Optimal Allocation of LDOs and Decoupling Capacitors within a Distributed On-Chip Power Grid

SAYED ABDULLAH SADAT, University of Utah
MUSTAFA CANBOLAT, State University of New York at Brockport
SELÇUK KÖSE, University of South Florida

Parallel on-chip voltage regulation, where multiple regulators are connected to the same power grid, has recently attracted significant attention with the proliferation of small on-chip voltage regulators. In this article, the number, size, and location of parallel low-dropout (LDO) regulators and intentional decoupling capacitors are optimized using mixed integer non-linear programming formulation. The proposed optimization function concurrently considers multiple objectives such as area, power noise, and overall power consumption. Certain objectives are optimized by putting constraints on the other objectives with the proposed technique. Additional constraints have been added to avoid the overlap of LDOs and decoupling capacitors in the optimization process. The results of an optimized LDO allocation in the POWER8 chip is compared with the recent LDO allocation in the same IBM chip in a case study where a 20% reduction in the noise is achieved. The results of the proposed multi-criteria objective function under a different area, power, and noise constraints are also evaluated with a sample ISPD'11 benchmark circuits in another case study.

CCS Concepts: • **Hardware** → **On-chip resource management**; *VLSI design manufacturing considerations*; *Chip-level power issues*; Network on chip; Circuits power issues;

Additional Key Words and Phrases: Power delivery network (PDN), distributed on-chip voltage regulator, current sharing, physical design, decoupling capacitors

1 INTRODUCTION

Integrated circuit designers have recently shifted their interest from solely speeding up a single processor to utilizing the potentials of parallel computing with multiple cores on a chip to better balance system performance, heat dissipation, and power efficiency (Li et al. 2017). High-performance circuits consume higher current, operate at higher speeds, and have lower
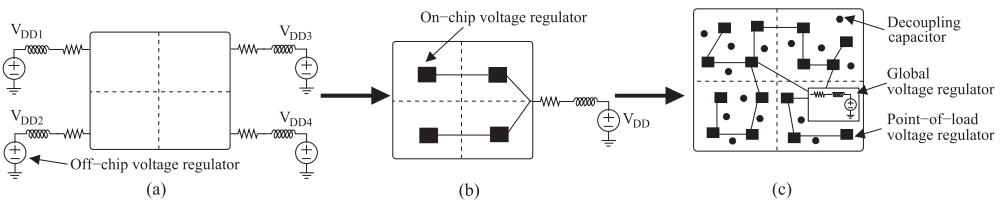
Fig. 1. Evolution of integrated voltage regulators. (a) Traditional off-chip voltage regulators. (b) On-chip voltage regulators to reduce the number of input pins and increase power efficiency. (c) Simultaneous placement of distributed point-of-load regulators and decoupling capacitors to enhance the overall signal integrity, power dissipation, and performance.

noise tolerance with the introduction of each new technology generation (Vaisband et al. 2016). The high-quality power delivery system, which is responsible for delivering sufficient and stable power from an off-chip source to all of the on-chip functional units and realizes fast reliable power management, plays a crucial role in guaranteeing the proper functionality of the whole system (Li et al. 2017). As much as it is a fundamental requirement of all integrated circuits, it remains a significant design challenge for integrated circuit designers (Wang et al. 2018; Zhan et al. 2016).

Novel voltage regulator topologies have recently been proposed (Guo and Leung 2010; Hazucha et al. 2005; Köse et al. 2013; Leung and Mok 2003; Ramadass et al. 2010), enabling not only the integration of on-chip voltage regulators but also multiple distributed on-chip point-of-load voltage regulators (Köse and Friedman 2012; Lai et al. 2013; Sanders et al. 2013; Zhou et al. 2014). These on-chip point-of-load voltage regulators provide the necessary voltage close to the load circuits, greatly reducing the parasitic impedance between the load circuits and voltage regulators and thereby enhancing the efficiency of the overall power delivery system (Karnik et al. 2013).

The low-dropout (LDO) regulator is suitable for on-chip integration and is a key component in on-chip power management as it exhibits a fast load regulation, high power efficiency, as well as stability over a wide range of current loads and process, voltage, and temperature (PVT) variations (Vaisband and Friedman 2016).

Next generation power delivery networks within high-performance circuits will contain hundreds of on-chip voltage regulators supported by local decoupling capacitors to satisfy the current demand of billions of load circuits within different voltage islands (Köse and Friedman 2010a; Kurd et al. 2014), as illustrated in Figure 1. The design of these complex systems would be greatly enhanced if the available resources, such as the physical area, number of metal layers, and power budget, were not severely limited. The continuous demand over the past decade for greater functionality within a small form factor has imposed tight resource constraints while achieving aggressive performance and noise targets (Wang and Marek-Sadowska 2005). Heterogeneous architectures have emerged as a promising solution to enhance energy-efficiency by allowing each application to run on a core that matches resource needs more closely than a one-size-fits-all core (Tavana et al. 2015; Wang et al. 2017).

Several techniques have been proposed for efficient power delivery systems, typically focusing on optimizing the power network (Khatamifard et al. 2017; Tan and Shi 2001; Wang and Marek-Sadowska 2005) and placement of the decoupling capacitors (Köse and Friedman 2010b, 2011a; Pant et al. 2002; Popovich et al. 2007). Zeng et al. (2010) proposed an optimization technique for designing power networks with multiple on-chip voltage regulators. Yu and Wong (2014) proposed a placement technique for LDO regulators by modelling the LDOs as ideal voltage sources and placing them one by one to minimize the IR voltage drops. The LDOs are placed to the nodes that have the maximum IR voltage drop. Zeng et al. (2013) proposed a co-design technique that

considers on-chip voltage regulators together with decaps. The model proposed is optimizing the LDO placement while evenly distributing the LDO block in the power delivery network, which is employing search based simulation-optimization algorithms. The design of these on-chip voltage regulators and the effect of these regulators on high-frequency voltage fluctuations and mid-frequency resonance have been investigated. Parallel voltage regulation utilizing LDO regulators has been investigated and a methodology has been proposed to ensure the stability of the parallel LDO system in Lai et al. (2013). Parallel voltage regulation utilizing switched capacitor (SC) voltage regulators has been recently investigated and a two-level optimization technique has been proposed to determine the size, location, and the number of required SC regulators that maximizes the power conversion efficiency of the system (Zhou et al. 2014). The interactions between the voltage regulators and decoupling capacitors, which can significantly affect the performance of an integrated circuit, are critical in producing a robust power distribution network. Decoupling capacitors and on-chip voltage regulators exhibit several distinct characteristics, such as the response time, area requirements, and parasitic output impedance. Circuit models of these components should accurately capture these characteristics while being sufficiently simple to not computationally constrain the optimization process. Simultaneously optimization of the location of a predefined number of on-chip LDO regulators and decoupling capacitor to minimize the power noise has been proposed in Köse and Friedman (2012).

