INTEGRATION the VLSI journal 61 (2018) 11–19

Contents lists available at ScienceDirect

INTEGRATION, the VLSI journal

journal homepage: www.elsevier.com/locate/vlsi

Exploratory design of on-chip power delivery for 14, 10, and 7 nm and beyond FinFET ICs☆

Kan Xu[a,*], Ravi Patel[a], Praveen Raghavan[b], Eby G. Friedman[a]

[a] Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY 14627, United States
[b] IMEC Research, 3001 Leuven, Belgium

ARTICLE INFO

Keywords:
On-chip power noise breakdown
Power delivery network
Exploratory PDN modelling
Advanced FinFET technologies
Graphene interconnect
Power noise suppression
Early stage power noise assessment

ABSTRACT

An exploratory modeling methodology is presented for estimating power noise in advanced technology nodes. The models are evaluated for 14, 10, and 7 nm FinFET technologies to assess the impact on performance. The power noise is composed of three parts, noise related to the global power grids, via stacks, and local power rails, based on the hierarchical nature of power distribution networks. In 14 nm technology, the global power noise dominates the total power noise. The power noise is lower and more evenly distributed in 10 nm technology. 7 nm technology is shown to be more sensitive to local power noise. To decrease the global power noise, extra metal layers are added to the global power grid. A 75% reduction in global power noise is observed in 14 nm technology. Stripes between local track rails are evaluated to reduce the local power noise, exhibiting up to 57% improvement in local power noise at the 7 nm technology node. As a promising alternative material for power network interconnects, few layer graphene is shown to exhibit good potential for reducing local power noise. The effects of different scaling scenarios of the local power rails on power noise are also discussed.

1. Introduction

The increasing demand for high density, high performance integrated circuits leads to aggressive technology scaling, enabling billions of transistors [1]. Due to the area and leakage current advantages as compared to planar CMOS, FinFETs have become the standard CMOS structure as technology is scaled below the 22 nm technology node [2]. While significant research effort is focused on deeply scaled transistors and emerging technologies, the RC interconnect impedance is challenging performance improvements brought by technology scaling. The parasitic capacitance of the local metal lines is less due to the adoption of low-k dielectrics and air gap interconnects [3]. The significantly increasing resistance of the local interconnects has however become the dominant limitation to performance improvements despite faster devices and greater levels of integration [4]. Scaling the cross sectional area of the local interconnects however quadractically increases the resistance. The resistivity of copper, used in traditional on-chip interconnects, sharply increases as the metal line pitch decreases [5], as illustrated in Fig. 1. In this case, the local power network is also highly resistive, leading to significant on-chip power noise.

Replacing copper interconnect with lower resistivity material interconnect is one way to reduce the effects of the "resistivity wall." Silver is one of these materials whose bulk level resistivity is lower than copper. Due to the excellent conductivity of both heat and electricity, and the negative temperature coefficient of carbon-based material graphene, few layer graphene (FLG) and graphene nanoribbons (GNRs) have been considered as an alternative material for on-chip interconnects [6,7]. Graphene material has been listed in the technology roadmap from ITRS 2015 [4] and many industrial research centers [8]. The thin film resistivity of three materials, silver, FLG, and GNRs, has been investigated, respectively, in [9–11]. A comparison of the resistivity of different materials with interconnect width scaling is also illustrated in Fig. 1. The thin film resistivity of silver increases significantly at 50 nm, and eventually becomes larger than copper as the metal line pitch is scaled to 10 nm. By intercalating FLG with ferric chloride, a sheet resistance of $8.8\Omega/\square$ has been reported [11]. Based on the thickness of five layer graphene, the resistivity of FLG is lower than copper, particularly when the metal line pitch is small, making FLG a promising material for a highly resistive local power network. GNR exhibits a higher resistivity comparable to copper when the metal line pitch is small (from 40 nm to 10 nm).

☆ This research is supported in part by the National Science Foundation under Grant Nos. CCF-1329374, CCF-1526466, and CCF-1716091, IARPA under Grant No. W911NF-14-C-0089, AIM Photonics under Award No. 059447-007, the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI), and by grants from Cisco Systems, Qualcomm, and OeC.

* Corresponding author.
E-mail addresses: kxu8@ece.rochester.edu (K. Xu), rapatel@ece.rochester.edu (R. Patel), Praveen.Raghavan@imec.be (P. Raghavan), friedman@ece.rochester.edu (E.G. Friedman).

https://doi.org/10.1016/j.vlsi.2017.10.007
Received 16 December 2016; Received in revised form 5 October 2017; Accepted 15 October 2017
Available online 09 November 2017

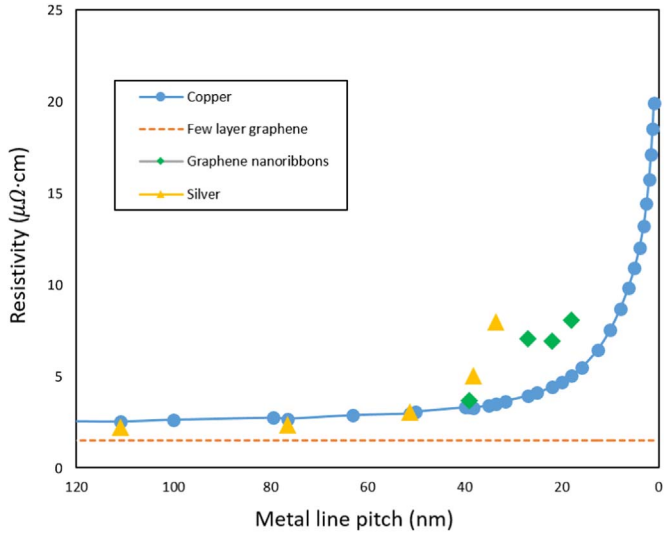


Fig. 1. Interconnect resistivity of different materials versus line width.

The "resistivity wall" phenomenon also leads to a more challenging power network design process due to the significantly resistive local power and ground rails. Reliable and energy efficient power distribution networks are necessary in high performance computing systems [12]. Decreasing supply voltages lead to smaller noise margins. Higher current densities and clock frequencies increase both resistive and inductive power supply noise. Moreover, a primary source of power noise is due to the highly resistive local power metal lines and vias between adjacent metal layers in advanced technology nodes. Poorly or overdesigned power networks either damage the reliability or decrease the performance of integrated circuits. Early assessment of the effects of the structure and material of the power networks supports tradeoffs among power noise, performance, and technology choice.

The rest of this paper is organized as follows. The structure of a typical standard cell based power distribution network is presented in Section 2. A modeling approach is discussed in Section 3. The components of power noise in advanced technology nodes is described in Section 4. Power noise suppression methods are presented in Section 5, followed by some conclusions in Section 6.

2. Standard cell based power network

The structure and impedance characteristics of power grids are presented in Section 2.1. The topology of a standard cell circuit influences the design of the power network, and therefore an overview of the structure is provided in Section 2.2.

