable topology achieves more than 80% reduction in peak and RMS power gating noise. In traditional topology, a significant amount of in-rush current flows not only for the activated circuit, but also to charge the associated decoupling capacitors. Alternatively, in the reconfigurable topology, the decoupling capacitors are connected to the global V_{DD} grid when the plane is power gated. Thus, even if the plane is power gated, these capacitors remain charged (significantly reducing in-rush current) and can behave as charge reservoir once the plane transitions to active state. The transient behavior of the power gating noise is illustrated in Fig. 7 for scenario 3 where the transition happens at 1 ns.

TABLE IV

Power supply noise and power gating noise obtained from each scenario and noise reduction achieved by proposed topology.

	Power status			Power supply noise (mV)						Power gating noise (mV)					
	Top	Mid.	Btm.	Traditional		Reconfigurable				Traditional		Reconfigurable			
				Peak	RMS	Peak	Redtn.	RMS	Redtn.	Peak	RMS	Peak	Redtn.	RMS	Redtn.
Scenario 1	on	on	on	50	30.27	50	N/A	28.33	6.4%	N/A		N/A			
Scenario 2	on	off	on	48.94	23.41	42.90	12.3%	16.79	28.3%	90.73	72.83	15.55	82.8%	11.46	84.2%
Scenario 3	off	off	on	52.86	17.53	38.34	27.4%	9.39	46.4%	103.2	78.76	15.66	84.8%	11.57	85.3%

TABLE V

Overall Average Power Consumption (when all of the decoupling capacitors are implemented as MOS capacitors)

	Traditional	Reconfigurable	
	Power (mW)	Power (mW)	Overhead (%)
Scenario 1	26.45	26.81	1.36%
Scenario 2	17.57	18.07	3.41%
Scenario 3	8.89	9.41	5.85%

B.4 **Power Overhead**

To quantify the power overhead of the reconfigurable topology, both topologies are simulated for each scenario and the overall average power consumption is determined. All of the decoupling capacitors are implemented as MOS capacitors to consider MOS-C leakage. Results are listed in Table V. The smallest overhead occurs in scenario 1 where all of the planes are active. This overhead is due to increased capacitance and switches (for the reconfigurable topology only). Power overhead increases to 5.85% in scenario 3 due to the leakage current of MOS capacitors that are connected to the global grid. The absolute power is, however, significantly smaller in this scenario compared with the peak power in scenario 1. Thus, the maximum power increases by only 1.36%. Also note that this power overhead can be further reduced if low leakage decoupling capacitor implementations are adopted such as metal-insulator-metal (MIM) capacitors.

IV. **Conclusions**

In 3D ICs with power gating, traditional decoupling capacitors within a power gated block/plane are disconnected from the global grid, and therefore are ineffective for neighboring active blocks/planes. Reconfigurable capacitors have been proposed to alleviate this issue and significantly enhance system-wide power integrity. Analysis results demonstrate up to 46% and 85% reduction in, respectively, RMS power supply and power gating noise at the expense of 1.55% increase in physical area and 1.36% increase in peak power consumption.

References

[1] V. F. Pavlidis and E. G. Friedman, *Three-dimensional Integrated Circuit Design*. Morgan Kaufmann, 2010.

[2] E. Salman and E. G. Friedman, *High Performance Integrated Circuit Design*. McGraw-Hill, 2012.

[3] G. Huang, M. Bakir, A. Naeemi, H. Chen, and J. D. Meindl, "Power Delivery for 3D Chip Stacks: Physical Modeling and Design Implication," in *Proceedings of the IEEE Electrical Performance of Electronic Packaging*, October 2007, pp. 205–208.

[4] S. M. Satheesh and E. Salman, "Power Distribution in TSV Based 3D Processor-Memory Stacks," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 4, pp. 692–703, December 2012.

[5] M. Healy and S.-K. Lim, "Distributed TSV Topology for 3-D Power-Supply Networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 20, no. 11, pp. 2066–2079, November 2012.

[6] P. Zhou, K. Sridharan, and S. Sapatnekar, "Optimizing Decoupling Capacitors in 3D Circuits for Power Grid Integrity," *IEEE Design Test of Computers*, vol. 26, no. 5, pp. 15–25, September 2009.

[7] K. Kim, J. S. Pak, H. Lee, and J. Kim, "Effects of On-chip Decoupling Capacitors and Silicon Substrate on Power Distribution Networks in TSV-based 3D-ICs," in *Proceedings of the IEEE Electronic Components and Technology Conference*, 2012, pp. 690–697.

[8] H. Jiang, M. Marek-Sadowska, and S. Nassif, "Benefits and Costs of Power-gating Technique," in *Proceedings of IEEE International Conference on Computer Design*, October 2005, pp. 559 – 566.

[9] A. Todri, S. Kundu, P. Girard, A. Bosio, L. Dilillo, and A. Virazel, "A Study of Tapered 3-D TSVs for Power and Thermal Integrity," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 2, pp. 306–319, February 2013.

[10] M. Popovich, M. Sotman, A. Kolodny, and E. G. Friedman, "Effective Radii of On-Chip Decoupling Capacitors," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 16, no. 7, pp. 894–907, July 2008.

[11] E. Salman, E. Friedman, R. Secareanu, and O. Hartin, "Worst Case Power/Ground Noise Estimation Using an Equivalent Transition Time for Resonance," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 5, pp. 997–1004, May 2009.

[12] T. Xu, P. Li, and B. Yan, "Decoupling for Power Gating: Sources of Power Noise and Design Strategies," in *Proceedings of the ACM/IEEE Design Automation Conference*, June 2011, pp. 1002 –1007.

[13] H. Wang and E. Salman, "Power Gating Methodologies in TSV Based 3D Integrated Circuits," in *Proceedings of the ACM/IEEE Great Lakes Symposium on VLSI*, May 2013, pp. 327–328.

[14] D.-S. Chiou, S.-H. Chen, and S.-C. Chang, "Sleep Transistor Sizing for Leakage Power Minimization Considering Charge Balancing," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 17, no. 9, pp. 1330–1334, 2009.

[15] "FreePDK45." [Online]. Available: http://www.eda.ncsu.edu/wiki /FreePDK45:Contents

[16] M. Gupta, J. Oatley, R. Joseph, G.-Y. Wei, and D. Brooks, "Understanding Voltage Variations in Chip Multiprocessors using a Distributed Power-Delivery Network," in *Proceedings of the Design, Automation Test in Europe Conference Exhibition*, April 2007, pp. 1 –6.

[17] I. Savidis and E. G. Friedman, "Closed-Form Expressions of 3-D Via Resistance, Inductance, and Capacitance," *IEEE Transactions on Electron Devices*, vol. 56, no. 9, pp. 1873–1881, September 2009.

[18] X. Zhang, T. Tong, S. Kanev, S. Lee, G.-Y. Wei, and D. Brooks, "Characterizing and Evaluating Voltage Noise in Multi-Core Near-Threshold Processors," in *Proceedings of International Symposium on Low Power Electronics and Design*, September 2013, pp. 82–87.

[19] Q. Zhu, *Power Distribution Network Design for VLSI*. Wiley, 2004.

[20] H. Wei, T. Wu, D. Sekar, B. Cronquist, R. Pease, and S. Mitra, "Cooling Three-dimensional Integrated Circuits Using Power Delivery Networks," in *Proceedings of the IEEE International Electron Devices Meeting*, December 2012, pp. 14.2.1–14.2.4.