In this article, the application of mixed integer non-linear programming (MINLP) is examined in minimizing the maximum voltage drop, total area, the response time for particular circuit blocks, and total power consumption. The contribution of this article is to concurrently determine an optimum number, size, and location of parallel LDO voltage regulators and decoupling capacitors for different design constraints, while preventing the overlapping of any two or more regulators and decoupling capacitors (Daskin 1995; Drezner and Hamacher 2002; Farahani et al. 2010). The constraints of this power network co-design problem depend on the application and performance objectives. Multiple optimization goals that are controlled by different weighting parameters are applied in the proposed function.

The remaining part of the article is organized as follows. Distributed power delivery with parallel LDO regulators is explained in Section 2. The three objective components are explained and the proposed methodologies to concurrently determine the optimum location, size, and the number of the parallel LDO regulators and decoupling capacitors are examined in Section 3. The optimum location, size, and the number of the LDO regulators and decoupling capacitors, exemplified on several benchmark circuits such as IBM POWER8 (Toprak-Deniz et al. 2014) and ISPD'11 benchmark circuit are presented in Section 4. A brief discussion of the proposed optimization technique and possible enhancements are offered in Section 5. The article is concluded in Section 6.

2 DISTRIBUTED POWER DELIVERY WITH PARALLEL LDOS

LDO voltage regulators are widely used for on-chip voltage regulation due to the small area requirement, easy control structure, and fast response time (Guo and Leung 2010; Hazucha et al. 2005; Leung and Mok 2003). Challenges such as device mismatch, offset voltages among parallel regulators, overall system stability, and balanced current sharing need to be considered to achieve a stable system of parallel LDO regulators. Lai et al. recently proposed a methodology based on the hybrid stability theory to ensure the stable operation of parallel LDO regulators (Lai et al. 2013). An LDO regulator that is tailored for parallel voltage regulation (Lai and Li 2012) is used to evaluate the efficacy of the proposed optimization techniques.

In our proposed technique, as part of the case studies we use LDO regulators of different sizes. The LDO regulator models are with a 90.5% peak power efficiency and identical to the one proposed by Toprak-Deniz et al. (2014). The conventional structure of the LDO regulator and some of its
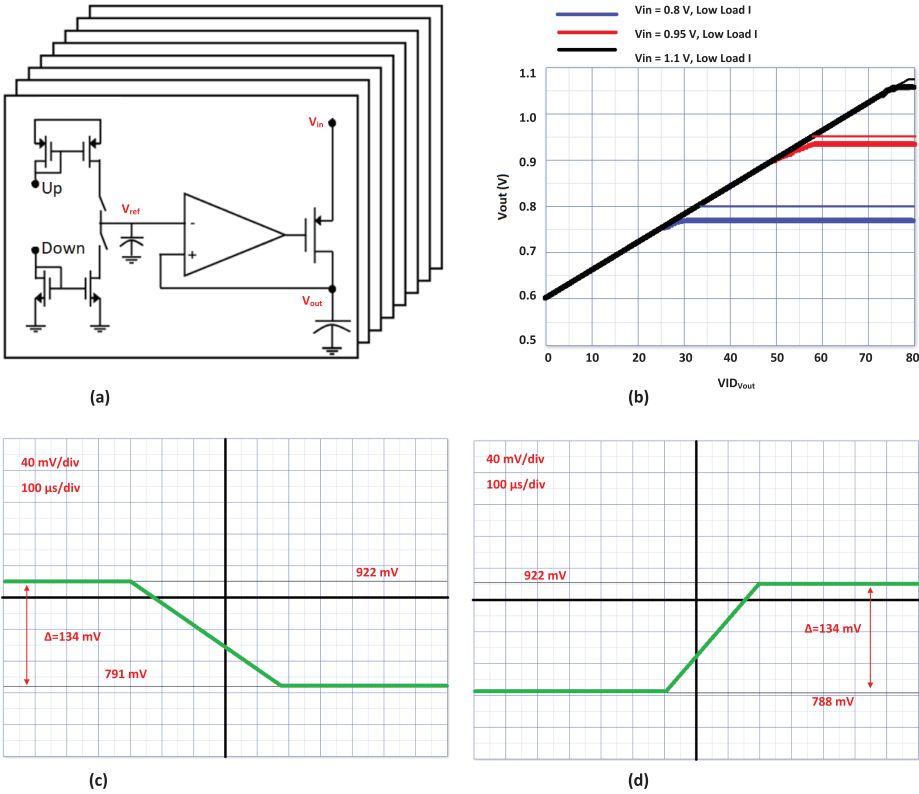
(a)

(b)

(c)

(d)

Fig. 2. Specifications of the LDO regulator. (a) Simplified conventional structure of LDO regulators. (b) Measured V_{out} voltage as a functional of $VID_{V_{out}}$. Only for case with deviations $V_{in} - V_{out} > 50$mV are plotted. (c) Measured V_{out} voltage showing 12.5mV steps in downward and (d) upward directions (Toprak-Deniz et al. 2014).
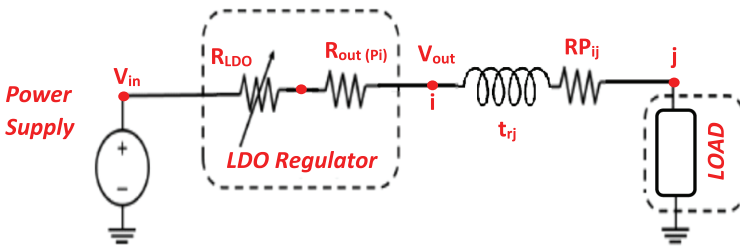
Fig. 3. Simplified model diagram of the LDO regulator with inductance effect represented by the rise time t_{r_j}.

characteristics are shown in Figure 2. For the ISPD's benchmark application, we have modeled LDOs with an RC model similar to the models in Köse and Friedman (2012), Vaisband et al. (2016), and Zhang et al. (2015). Figure 3 shows a simplified model diagram of an LDO with all its basic and equivalent parameters identified. The circuit level schematic of this LDO regulator is shown in Figure 4. The current rating of these LDOs are assumed to vary from 0.8A to 4A. The size of these parallel regulators, as well as their location and number, are determined simultaneously with the decoupling capacitors that can have different physical sizes determined by the proposed optimization.