2.1. Hierarchy of power grids

The resistance of the power metal lines is affected by the structure of the power grids. An on-chip power grid is a hierarchical structure consisting of a global interdigitated mesh, local power and ground rails, and a via stack connecting the global power grid to the local power rails, as illustrated in Fig. 2. A typical global power grid for high performance ICs uses two layers of orthogonal metal lines to form a mesh structure, as illustrated in Fig. 2(a). Adding global metal layers decreases the grid impedance, reducing power noise in the global power grid. The total number of on-chip metal layers is however limited by the technology. A mesh structure increases the reliability and robustness of a power network due to the multiple redundant paths. The mesh structure also reduces the resistance and parasitic capacitance of the power grids. Each metal layer in a mesh consists of parallel P/G pairs separated from adjacent pairs by tens of micrometers [4]. The pitch of each adjacent P/G pair is a design tradeoff between the

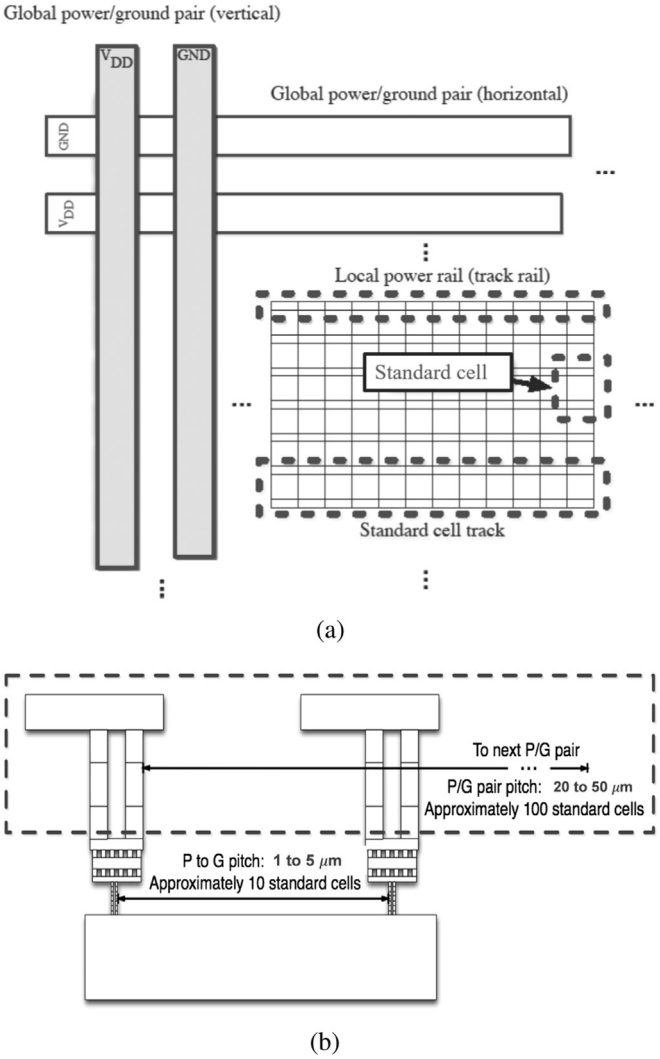


Fig. 2. Topology of a standard cell based power network: a) planar view, b) profile view.

power distribution network and signal/clock routing. The impedance of the global power network typically exhibits low resistance and significant inductance due to the mesh structure. As the current density and clock frequency increases with technology scaling, inductive $L\ di/dt$ noise becomes comparable to the resistive noise [13,14]. The power noise contributed by the global power grid should therefore be carefully evaluated to satisfy the strict power noise requirements in advanced technology nodes.

2.2. Standard cell based power rails

An individual standard cell track is structured as a row with a substrate region patterned between the power and ground rails, as illustrated in Fig. 2(a). Gates within a cell library are structured to fit within a constant height track with transistors patterned within the substrate. The height of a standard cell is typically controlled by lithographic limits introduced by double and quadruple patterning processes [15]. Standard cell gates are mirrored to ensure that two tracks share a common power rail, doubling the effective current load on the line. After the gates are placed, the interconnections are routed among the internal gates, constraining the available metal resources. The power rail impedances are dominated by the metal resistance and decoupling capacitance. On-chip power noise is caused by current switching on the track rails with the greatest contribution arising from the clocked gates and buffers [16]. Most notably, local power noise is

contributed by the *IR* drop within the power rails when multiple loads simultaneously switch. Reducing the resistance of the local power network is therefore an effective approach to mitigate resistive power noise. Early impedance characterization and power noise analysis can therefore be used to evaluate different metallization schemes and material alternatives in advanced technology nodes. The local power rails are typically not connected to each other to alleviate routing congestion in local metal layers. The global power grid is connected to the local power rails by a via stack, as illustrated in Fig. 2(b). The size and resistance of these vias are determined by the overlap area between metal lines and the thickness of the metallic barrier. The via is assumed to be cylinder shaped with a layer of metallic barrier, where the diameter is the same as the width of the adjacent power line. As technology is scaled, notably, the resistance of the via increases significantly due to the smaller cross sectional area and highly resistive metallic barrier of the via. The impedance characteristics of the on-chip power network affect the power noise generated in the three different parts of a power network.

3. Circuit models

The overall grid model consists of a load model, a local rail model, and a global mesh model, as illustrated in Fig. 3. Due to the symmetric characteristics of power distribution networks, only the V_{DD} portion of the power network is illustrated in Fig. 3. The digital load is modeled as a current source. The local power rail is modeled as a system composed of distributed resistors and capacitors. The global grid is modeled as an interdigitated mesh with the parameters described in [17]. The mesh size is based on the space between the pads. The model considers the physical area, supply current, and stage delay for each technology node (14, 10, and 7 nm). The load models, track rail, and stripes across the power rails are discussed in the following sections.

3.1. Load model

The peak power noise is dependent on the clock network [18]. The load model is based on the current demands of a register and adjacent gates within a standard cell track. A model of an interdigitated power and ground distribution network is discussed in [19]; however, only a global power network is considered. On-chip power noise in a high performance system-on-chip based IC is evaluated in [20]. A lumped model is utilized where the load is modeled at the block level. A distributed on-die power grid model is introduced in [21], where the on-die power noise is dependent on the microarchitecture and current profile within different blocks. An individual load on a track rail is modeled as a current source with a triangular load characteristic [22,23].