Fig. 4. LDO regulator for the parallel voltage regulation proposed in Lai and Li (2012).

3 PROPOSED OPTIMIZATION METHODOLOGY

The primary objective of the proposed optimization methodology is to *concurrently* determine the optimal location, number, and size of the on-chip LDO regulators and decoupling capacitors that minimize (i) maximum power noise, (ii) power conversion loss, and (iii) required physical area. A weighted optimization function balances these often conflicting objectives as explained in this section.
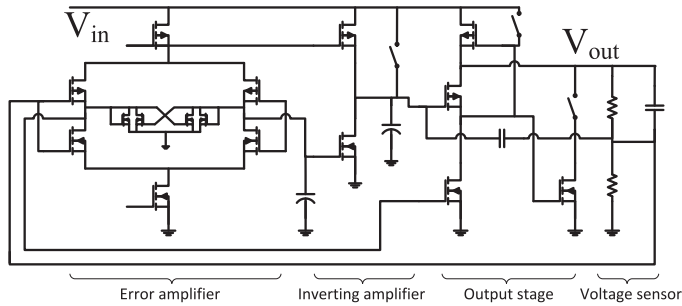
Euclidean or Manhattan distance is widely used in facility location problems to determine a cost function. Alternatively, the cost of delivering power from a local voltage regulator or a decoupling capacitor to a load circuit depends on the parasitic impedance of the power distribution network, the amount of current delivered to the load circuit, and the parasitic impedance to the regulators and decoupling capacitors. A closed-form impedance model, proposed in Köse and Friedman (2011b), is utilized to determine the effective impedance within the power grid from the regulators and decoupling capacitors to the load circuits. The physical distances and power grid characteristics are included within this effective impedance model. Multiple LDO regulators and decoupling capacitors can provide current to a single load circuit, depending on the physical distances among these components (Chiprout 2004). The contribution from the regulators and decoupling capacitors to a load circuit are based on the requirements of the load circuit. For example, when the current profile exhibits a fast transition time, a decoupling capacitor is a better choice due to the faster response of these structures. The definitions of the parameters used in this article are listed in Table 1 and of variables used in this article are listed in Table 2.

Multiple parameters such as the parasitic impedance of the power network, output impedance of the on-chip voltage regulator, effective series resistance of a decoupling capacitor, and load current characteristics significantly affect the power noise. These parameters are therefore considered in the proposed optimization function where the parasitic impedance of the power network is characterized by the closed-form effective impedance model (Köse and Friedman 2011b). In the following subsections, the components of the optimization function along with their formulation will be explored.

3.1 Noise-Aware Physical Design

The proposed optimization function considers both static and transient power noise. The LDO regulators are placed closer to the load circuits to reduce the static voltage drop by minimizing the parasitic impedance between the LDO and the load circuit. The overall current consumption by the load circuits is provided by multiple LDO regulators and each LDO contributes a different amount of charge to the load circuit based on the size of the load and the effective resistance between the LDO and the load. A term CP_{ij} is defined to approximate the current contribution from each LDO

Table 1. Definition of the Parameters in (2)–(20)

Parameter	Definition
x_j, y_j	x-coordinate for circuit j and y-coordinate for circuit j
$P_i \ (i \in I)$	i^{th} LDO
$D_k \ (k \in K)$	k^{th} decoupling capacitor (decap)
$L_j \ (j \in J)$	j^{th} circuit block
$R_{out}(P_i)$	output resistance of the ith LDO in ohms Ω
$R_{esr}(D_k)$	effective series resistance of the kth decap in ohms Ω
$K1$	area of an LDO (excluding the pass transistor and output capacitor) in m^2
KP	area of the pass transistor within an LDO in m^2
KC	area of the output capacitor an LDO in m^2
A_{P_i}	total area of ith LDO in m^2
A_{D_k}	total area of kth decoupling capacitor in m^2
$capP_i$	normalized maximum load current of ith LDO in A
$capD_k$	capacitance of kth decap in F
I_j	DC current demand at circuit j in A
Np	limit on the number of LDOs to be located
Nd	limit on the number of decaps to be located
$T_{cap}P$	limit on the maximum output current of an LDO in A
$T_{cap}D$	limit on the total capacitance of decaps in F
KT	technology dependent parameter to determine the area of a decap from its capacitance
TA_{PD}	limit on the total area of LDOs and decaps in m^2
TA_P	maximum area of a single LDO in m^2
TA_D	maximum area of a single decoupling capacitor in m^2
$V_{drop}P_i$	dropout voltage of the ith LDO in V
t_{r_j}	rise time of the jth load current in s
RP_{ij}	Effective resistance between nodes i and j for LDO placement in ohms Ω
RD_{kj}	Effective resistance between nodes k and j for decap placement in ohms Ω
$W_{area}, W_{noise}, W_{power}$	weighting parameters for physical area, power noise, and power consumption

Table 2. Definition of the Decision Variables

Variable	Definition
xp_i, yp_i	x-coordinate and y-coordinate for the ith LDO
xd_k, yd_k	x-coordinate and y-coordinate for the kth capacitor
vp_i	binary decision variable, equal to 1 if i^{th} regulator is selected and 0 otherwise
vd_k	binary decision variable, equal to 1 if k^{th} decap is selected and 0 otherwise
CP_{ij}	contribution of ith LDO to j^{th} load in A
CD_{kj}	contribution of kth decap to j^{th} load in A

to the load circuits as

$$CP_{ij} = I_j * \frac{G_{ij}}{\sum_{j=1}^{J} G_{ij}}, \tag{1}$$

where G_{ij} is the equivalent conductance between the ith LDO and jth current load, which is the reciprocal of RP_{ij}, i.e., $G_{ij} = 1/RP_{ij}$, and the equation for RP_{ij} is given in Constraint (3).