Those gates are spatially adjacent to the register and are likely to switch at approximately the same time as the register, thereby contributing to the local current. If an adjacent gate at the load switches before the track rail is recharged to the supply voltage, the magnitude of the noise increases [24]. If the gate does not switch before the voltage is restored to V_{DD}, the gate does not contribute to the peak noise [25,26]. Recharging determines the noise window (t_{window}) during which the loads that switch within the window are summed and the gates that switch outside of the window are ignored. The noise

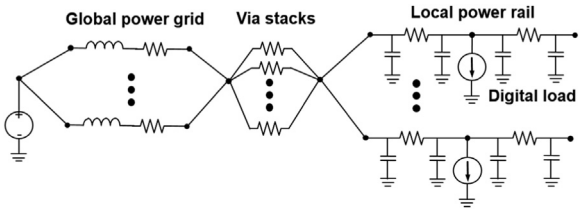


Fig. 3. Model of power network.

window, which determines the recharge time of a track rail, is approximated by three *RC* time constants,

$$t_{window} \approx 3\frac{N_{cell}^2}{4}R_{cell}(C_{cell} + C_{decap}), \tag{1}$$

where N_{cell} is the number of cells between each P/G pair, R_{cell} and C_{cell} are, respectively, the resistance and capacitance of the track rail within a standard cell, and C_{decap} is the decoupling capacitance per cell. No additional decoupling capacitors are considered in this work. The placement and optimization of decoupling capacitors have been investigated in [24,27].

Only those adjacent logic gates that switch within the noise window contributes to the peak power noise. The delay of the adjacent gates is approximated by the delay of an inverter. The load current is

$$I_{Load} = \frac{\alpha 2 t_{window}}{t_{inv1}} I_{inv1} + 2I_{invd4}, \tag{2}$$

where t_{window} is the noise window, t_{inv1} and I_{inv1} are, respectively, the delay and peak current of a 1X inverter, I_{inv4} is the peak current of a 4X inverter, and α is the switching factor of the circuit. Note that the intention of the load model is not to precisely emulate billions of load changes across the entire power network, but rather to mimic the peak power noise under realistic conditions when detailed block or load information is not available.

3.2. Rail model

Each local rail is modeled as a distributed resistor-capacitor with multiple loads, with the length of the rail determined by the space between two P/G pairs in the global power network. At least one load is placed at the center of the rail to model a single register assuming the worst case position. The number of loads and the space between loads are determined by the target clock frequency. An individual logic gate is modeled with an inverter delay (t_{inv}) where the logic depth (D) at a target frequency (f_{clock}) is

$$D = \frac{1}{f_{clock}\, t_{inv}(1 + U)}, \tag{3}$$

where U is the delay uncertainty. The logic depth is the number of gates between adjacent loads on a rail. The width of an inverter is used to estimate the size of a standard cell. The physical distance between loads on a local rail is therefore known. Based on this assumption, the total number of active loads and the impedance between each active load can be estimated. The logic depth is also used to determine the decoupling capacitance,

$$C_{decap} = C_{gate}(1 - \beta)D, \tag{4}$$

where C_{gate} is the gate capacitance of an inverter, and β is the fill factor of the standard cell layout. The fill factor, the fraction of silicon area occupied by the standard cells, is a common metric for characterizing the efficiency of standard cell circuits [28].

3.3. Striping of power rail

Each track rail is typically distinct. Recently, however, low impedance connections between adjacent track rails have been used to reduce the local rail resistance and any associated power noise, as illustrated in Fig. 4(a). These connections between the power and ground rails, called stripes, ensure that loads on the adjacent rails interact. For any interaction, however, the worst case power noise of a single local power rail (described as local power noise in the following section) occurs when the power rail is not connected with striping. The greatest reduction in power noise from striping occurs when the adjacent rails are not affected by simultaneously switching signals. These two conditions, therefore, bound the noise generated by a circuit.

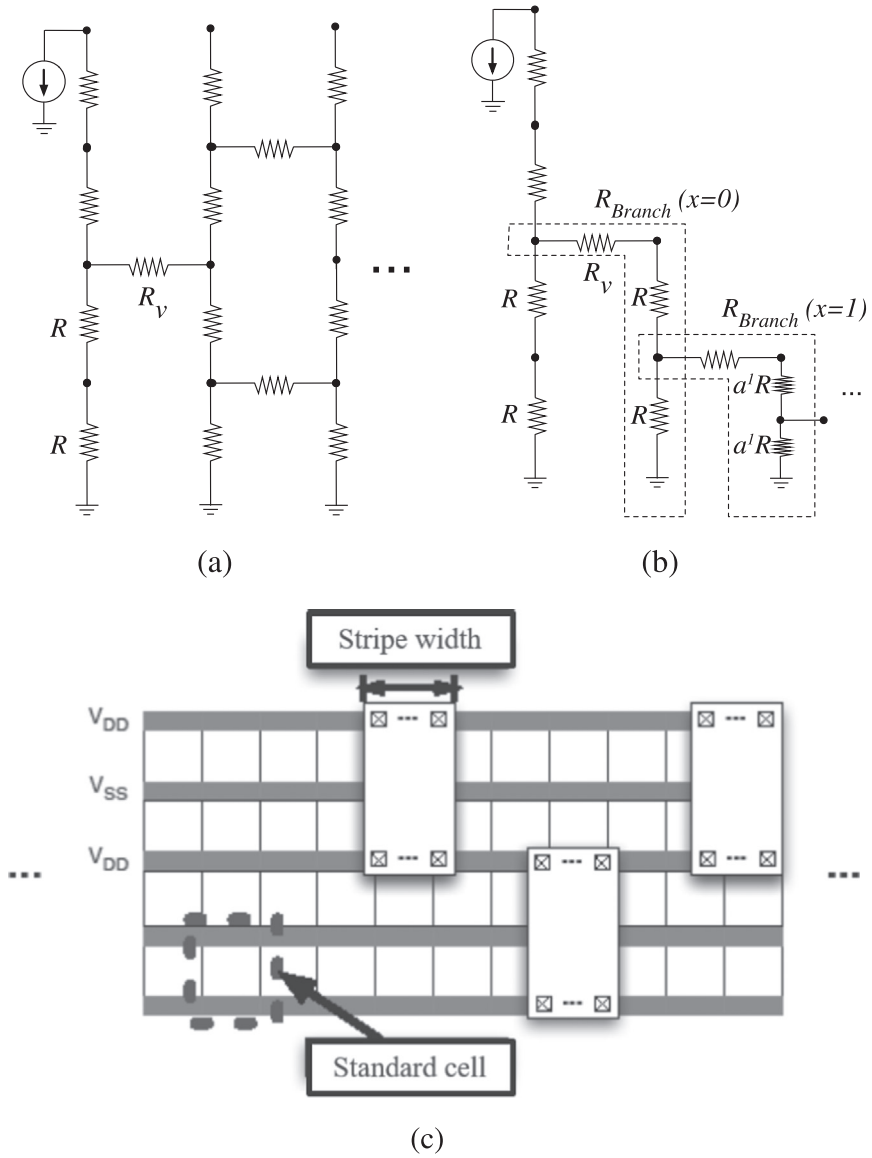


Fig. 4. Circuit models and physical structure of striping between the local power rails. a) comprehensive circuit model, b) R_{branch} approximated circuit model, and c) physical structure of a stripe.