The transient noise is reduced by a careful placement of decoupling capacitors closer to the fast switching load circuits. The RC time constant determined by the capacitance of the decoupling capacitors and the parasitic resistance between the decap and load circuit are multiplied by the

inverse of the normalized transient rise time of the load current. This approach enables the placement of the decoupling capacitors closer to the load circuits that have faster transients. Intuitively, since the transition time of the current within the blocks with a fast switching activity is smaller, reducing the effective impedance between the decoupling capacitors and these blocks decreases the cost function. Moving the decoupling capacitors close to those circuit blocks requiring a faster transition time minimizes the objective function.

For a decap the transient power noise is determined by multiplying the current from a decap with the effective series resistance of the decap and the effective resistance between the decap and the load. It is to be noted that the transient power noise is also multiplied with a term $(1/t_{r_j})$ to consider the rise time of the load current. The rise time can be obtained through simulations under different load conditions. By considering the rise time of the load current in the second term, the inductive effects are also considered in objective functions (2) and (13). Although this is a first order approximation of modeling the on-chip inductive effects, we have seen a good agreement of this modeling after extensive simulations.

3.2 Formulation to Minimize the Maximum Power Noise

In this part, the objective function (2) of the optimization minimizes the maximum power noise by optimally placing the LDO regulators decaps.

$$\text{Minimize} \qquad \sum_{i=1}^{I} \sum_{j=1}^{J} v_{p_i} CP_{ij}(R_{out} + RP_{ij}) + \sum_{k=1}^{K} \sum_{j=1}^{J} v_{d_k} CD_{kj}(R_{esr} + RD_{kj}) * (1/t_{r_j}), \qquad (2)$$

subject to

$$RP_{ij} = \frac{1}{2\pi}[ln((xp_i - x_j)^2 + (yp_i - y_j)^2) + 3.44388] - 0.033425, \qquad (3)$$

$$RD_{kj} = \frac{1}{2\pi}[ln((xd_k - x_j)^2 + (yd_k - y_j)^2) + 3.44388] - 0.033425, \qquad (4)$$

$$|xp_i - xp_a| \geq v_{p_i}\frac{\sqrt{A_{P_i}}}{2} + v_{p_a}\frac{\sqrt{A_{P_a}}}{2} \qquad \forall i \in I, a \in I, i \neq a, \qquad (5)$$

$$|yp_i - yp_a| \geq v_{p_i}\frac{\sqrt{A_{P_i}}}{2} + v_{p_a}\frac{\sqrt{A_{P_a}}}{2} \qquad \forall i \in I, a \in I, i \neq a, \qquad (6)$$

$$\sum_{j=1}^{J} CP_{ij} \leq capP_i v_{p_i}, \qquad \forall i \in I, \qquad (7)$$

$$\sum_{i=1}^{I} CP_{ij} \geq I_j, \qquad \forall j \in J, \qquad (8)$$

$$\sum_{i=1}^{I} v_{p_i} \leq Np, \qquad (9)$$

$$\sum_{i=1}^{I} v_{p_i} A_{P_i} \leq TA_P, \qquad (10)$$

where the definitions of the aforementioned parameters are listed in Table 1. The area of ith LDO regulator and the area of kth decoupling capacitor are given as:

$$A_{P_i} = K1 + (KP + KC) * capP_i, \tag{11}$$

$$A_{D_k} = KT * capD_k. \tag{12}$$

Constraints (3) and (4) are used to calculate effective resistance between any two arbitrary nodes Rx,y within a mesh when k approaches 1. Köse and Friedman (2011b) has solely focused on deriving this equation. Constraints (5) and (6) ensures that the LDOs don't overlap both in the x-axis as well as in the y-axis. Constraint (7) ensures that the total contribution of current from an LDO voltage regulator cannot exceed the capacity of that particular LDO. The total current demand from all of the load circuits is equal to the total contribution from all of the LDO regulators, as guaranteed by Constraint (8). The number of LDO regulators that can be placed within the circuit is limited by the N_p as shown in Constraint (9). Constraint (10) ensures that the area of the largest LDO regulator is smaller than or equal to the maximum allowed size.

3.3 Area-Aware Physical Design

The physical area of the LDO proposed in Lai and Li (2012) has been divided into three components such as pass transistors, output capacitors, and remaining active circuitry. The size of the pass transistor and the output capacitor is assumed to increase linearly with the output current whereas the size of the remaining active circuitry is assumed to be constant, as shown in Constraint (11). The validity of this assumption has been investigated with simulations and the error is within 10% of the approximation. Alternatively, the physical area of a decoupling capacitor is proportional to the capacitance determined by a technology dependent constant, as shown in Constraint (12).

3.4 Power-Aware Physical Design

Since the decoupling capacitors are passive circuit elements, the power consumed by the LDO regulators is assumed to be the only power loss mechanism considered in this article. Since all the LDO regulators are assumed to have the same characteristics, the overall power conversion loss can be approximated as the total output current times the dropout voltage. At lower load current, the power efficiency of LDO regulators is degraded by the low current efficiency. In this article, without losing the generality and for the sake of simplicity, the current efficiency of the LDO regulators is assumed to be 100%. The optimization function can, however, be modified to also consider other power loss mechanisms.

3.5 Formulation to Minimize All Three Components Using LDOs and Considering Decoupling Capacitors

An objective function, which is a weighted sum of three optimizations, is proposed to determine the optimum number and location of the voltage regulators and decoupling capacitors. The first term that is multiplied with W_{area} is used to minimize the physical area occupied by the regulators and decoupling capacitors. The second term is also comprised of two terms to minimize the average and transient power noise. $(\sum_{i=1}^{I} \sum_{j=1}^{J} v_{p_i} CP_{ij}(R_{out} + RP_{ij}))$ is minimized to reduce the static IR voltage drop and $(\sum_{k=1}^{K} \sum_{j=1}^{J} v_{d_k} CD_{kj}(R_{esr} + RD_{kj}) * (1/t_{r_j}))$ is minimized to reduce the transient power noise. The last term minimizes the power conversion loss of the voltage regulator. The

proposed objective function is

$$
\text{Minimize} \quad W_{area} * \left[\sum_{i=1}^{I} v_{p_i} A_{p_i} + \sum_{k=1}^{K} v_{d_k} A_{D_k} \right]
$$