The number of interacting rails is determined by approximating a set of rails as a resistive tree, as illustrated in Fig. 4(b). The resistance from the center load to the edge of the track rail is

$$R_{branch}(x) = R_v + a^x{*}R + \left(\frac{1}{a^x{*}R} + \frac{1}{R_{branch}(x+1)}\right)^{-1}, \quad (5)$$

where R_v is the resistance of a stripe, x is the number of additional branches, and a is the scaling factor of the resistance. As x increases, the error decreases. Note that (5) is used to estimate the maximum number of rails that minimizes the error. A distributed resistance across the rail is included in the model.

The proposed power methodology produces a general circuit model to evaluate peak power noise during the early exploratory design stage when floorplan and placement information is unavailable. This power network model is not intended to be integrated within a power network synthesis and optimization flow [29–31] or to compete with fast simulation algorithms within power network solvers which support many billions of nodes [32,33].

4. Evaluation of power noise

The model has been evaluated for power networks in 14 (N14), 10 (N10), and 7 nm (N07) CMOS FinFET technologies with a clock frequency ranging from DC to 5 GHz. The global power grid dimensions are based on the 14 nm technology node. The pitch of the global grid is subsequently linearly scaled to N10 and N07 based on the global grid in 14 nm technology. Model generation and simulation are based on MATLAB and Cadence Spectre. The contribution of power noise in advanced technology nodes is discussed in Section 4.1. A comparison of the local power noise for different technology nodes is provided in Section 4.2.

4.1. Power noise components

The power noise is assumed to be the peak power noise due to simultaneously switching loads. The total on-chip power noise is the voltage variation from the power pad to the local V_{DD}. The entire power distribution network structure, including the global power grids, local power rails, and via stacks, contributes to the power noise. To

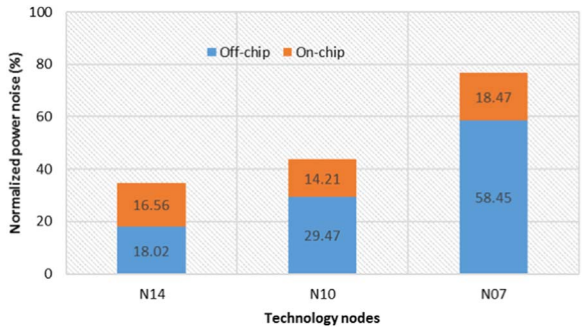


Fig. 5. Comparison between on-chip and off-chip power noise in 14, 10, and 10 nm technology nodes.

comprehensively evaluate the power noise from the perspective of technology scaling, the total power noise considers the hierarchy of the power network. The global power noise consists of *IR* and *L di/dt* noise introduced by the mesh grid and high density transient currents. The local power noise is due to the highly resistive power rails and is dominant in advanced technology nodes. The via stack power noise is due to *IR* drops across stacked vias connecting the global power grid to the local power rails. The resistance of each local via is significant. The resistance of a via between metal 1 and metal 2 in the 10 nm technology node can reach 30 Ω [4].

The package inductance is important not only at the package and board levels but also at the IC level. A comparison between on-chip and off-chip power noise for different technology nodes is illustrated in 5. Note that the assumed package impedance is based on [21]. The power noise is averaged across clock frequencies ranging from DC to 5 GHz.

The total on-chip power noise ranges from 14.2% to 18.5% in 14, 10, and 7 nm technology nodes with a trend of increasing power noise with technology scaling, although the 10 nm node exhibits lower power noise than the other two nodes (see Fig. 6). The reason is that the reduction in power noise in global power grids is larger than the increase in power noise in local power rails and via stacks.

The distribution of the three power noise components vary with technology scaling, as illustrated in Fig. 6. The power noise is averaged across clock frequencies ranging from DC to 5 GHz. The via stack power noise and local power noise exhibit the same trend of increasing noise as technology scales due to the significant resistance of the vias and local power rails. The global power noise, however, decreases 4.6% and 4.0%, respectively, in N10 and N07 as compared with N14, which is 10.8%. This reduction in global power noise is due to the lower resistance and inductance of the global power grid in N10 and N07 due to the decreasing global power/ground dimensions. Notably, global power noise in N14 is 10.8%, which dominates the total power noise, as compared with 6.2% and 6.8%, respectively, in N10 and N07. A metalization scheme which reduces the global power noise is therefore

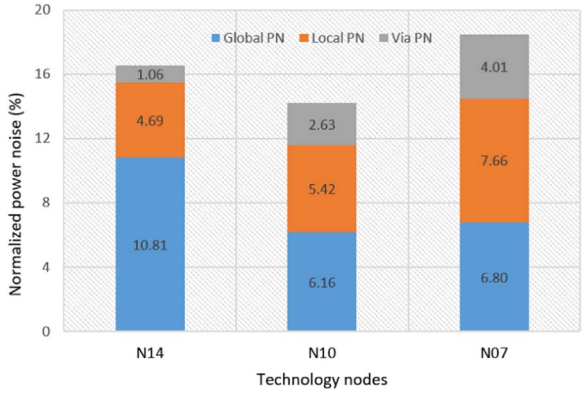


Fig. 6. Components of on-chip power noise in 14, 10, and 10 nm technology nodes.

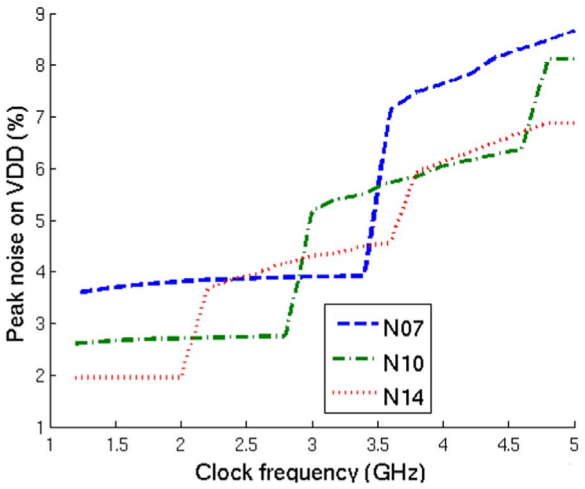


Fig. 7. Local peak power noise in 14 nm, 10 nm, and 7 nm technologies with increasing clock frequency.

preferable in N14. For N07, local power noise is the largest contributor to the total power noise, indicating methods to reduce local power noise are needed in N07. As an effective method to mitigate local power noise, the effects of graphene interconnects on power noise suppression are discussed in the following section.

4.2. Different technology nodes

The local V_{DD} rails exhibit a peak power noise that ranges from 3% to 10% of V_{DD} with a trend of increasing power noise with technology scaling. As the clock frequency supported by the track increases, the power noise increases in discrete steps, as illustrated in Fig. 7. Each step is due to the larger number of loads that simultaneously switch on a track rail, which corresponds to a relative decrease in logic depth. Local noise levels also increase with each technology node, although the magnitude of the noise is strongly dependent on the clock frequency and number of loads per rail. At lower frequencies with only a single load switching per rail, N10 and N07 exhibit, respectively, power noise increases of 0.7% and 1.8% as compared to N14. At higher frequencies with two loads per rail, the power noise increases, respectively, by 1.8% and 4.1%. This behavior is expected as the width of a standard cell gate is proportionally larger with scaled technologies, producing a larger track resistance per cell.

To measure the effects of power noise on circuit performance, a five stage ring oscillator (RO) is driven with power noise injected into both the power and ground rails. The per cent reduction in ring oscillator frequency is depicted in Fig. 8. As the power noise increases with frequency, the performance of the ring oscillator decreases. As expected, the RO performance increases with each technology generation and drops discretely with increasing clock frequency. Notably, the magnitude of the decrease in oscillator frequency is much higher in N07 than in N10 and N14, indicative of the higher sensitivity to power noise with device scaling. At frequencies above 3 GHz, the performance of the N07 ring oscillator drops below the performance of the N10 ring oscillator operating at a lower clock frequency. Intuitively, the delay of an N07 circuit degrades, losing the advantages of scaling. Maintaining the same performance requires a proportionally smaller P/G pitch that is more aggressive than a linearly scaled grid.

5. Power noise suppression

The dependence of power noise on additional global power metal layers, stripes, graphene interconnect, and local interconnect scaling is discussed in this section. Methods to suppress power noise in power distribution networks are discussed in the following subsections. The

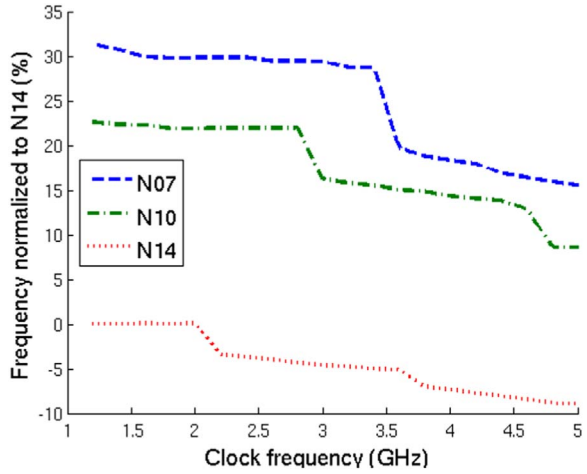


Fig. 8. Per cent decrease in performance of average power noise of a five stage ring oscillator in 14 nm, 10 nm, and 7 nm technologies normalized to an N14 ring oscillator.

effectiveness of additional power metal layers to reduce global power noise is presented in Section 5.1. The striping technique is discussed in Section 5.2. Reductions in power noise due to graphene is evaluated in Section 5.3. The scaling scenario for local power rails affects the local and via stack power noise, which is discussed in Section 5.4. A preferable metalization scheme for different technology nodes to reduce the total power noise is discussed in Section 5.5.

5.1. Additional global power metal layers

As technology is scaled, the number of on-chip metal layers increases, resulting in multiple metal layers available for the global power network [34]. Adding metal layers to the global power grid introduces more paths for the current to flow, lowering the grid impedance. In this section, reductions in global power noise due to adding layers is evaluated for different advanced technology nodes.

These additional power metal layers are oriented orthogonal to the adjacent metal layers to lower inductive coupling, thereby producing a mesh structure [35]. The size of the metal line and the pitch between adjacent metal lines are assumed the same. As expected, global power noise decreases as more global metal layers are added, as illustrated in Fig. 9. The rate of global power noise reduction however decreases with increasing number of additional global layers. Further increase the number of dedicated layers is not efficient to reduce global power noise. Note in Fig. 9 that the baseline of the global power grid is two layers. The greatest reduction in global power noise for N14, N10, and N07 is, respectively, 8.1%, 4.6%, and 5.2% when an additional six metal layers are dedicated to the global power grid. Adding power layers is shown to

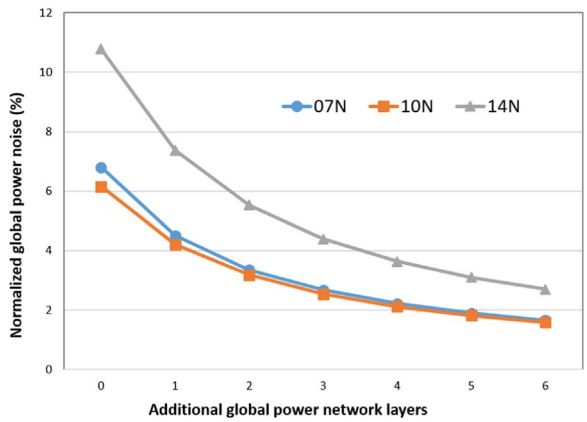


Fig. 9. Degradation in global power noise versus additional global power metal layers.

be more advantageous in N14 as compared with N10 and N07 where global power noise is less significant.

5.2. Stripes technique

One method to reduce local power noise is applying multiple stripes to adjacent track rails. As a primary component of on-chip power noise, local power noise become dominant in the N07 node. To reduce local power noise, an individual track rail can use multiple stripes to the adjacent rails, each with a variable width. The noise exhibited by a 3.6 GHz circuit with striping for variable width and count is illustrated in Fig. 10. For reference, the peak noise of a 3.6 GHz circuit without striping for the N14, N10, and N07 technology nodes is, respectively, 4.6%, 5.7%, and 7.1%. The stripe count is the number of stripes per track rail, and the stripe width is the pitch of a stripe with additional vias. The stripe count and stripe width are both normalized to the minimum metal pitch of the technology node.

Introducing striping reduces power noise by almost a factor of two for each technology node, with a slight increase in noise reduction with each technology generation. The maximum stripe width and count, with nine stripes at a stripe width of ten, is impractical in conventional circuits for any technology node. In these cases, ten cells are between each stripe, and each stripe is approximately the size of four inverter cells. These additional interconnects cause significant routing congestion and area overhead.

Much benefit, however, can be achieved with wide stripes. A single stripe with a stripe width of ten reduces the power noise by almost a third for N14, N10, and N07. This reduction in noise is due to the relatively large resistance of the via for each stripe. As the stripe width increases, additional vias can be added, reducing the effective resistance of the stripe, thereby lowering the resistance of the path to the power supply. At stripe counts greater than five, there are diminishing returns on the reduction in power noise. In this case, a stripe width above six reduces much of the power noise without incurring excessive overhead.

5.3. Graphene interconnects

Another method to reduce power noise is exploiting lower resistivity material in power grids to reduce the effects of the “resistivity wall.” As illustrated in Fig. 1, the resistivity of GNRs is comparable to copper, and the resistivity of FLG is lower than copper in deeply scaled metal lines [10]. Although integrating graphene with CMOS technology is not yet practical, graphene as an interconnect replacement significantly reduces power noise.