$$
+ W_{noise} * \left[\sum_{i=1}^{I} \sum_{j=1}^{J} v_{p_i} CP_{ij}(R_{out} + RP_{ij}) + \sum_{k=1}^{K} \sum_{j=1}^{J} v_{d_k} CD_{kj}(R_{esr} + RD_{kj}) * (1/t_{r_j}) \right]
$$

$$
+ W_{power} * \sum_{i=1}^{I} \sum_{j=1}^{J} v_{p_i} CP_{ij} V_{drop} P_i, \tag{13}
$$

subject to Equations (3)–(10) and the following additional constraints:

$$
|xd_k - xd_b| \geq v_{d_k} \frac{\sqrt{A_{D_k}}}{2} + v_{d_b} \frac{\sqrt{A_{D_b}}}{2} \qquad \forall k \in K, b \in K, k \neq b, \tag{14}
$$

$$
|yd_k - yd_b| \geq v_{d_k} \frac{\sqrt{A_{D_k}}}{2} + v_{d_b} \frac{\sqrt{A_{D_b}}}{2} \qquad \forall k \in K, b \in K, k \neq b, \tag{15}
$$

$$
\sum_{j=1}^{J} CD_{kj} \leq capD_k vd_k, \qquad \forall k \in K, \tag{16}
$$

$$
\sum_{k=1}^{K} vd_k \leq Nd, \tag{17}
$$

$$
\sum_{i=1}^{I} vp_i A_{P_i} + \sum_{k=1}^{K} vd_k A_{D_k} \leq TA_{PD}, \tag{18}
$$

$$
\sum_{k=1}^{K} vd_k A_{D_k} \leq TA_D, \tag{19}
$$

$$
W_{area} + W_{noise} + W_{power} = 1. \tag{20}
$$

Equation (13) is the objective function that is the weighted sum of effects of three different design constraints: area, noise, and power. Each of these constraints can be calculated individually. The first term, area, is the area of all of the on-chip LDOs and decoupling capacitors. The second term, noise, is the sum of the static power noise and transient power noise. The static power noise is determined by multiplying the amount of current (CP_{ij}) from an LDO with the sum of the LDO output resistance and the effective impedance between the LDO and the load. Constraints (14) and (15) ensures that the decaps don't overlap both in the x-axis as well as in the y-axis. The number of decoupling capacitors that can be placed within the circuit is limited by the N_d as shown in Equation (17). Additionally, by applying Constraint (18), the total area of the LDO voltage regulators and decoupling capacitors is maintained smaller than or equal to the total permitted area for power delivery network. Constraint (19) ensure that the area of the largest decoupling capacitor is smaller than or equal to the maximum allowed size.

In the proposed optimization function, the weighting terms W_{area}, W_{noise}, and W_{power} (see Table 1) provide the flexibility to optimize the power distribution system for different objectives. When only W_{noise} in the non-zero weighting parameter, the location of the LDO voltage regulators and decoupling capacitors is chosen to minimize the power noise. The algorithm will try to put the
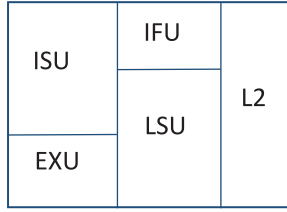
Fig. 5. A schematic diagram demonstrating the floor-plan of a core in an IBM POWER8 chip, showing IFU, LSU, ISU, EXU, and L2.

Table 3. Maximum Current Variations for IFU, LSU, ISU, EXU, L2, and L3 within a Core

	IFU	LSU	ISU	EXU	L2	L3
Max. Load current (A)	2.884	5.253	1.339	5.974	2.472	2.781

maximum number of LDO regulators and decoupling capacitors that would increase the area requirement and power conversion loss. The authors observed meaningful results when none of the weighting terms are nonzero and greater than 0.1. When the maximum load current of the closest LDO regulator is smaller than the total current demand of the integrated circuit, the current is supplied from other LDO regulators. For example, when the total power consumption is the primary bottleneck, adding more decoupling capacitors instead of LDO regulators is a better option if the noise constraints are satisfied. In this case, W_{noise} and W_{power} should be greater than W_{area} to ensure that the weight of the second and third terms in the objective function (13) is greater than the weight of the first term in the objective function (13).

4 CASE STUDY

4.1 IBM POWER8

The schematic diagram of a core is shown in Figure 5, which contains a private L2, an instruction scheduling unit (ISU), an execution unit (EXU), a load store unit (LSU), and an instruction fetch unit (IFU). L1 data cache is a part of LSU, while L1 instruction cache resides inside IFU. The static and transient noise and power conversion loss have been determined for an IBM POWER8-like from SPLASH2X (Bienia et al. 2008). The benchmarks experimented represent typical application domains and features. Eight threads are involved in the simulations and analysis is limited to the region-of-interest of the benchmarks.

An IBM POWER8-like (Fluhr et al. 2014) processor is modeled to quantitatively characterize unbalanced current sharing effects. The maximum load current for various blocks in the core are shown in Table 3.

The formulation presented in Section 3.2, is used to optimally place 64 LDOs as compared to their placement presented in Toprak-Deniz et al. (2014) for the core domains part. The mimicked floor-plan diagram of the core domain part of the work presented in Toprak-Deniz et al. (2014) is shown in Figure 6(a). Figure 6(b) shows the diagram after optimizing the placement of the same number of LDOs (i.e., 64 LDOs) for the objective of reducing noise. CP_{ij} for the Figure 6(a) is calculated based on minimizing the noise to obtain a fair comparison. The results show that the overall noise has been reduced by 20%. The overall noise calculated for Figure 6(a) is 23mV, which is reduced to 18mV when the placement of LDOs are optimized. The runtime was noted as 0.236s.
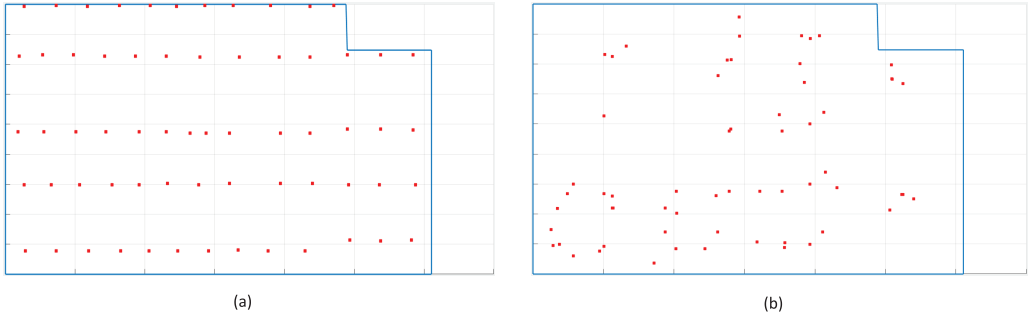
Fig. 6. The blue line covers IFU, LSU, ISU, EXU, and L2 blocks in a floorplan of a core, with red cubes showing the location of LDOs. Figure 6(a) shows a mimicked floorplan of the core domain part of the work presented in Toprak-Deniz et al. (2014). Figure 6(b) shows the diagram after optimizing the placement of the same number of LDOs (i.e., 64 LDOs) for the objective of reducing power noise in the same.
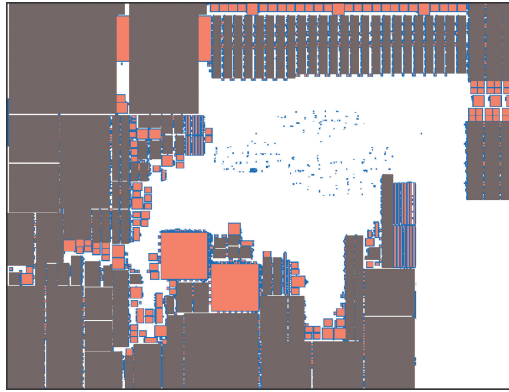