Power noise is evaluated for the 7 nm technology with five stripes across the power ground rails for three different materials. The resistivity of GNRs is extracted from experimental data based on the local interconnect width used in 7 nm technology [10]. The resistivity of FLG is determined based on the sheet resistance reported in [11] and the thickness of typical five layer graphene. The third material is copper. As illustrated in Fig. 11, a large difference in power noise between FLG and copper is exhibited since the difference in resistivity is significant in 7 nm interconnect technology. A 59.1% reduction in peak noise is achieved with FLG as compared with copper. The bottleneck is the vias between two adjacent metal layers, which is highly resistive in advanced technology nodes as compared to the resistance of the metal lines.

5.4. Scaling of local power rails

For a global mesh structured power grid, the pitch of each power/ground pair decreases with technology scaling. The width of the global power/ground interconnect is however fixed in advanced technology nodes to prevent an increase in the impedance of the global power grid. Widening the global power grid significantly increases on-chip area

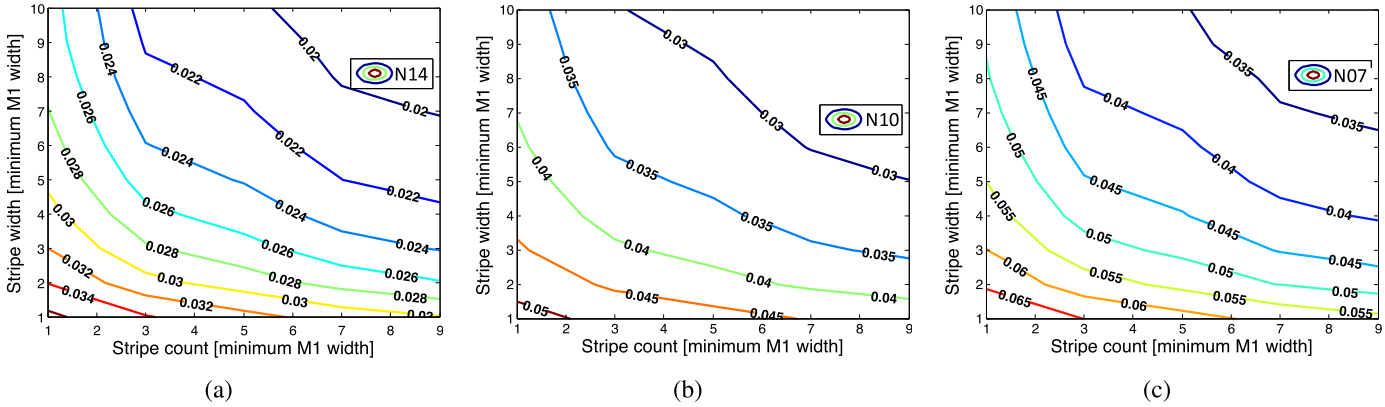


Fig. 10. Effect of track stripe count and stripe width on normalized power noise, a) 14 nm, b) 10 nm, and c) 7 nm technologies. A 3.6 GHz frequency is assumed.

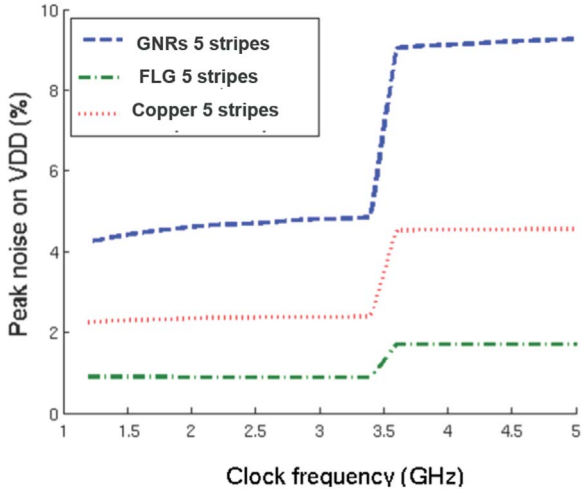


Fig. 11. Peak power noise in GNRs, FLG, and copper power grids with increasing clock frequencies in 7 nm technology.

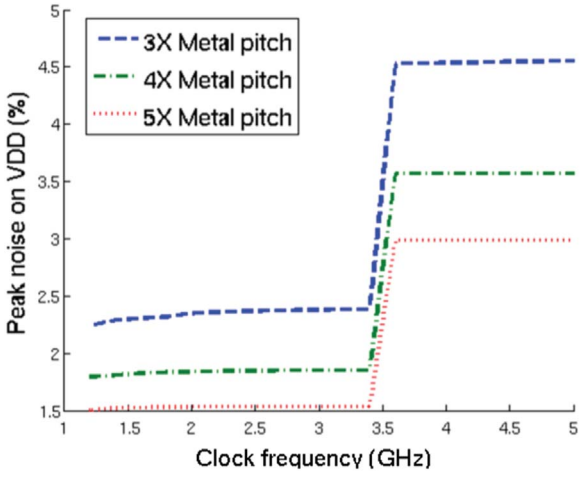


Fig. 12. Peak noise in 3X, 4X, and 5X minimum metal pitch interconnect scaling scenarios with increasing clock frequencies in 7 nm technology.

while also introducing a larger parasitic capacitance between adjacent metal layers. For interdigitated local power rails, the pitch of the adjacent power and ground rail is proportional to the gate pitch to match the standard cell height for each technology node. The width of the local power rail is proportional to the minimum metal pitch of each technology. This scaling process is a primary source of power noise due to IR drops.

In evaluating power noise, the width of the local power rail is set to three times the minimum metal pitch of each technology. A tradeoff should be considered between the physical area and the impedance characteristics of the local power rails to satisfy power noise budgets in advanced technology nodes. A smaller standard cell height allows the on-chip area of the local power rails to be increased while maintaining performance improvements. As illustrated in Fig. 12, a 32.4% reduction in peak noise is exhibited after increasing the power rail width from three times to five times the minimum metal pitch in a 7 nm technology. As compared with Fig. 11, the reduction in power noise with larger power rail widths is lower than exploiting new interconnect materials. Changing the metal width is however more practical since this change does not require novel fabrication and integration technologies. Increasing the width of the local power rail degrades performance due to the large area overhead of the local metal layer.

5.5. Metalization schemes for advanced technology nodes

In this section, power noise is compared for four different scenarios (A: baseline case, B: adding two extra metal layers for the global power network, C: increasing the local metal line width to five times the minimum metal pitch, D: adding five stripes to each power track). For reference, in the baseline case, the power network utilizes two metal layers for the global grid. The local metal width is three times the minimum metal pitch without striping.

Comparing scenarios B, C, and D with A, the total power noise in N14, N10, and N07 is suppressed. The greatest reduction in power noise is, respectively, 5.1% (case B), 4.6% (case C), and 1.1% (case D), as illustrated in Fig. 13. An additional two power layers significantly decrease power noise in N14 but has limited effect on N10 and N07, indicating in N14 that adding global power layers is preferable to reducing power noise. A 5X metal line width achieves the greatest reduction in power noise (5.1%) as compared with adding five stripes (3.0%) in 7 nm technology. A wider local metal line more efficiently

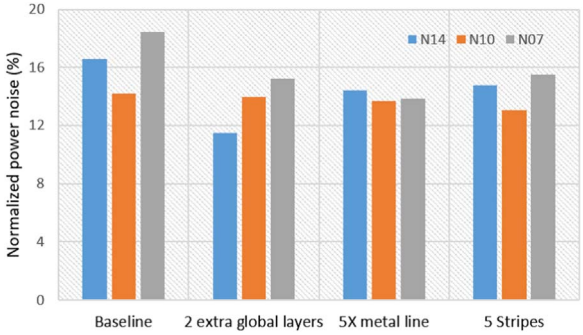


Fig. 13. Total power noise in 14, 10, and 7 nm technology nodes for four different scenarios.

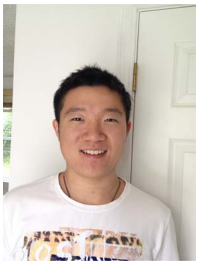
suppresses power noise in N07 than the stripping technique. Local power noise is reduced by wider metal lines, as well as via stacking due to the larger vias. As a result, widening the metal line has the greatest potential to reduce the total power noise in 7 nm technology. Comparing scenarios B, C, and D for N10, a 5X metal line width is less efficient in suppressing power noise than the five stripes technique. Widening the metal line is less advantageous due to the relatively high resistance via stack in N10 as compared with N07.

6. Conclusions

An exploratory modeling methodology is proposed for assessing power noise in standard cell digital circuits. Models are discussed for 14, 10, and 7 nm technologies to evaluate noise trends. Local resistive noise is shown to increase with technology scaling and starts to dominate the total power noise at the 7 nm node. The effects of local stripes are evaluated on power grids, exhibiting a 2X reduction in local power noise. Adding global power metal layers is an effective method to reduce global power noise. Exploiting new materials in on-chip interconnect exhibits good potential to lower power noise in advanced technology nodes. Tradeoffs between power noise and performance need to be carefully considered when scaling the width of the local power rails. In 14 nm technology, providing additional global metal layers is preferable to lowering power noise, where a 30.6% reduction in power noise is exhibited. Below 10 nm, local power rails with wider metal lines are effective in suppressing local and via stack power noise.

References

[1] K. Schuegraf, et al., Semiconductor Logic Technology Innovation to Achieve Sub-10 nm Manufacturing, IEEE Transactions on Electron Devices 1 (3) (2013) 66–75.
[2] J. Whitehouse, E. John, Leakage and Delay Analysis in FinFET Array Multiplier Circuits, Proceedings of the IEEE International Midwest Symposium on Circuits and Systems, August 2014, pp. 909–912.
[3] S. Natarajan et al., A 14nm Logic Technology Featuring 2nd-Generation FinFET, Air-Gapped Interconnects, Self-Aligned Double Patterning and A 0.0588 m^2 SRAM Cell Size, Proceedings of the IEEE International Electron Devices Meeting, December 2014, pp. 3.7.1-3.7.3.
[4] ITRS Technology Working Groups, International Technology Roadmap for Semiconductors (ITRS), 2015.
[5] W. Steinhögl, et al., Comprehensive Study of the Resistivity of Copper Wires with Lateral Dimensions of 100 nm and Smaller, Journal of Applied Physics 97 (2) (2005) 023706-1–023706-7.
[6] Pirjo Pasanen, et al., Graphene for Future Electronics, Physica Scripta 2012 (T146) (2012) 014025.
[7] J.S. Moon, D.K. Gaskill, Graphene Its Fundamentals to Future Applications, IEEE Transactions on Microwave Theory and Techniques 59 (10) (2011) 2702–2708.
[8] A.C. Ferrari, et al., Science and Technology Roadmap for Graphene, Related Two-Dimensional Crystals, and Hybrid Systems, Nanoscale 7 (2015) 4598–4810.
[9] Fred Lacy, Developing A Theoretical Relationship between Electrical Resistivity, Temperature, and Film Thickness for Conductors, Nanoscale Research Letters 6 (1) (2011) 1–14.
[10] R. Murali, et al., Resistivity of Graphene Nanoribbon Interconnects, IEEE Electron Device Letters 30 (June (6)) (2009) 611–613.
[11] I. Khrapach, et al., Novel Highly Conductive and Transparent Graphene-Based Conductors, Advanced Materials 24 (21) (2012) 2844–2849.
[12] E. Salman, E.G. Friedman, High Performance Integrated Circuit Design, McGraw Hill Professional, 2012.
[13] A.V. Mezhiba, E.G. Friedman, Scaling trends of on-chip power distribution noise, IEEE Transactions on Very Large Scale Integration (VLSI) Systems 12 (4) (2004) 386–394.
[14] N. Srivastava, Xiaoning Qi, K. Banerjee, Impact of On-Chip Inductance On Power Distribution Network Design for Nanometer Scale Integrated Circuits, Proceedings of the IEEE International Symposium on Quality Electronic Design, March 2005, pp. 346–351.
[15] K. Vaidyanathan, et al., Design and Manufacturability Tradeoffs in Unidirectional and Bidirectional Standard Cell Layouts in 14 nm Node, Proceedings of SPIE 8327 (2012) (pp. 83270K–83270K-12).
[16] S. Lin, N. Chang, Challenges in Power-Ground Integrity, Proceedings of the IEEE/ACM International Conference on Computer-Aided Design, pp. 651–654, November 2001.
[17] R. Jakushokas, E.G. Friedman, Multi-Layer Interdigitated Power Distribution Networks, IEEE Transactions on Very Large Scale Integration (VLSI) Systems 19 (5) (2011) 774–786.
[18] I.S. Kourtev, B. Taskin, E.G. Friedman, Timing Optimization Through Clock Skew Scheduling, Springer International Publishing, 2009.
[19] R. Jakushokas, E.G. Friedman, Inductance Model of Interdigitated Power and Ground Distribution Networks, IEEE Transactions on Circuits and Systems II: Express Briefs 56 (7) (2009) 585–589.
[20] M. Popovich, E.G. Friedman, R. Secareanu, O.L. Hartin, On-Chip Power Noise Reduction Techniques in High Performance SoC-Based Integrated Circuits, Proceedings of the IEEE International SOC Conference, September 2005, pp. 309–312.
[21] M.S. Gupta, et al., Understanding voltage variations in chip multiprocessors using a distributed power-delivery network, in: Proceedings of the IEEE Design, Automation Test in Europe, pp. 1–6, April 2007.
[22] E. Salman, E.G. Friedman, R.M. Secareanu, O.L. Hartin, Worst Case Power/Ground Noise Estimation Using an Equivalent Transition Time for Resonance, IEEE Transactions on Circuits and Systems I: Regular Papers 56 (5) (2009) 997–1004.
[23] R. Patel, P. Raghavan, E.G. Friedman, Power Noise in 14, 10, and 7 nm FinFET CMOS Technologies, Proceedings of the 2016 IEEE International Symposium on Circuits and Systems, May 2016, pp. 37–40.
[24] M. Popovich, M. Sotman, A. Kolodny, E.G. Friedman, Effective Radii of On-Chip Decoupling Capacitors, IEEE Transactions on Very Large Scale Integration (VLSI) Systems 16 (7) (2008) 894–907.
[25] I.P. Vaisband, R. Jakushokas, M. Popovich, A.V. Mezhiba, S. Kose, E.G. Friedman, On-Chip Power Delivery and Management, Springer International Publishing, 2016.
[26] R. Jakushokas, M. Popovich, A.V. Mezhiba, S. Kose, E.G. Friedman, Power Distribution Networks with On-Chip Decoupling Capacitors, Springer International Publishing, 2011.
[27] A. Todri, M. Marek-Sadowska, F. Maire, C. Matheron, A Study of Decoupling Capacitor Effectiveness in Power and Ground Grid Networks, Proceedings of the IEEE International Symposium on Quality Electronic Design, March 2009, pp. 653–658.
[28] X. Yang, B.K. Choi, M. Sarrafzadeh, Routability-Driven White Space Allocation for Fixed-Die Standard-Cell Placement, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 22 (4) (2003) 410–419.
[29] S.S.Y. Liu et al., Effective Power Network Prototyping via Statistical-based Clustering and Sequential Linear Programming, Proceedings of the IEEE Conference on Design, Automation and Test in Europe, March 2013, pp. 1701–1706.
[30] C.C. Huang et al., Improving Power Delivery Network Design by Practical Methodologies, Proceedings of the IEEE International Conference on Computer Design (ICCD), October 2014, pp. 237–242.
[31] L.T. Wang, Y.W. Chang, K.T. Cheng, Electronic Design Automation: Synthesis, Verification, and Test, Morgan Kaufmann Publishing, 2009.
[32] Z. Zeng, T. Xu, Z. Feng, P. Li, Fast Static Analysis of Power Grids: Algorithms and Implementations, Proceedings of the IEEE International Conference on Computer-Aided Design (ICCAD), November 2011, pp. 488–493.
[33] J. Yang, Z. Li, Y. Cai, Q. Zhou, PowerRush: A Linear Simulator for Power Grid, Proceedings of the IEEE International Conference on Computer-Aided Design (ICCAD), pp. 482–487, November 2011.
[34] C.H. Lin et al., High Performance 14nm nm SOI FinFET CMOS Technology with 0.0174 m^2 Embedded DRAM and 15 Levels of Cu Metallization, Proceedings of the IEEE International Electron Devices Meeting, December 2014, pp. 3.8.1-3.8.3.
[35] A.V. Mezhiba, E.G. Friedman, Impedance Characteristics of Power Distribution Grids in Nanoscale Integrated Circuits, IEEE Transactions on Very Large Scale Integration (VLSI) Systems 12 (11) (2004) 1148–1155.