Fig. 7. Floorplan of the ISPD'11 circuit (Viswanathan et al. 2011), superblue5. The light-red shaded boxes with blue boundary represent the rectangular fixed nodes and the gray shaded boxes represent the non-rectangular fixed nodes in the design. The White Space indicates Free-Space.

4.2 Superblue5

The optimum number and location of the LDO voltage regulators and number, location, and size of the decoupling capacitors that minimize the three competing objectives such as (i) the physical area, (ii) static and transient noise, and (iii) power conversion loss have been determined for a sample ISPD'11 placement benchmark suite circuit, superblue5. Superblue5 has a quite asymmetric floorplan as illustrated in Figure 7 and therefore serves as a convenient circuit for placement evaluation. The floorplan of this circuit is comprised of more than 95,000 individual circuit blocks. As shown in Figure 7, a significant portion of the floorplan is occupied by several large circuit blocks. To reduce the complexity of the proposed optimization problem, only the large circuit blocks are considered in the proposed co-design methodology. The actual and reduced number of circuit blocks are listed in Table 4. Although the reduced number of blocks corresponds to less than 0.1% of the actual number of blocks, these fewer number of blocks occupy more than 82% of the total active circuit area.

The size of the power distribution networks and the total number of nodes in superblue5 are listed in Table 4. Each circuit block is modeled as a single current load where the maximum current

Table 4. Properties of the ISPD Benchmark Circuit, Superblue5

circuit	# of blocks	Reduced # of blocks	Coverage of reduced floorplan	Power grid size	# of nodes in the power grid
superblue5	95,041	89	82.46%	774 X 713	551,862

Table 5. Summary of the Results When Different Weights Are Applied for W_{area}, W_{noise}, and W_{power}

W_{area}	W_{noise}	W_{power}	# of LDOs	# of decaps	Total LDO area (mm^2)	Total decap area (mm^2)	Max noise (mV)	Power loss (W)	Run-time (seconds)
0.33	0.34	0.33	5	20	0.049	0.005	44	0.013	309.281
0.8	0.1	0.1	2	3	0.034	0.0008	96	0.008	103.483
0.1	0.8	0.1	8	100	0.064	0.025	22	0.019	296.851
0.1	0.1	0.8	1	100	0.029	0.025	125	0.005	209.412

demand is proportional to the size of the circuit block. Each current load, representing a circuit block, is connected to the power grid from the node physically closest to the center of that particular circuit block.

The general algebraic modeling system (GAMS) is used as the optimization tool (Brooke et al. 1998). The proposed optimization methodology is modeled as a mixed integer nonlinear programming problem. Different weights are given to the competing functions W_{area}, W_{noise}, and W_{power} that satisfy (20) to evaluate their impact on the area, power, and noise characteristics of the ISPD'11 benchmark circuit, superblue5. The model is solved using GAMS on an Intel(R) Core(TM) i7-7700 with processors at 3.6GHz and 16GB RAM. The results are listed in Table 5.

By assigning W_{area}, W_{noise}, and W_{power} to 0.33, 0.33, and 0.34, respectively, a good balance is obtained between the three competing objectives of the area, noise, and power. When one of these terms is given a greater weight, a different optimum number of LDO regulators and decoupling capacitors is determined by the optimization function. For example, when the area is the limiting constraint, a greater value is assigned to W_{area} (0.8) than W_{power} and W_{noise}. The overall area occupied by the LDO regulators and decoupling capacitors is therefore reduced at the expense of higher power noise and power conversion loss. Similar results are observed when power and noise are the limiting constraints and corresponding weighting parameters are increased accordingly, as listed in Table 5.

5   DISCUSSION

Delivering a robust power supply voltage to circuits with varying noise and voltage constraints is crucial to maintaining the performance of next generation integrated circuits. Local supply voltages are generated and regulated by local voltage regulators within a distributed power delivery system. Since the physical distance among the voltage regulators and load circuits is less with a distributed power delivery system, the inductive $L\,di/dt$ and resistive IR power noise are reduced.

In the proposed physical design methodology to allocate on-chip voltage regulators and decoupling capacitors, minimizing the area, power conversion loss, and power noise is the primary

optimization constraints. The distinctive properties of the on-chip voltage regulators and decoupling capacitors should be further exploited to satisfy these constraints while using limited system resources. Although the voltage regulators and decoupling capacitors both provide local charge to the load circuitry, a decoupling capacitor requires a power source to recharge after each clock cycle. The decoupling capacitors provide a faster response with minimal power consumption. Alternatively, the voltage regulators dissipate significant power during voltage down-conversion and regulation. A voltage regulator, however, can provide continuous charge and does not need to be recharged after each clock cycle. The use of weights of different components in the objective function can be decided based on the priority and importance of that particular component under the given circumstances. However, keeping a balanced weight tends to provide more reasonable solution as opposed to the use of extreme weights. This can be clearly seen in Table 5, when the W_{power} is made 80%, it returns the number of LDOs as 1, thereby reducing the area lower than the area obtained when the optimization stressed on the minimization of area, and the output seems practically less reasonable.