Kan Xu received the B.S. degree in electrical engineering from North China University of Water Resources and Electric Power, Zhengzhou, China, 2012, and the M.S. degree in electrical and computer engineering from the University of Rochester, Rochester, New York, in 2014. He is currently working towards the Ph.D. degree in electrical engineering from the University of Rochester, Rochester, New York under the supervision of Prof. Eby G. Friedman. His current research interests include the analysis and design of high performance integrated circuits, 3-D integration, on-chip power distribution network, on-package electromagnetic interference, and EDA design for photonics ICs.

Ravi Patel received the B.Sc., M.Sc. and PhD degreees in electrical and computer engineering from the University of Rochester, Rochester, NY, USA, in 2008, 2010, and 2016, specializing in memristor based computer circuits and memory. He was a Research Intern with Freescale Semiconductor Inc., Tempe, AZ, USA, in 2010 and 2012, and imec, Leuven, Belgium, in 2014, where he was investigating power network design for sub-10-nm integrated circuits. Currently, he is lead machine learning systems engineer for Kheiron Medical, Limited. His current research interests include memristors, as well as deep learning acceleration, binary neural networks, and high performance computing.

Praveen Raghavan obtained his PhD from KU Leuven in 2009, his MS from Arizona State University, USA and his Bachelors in Electrical Engineering from REC Trichy, India. In 2007, he was also a visiting researcher at Berkeley Wireless Research Center (BWRC), University of Berkeley, California. He is currently the principal engineer leading the group in design enabled technology exploration for device-design co-optimization on CMOS and beyond CMOS technologies. He has published over 250 conference and journal papers and holds more than 30 patents.

Eby G. Friedman received the B.S. degree from Lafayette College in 1979, and the M.S. and Ph.D. degrees from the University of California, Irvine, in 1981 and 1989, respectively, all in electrical engineering.

From 1979–1991, he was with Hughes Aircraft Company, rising to the position of manager of the Signal Processing Design and Test Department, responsible for the design and test of high performance digital and analog IC's. He has been with the Department of Electrical and Computer Engineering at the University of Rochester since 1991, where he is a Distinguished Professor, and the Director of the High Performance VLSI/IC Design and Analysis Laboratory. He is also a Visiting Professor at the Technion - Israel Institute of Technology. His current research and teaching interests are in high performance synchronous digital and mixed-signal microelectronic design and analysis with application to high speed portable processors, low power wireless communications, and high performance server farms.

He is the author of more than 500 papers and book chapters, 14 patents, and the author or editor of 18 books in the fields of high speed and low power CMOS design techniques, 3-D design methodologies, high speed interconnect, and the theory and application of synchronous clock and power distribution networks. Dr. Friedman is the Editor-in-Chief of the *Microelectronics Journal*, a Member of the editorial boards of the *Journal of Low Power Electronics* and *Journal of Low Power Electronics and Applications*, and a Member of the technical program committee of numerous conferences. He previously was the Editor-in-Chief and Chair of the steering committee of the *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, the Regional Editor of the *Journal of Circuits, Systems and Computers, a Member of the editorial board of* the *Proceedings of the IEEE*, *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, *Analog Integrated Circuits and Signal Processing*, and *Journal of Signal Processing Systems*, a Member of the Circuits and Systems (CAS) Society Board of Governors, Program and Technical chair of several IEEE conferences, and a recipient of the IEEE Circuits and Systems Charles A. Desoer Technical Achievement Award, a University of Rochester Graduate Teaching Award, and a College of Engineering Teaching Excellence Award. Dr. Friedman is a Senior Fulbright Fellow and an IEEE Fellow.