6 CONCLUSIONS

The number of integrated voltage regulators on the same die has increased from a single regulator to tens of regulators in the past couple of years. The distributed nature of these regulators will have a significant impact on the power efficiency, power noise, and area requirements. An MINLP optimization function is proposed that minimizes (i) the total area occupied by the LDO regulators and decoupling capacitors, (ii) maximum power noise, and (iii) power conversion loss during the down-conversion of different voltage domains.

In our technique, we defined an objective function that considers multiple objectives, not just the noise. We also place the LDO's not individually. This will help our technique alter the location of an LDO after a number of iterations. Our main difference with other works on optimizing on-chip LDO placement is that in our article we are proposing a placement technique considering different objective that were not considered together in any LDO placement article previously.

A number of constraints to prevent on-chip LDOs overlap in both horizontal and vertical axises were introduced. The current contributions among multiple scattered from voltage regulators and decoupling capacitors to the power grid are considered in the proposed optimization function. The location of the current demand is also considered. The optimal location of the on-chip voltage regulators to minimize the power noise in IBM POWER8 chip is found, where a 20% reduction in the noise is achieved. The number, location, and size of LDOs and decoupling capacitors to minimize all three components are determined for a sample ISPD'11 benchmark suite circuit, where up to 50% reduction in the noise is achieved when stressed on the weight of the noise component.

REFERENCES

C. Bienia, S. Kumar, J. P. Singh, and K. Li. 2008. The PARSEC benchmark suite: Characterization and architectural implications. In *Proceedings of the 17th International Conference on Parallel Architectures and Compilation Techniques (PACT'08)*. ACM, New York, NY, 72–81. DOI : http://dx.doi.org/10.1145/1454115.1454128

A. Brooke, D. Kendrick, A. Meeraus, and R. Raman. 1998. *GAMS: A USER'S GUIDE*. GAMS Development Corporation.

E. Chiprout. 2004. Fast flip-chip power grid analysis via locality and grid shells. In *Proceedings of the IEEE/ACM International Conference on Computer Aided Design (ICCAD'98)*. 485–488. DOI : http://dx.doi.org/10.1109/ICCAD.2004.1382626

M. S. Daskin. 1995. *Network and Discrete Location: Models, Algorithms, and Applications*. John Wiley & Sons.

Z. Drezner and H. Hamacher. 2002. *Facility Location*. Springer.

R. Z. Farahani, M. SteadieSeifi, and N. Asgari. 2010. Multiple criteria facility location problems: A survey. *Appl. Math. Model.* 34, 7 (2010), 1689–1709. DOI : http://dx.doi.org/10.1016/j.apm.2009.10.005

E. J. Fluhr, J. Friedrich, D. Dreps, V. Zyuban, G. Still, C. Gonzalez, A. Hall, D. Hogenmiller, F. Malgioglio, R. Nett, J. Paredes, J. Pille, D. Plass, R. Puri, P. Restle, D. Shan, K. Stawiasz, Z. T. Deniz, D. Wendel, and M. Ziegler. 2014. POWER8: A 12-core

server-class processor in 22nm SOI with 7.6Tb/s off-chip bandwidth. In *Proceedings of the 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC'14)*. 96–97. DOI: http://dx.doi.org/10.1109/ISSCC.2014.6757353

J. Guo and K. N. Leung. 2010. A 6-u W chip-area-efficient output-capacitorless LDO in 90-nm CMOS technology. *IEEE J. Solid-State Circ.* 45, 9 (Sep. 2010), 1896–1905. DOI: http://dx.doi.org/10.1109/JSSC.2010.2053859

P. Hazucha, T. Karnik, B. A. Bloechel, C. Parsons, D. Finan, and S. Borkar. 2005. Area-efficient linear regulator with ultra-fast load regulation. *IEEE J. Solid-State Circ.* 40, 4 (Apr. 2005), 933–940. DOI: http://dx.doi.org/10.1109/JSSC.2004.842831

T. Karnik, M. Pant, and S. Borkar. 2013. Power management and delivery for high-performance microprocessors. In *Proceedings of the International IEEE/ACM Design Automation Conference (DAC'13)*. 1–3. DOI: http://dx.doi.org/10.1145/2463209.2488931

S. K. Khatamifard, L. Wang, W. Yu, S. Köse, and U. R. Karpuzcu. 2017. ThermoGater: Thermally-aware on-chip voltage regulation. In *Proceedings of the IEEE International Symposium on Computer Architecture (ISCA'17)*. 120–132. DOI: http://dx.doi.org/10.1145/3079856.3080250

S. Köse and E. G. Friedman. 2010a. An area efficient fully monolithic hybrid voltage regulator. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*. 2718–2721. DOI: http://dx.doi.org/10.1109/ISCAS.2010.5537035

S. Köse and E. G. Friedman. 2010b. Simultaneous co-design of distributed on-chip power supplies and decoupling capacitors. In *Proceedings of the IEEE International System-on-Chip Conference (SoC'10)*. DOI: http://dx.doi.org/10.1109/SOCC.2010.5784662

S. Köse and E. G. Friedman. 2011a. Distributed power network co-design with on-chip power supplies and decoupling capacitors. In *Proceedings of the ACM/IEEE International Workshop on System Level Interconnect Prediction (SLIP'11)*. DOI: http://dx.doi.org/10.1109/SLIP.2011.6135434

S. Köse and E. G. Friedman. 2011b. Effective resistance of a two layer mesh. *IEEE Trans. Circ. Syst. II* 58, 11 (Nov. 2011), 739–743. DOI: http://dx.doi.org/10.1109/TCSII.2011.2168016

S. Köse and E. G. Friedman. 2012. Distributed on-chip power delivery. *IEEE J. Emerg. Select. Top. Circ. Syst.* 2, 4 (Dec. 2012), 704–713. DOI: http://dx.doi.org/10.1109/JETCAS.2012.2226378

S. Köse, S. Tam, B. McDermott, and E. G. Friedman. 2013. Active filter based hybrid on-chip DC-DC converters for point-of-load voltage regulation. *IEEE Trans. VLSI Syst.* 21, 4 (Apr. 2013), 680–691. DOI: http://dx.doi.org/10.1109/TVLSI.2012.2190539

N. Kurd, M. Chowdhury, E. Burton, T. P. Thomas, C. Mozak, B. Boswell, M. Lal, A. Deval, J. Douglas, M. Elassal, A. Nalamalpu, T. M. Wilson, M. Merten, S. Chennupaty, W. Gomes, and R. Kumar. 2014. Haswell: A family of IA 22nm processors. In *Proceedings of the IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC'14)*. 112–113. DOI: http://dx.doi.org/10.1109/ISSCC.2014.6757361

S. Lai and P. Li. 2012. A fully on-chip area-efficient CMOS low-dropout regulator with fast load regulation. *Analog Integr. Circ. Sign. Process.* 72, 2 (2012), 433–450. DOI: http://dx.doi.org/10.1007/s10470-012-9841-8

S. Lai, B. Yan, and P. Li. 2013. Localized stability checking and design of IC power delivery with distributed voltage regulators. *IEEE Trans. Comput.-Aid. Des. Integr. Circ. Syst.* 32, 9 (Sep. 2013), 1321–1334. DOI: http://dx.doi.org/10.1109/TCAD.2013.2256393

Ka Nang Leung and P. K. T. Mok. 2003. A capacitor-free CMOS low-dropout regulator with damping-factor-control frequency compensation. *IEEE J. Solid-State Circ.* 38, 10 (Oct 2003), 1691–1702. DOI: http://dx.doi.org/10.1109/JSSC.2003.817256

H. Li, X. Wang, J. Xu, Z. Wang, R. K. V. Maeda, Z. Wang, P. Yang, L. H. K. Duong, and Z. Wang. 2017. Energy-efficient power delivery system paradigms for many-core processors. *IEEE Trans. Comput.-Aid. Des. Integr. Circ. Syst.* 36, 3 (Mar. 2017), 449–462. DOI: http://dx.doi.org/10.1109/TCAD.2016.2584056

M. D. Pant, P. Pant, and D. S. Wills. 2002. On-chip decoupling capacitor optimization using architectural level prediction. *IEEE Trans. VLSI Syst.* 10, 3 (Jun. 2002), 319–326. DOI: http://dx.doi.org/10.1109/TVLSI.2002.1043335

M. Popovich, E. G. Friedman, R. M. Secareanu, and O. L. Hartin. 2007. Efficient placement of distributed on-chip decoupling capacitors in nanoscale ICs. In *Proceedings of the 2007 IEEE/ACM International Conference on Computer-Aided Design*. 811–816. DOI: http://dx.doi.org/10.1109/ICCAD.2007.4397365

Y. K. Ramadass, A. A. Fayed, and A. P. Chandrakasan. 2010. A fully-integrated switched-capacitor step-down DC-DC converter with digital capacitance modulation in 45nm CMOS. *IEEE J. Solid-State Circ.* 45, 12 (Dec. 2010), 2557–2565. DOI: http://dx.doi.org/10.1109/JSSC.2010.2076550

S. R. Sanders, E. Alon, H. P. Le, M. D. Seeman, M. John, and V. W. Ng. 2013. The road to fully integrated DC-DC conversion via the switched-capacitor approach. *IEEE Trans. Power Electron.* 28, 9 (Sep. 2013), 4146–4155. DOI: http://dx.doi.org/10.1109/TPEL.2012.2235084

X. D. S. Tan and C. J. R. Shi. 2001. Fast power/ground network optimization based on equivalent circuit modeling. In *Proceedings of the 38th Design Automation Conference*. 550–554. DOI: http://dx.doi.org/10.1145/378239.379021

M. K. Tavana, M. H. Hajkazemi, D. Pathak, I. Savidis, and H. Homayoun. 2015. ElasticCore: Enabling dynamic heterogeneity with joint core and voltage/frequency scaling. In *Proceedings of the 52nd Annual Design Automation Conference (DAC'15)*. ACM, New York, NY, Article 151, 6 pages. DOI: http://dx.doi.org/10.1145/2744769.2744833

Z. Toprak-Deniz, M. Sperling, J. Bulzacchelli, G. Still, R. Kruse, S. Kim, D. Boerstler, T. Gloekler, R. Robertazzi, K. Stawiasz, T. Diemoz, G. English, D. Hui, P. Muench, and J. Friedrich. 2014. Distributed system of digitally controlled microregulators enabling per-core DVFS for the POWER8 microprocessor. In *Proceedings of the 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC'14)*. 98–99. DOI:http://dx.doi.org/10.1109/ISSCC.2014.6757354

I. Vaisband and E. G. Friedman. 2016. Stability of distributed power delivery systems with multiple parallel on-chip LDO regulators. *IEEE Trans. Power Electron.* 31, 8 (Aug. 2016), 5625–5634. DOI:http://dx.doi.org/10.1109/TPEL.2015.2493512

I. Vaisband, R. Jakushokas, M. Popovich, A. V. Mezhiba, S. Köse, and E. G. Friedman. 2016. In *On-Chip Power Delivery and Management.* DOI:http://dx.doi.org/10.1007/978-3-319-29395-0_5

N. Viswanathan, C. J. Alpert, C. Sze, Z. Li, G.-J. Nam, and J. A. Roy. 2011. The ISPD-2011 routability-driven placement contest and benchmark suite. In *Proceedings of the International Symposium on Physical Design (ISPD'11).* ACM, New York, NY, 141–146. DOI:http://dx.doi.org/10.1145/1960397.1960429

K. Wang and M. Marek-Sadowska. 2005. On-chip power-supply network optimization using multigrid-based technique. *IEEE Trans. Comput.-Aid. Des. Integr. Circ. Syst.* 24, 3 (Mar. 2005), 407–417. DOI:http://dx.doi.org/10.1109/TCAD.2004.842802

L. Wang, S. K. Khatamifard, U. R. Karpuzcu, and S. Köse. 2018. Mitigation of NBTI induced performance degradation in on-chip digital LDOs. In *Proceedings of the IEEE Design, Automation and Test in Europe Conference and Exhibition (DATE'18).*

L. Wang, S. K. Khatamifard, O. A. Uzun, U. R. Karpuzcu, and S. Köse. 2017. Efficiency, stability, and reliability implications of unbalanced current sharing among distributed on-chip voltage regulators. *IEEE Trans. VLSI Syst.* 25, 11 (Nov. 2017), 3019–3032. DOI:http://dx.doi.org/10.1109/TVLSI.2017.2742944

T. Yu and M. D. F. Wong. 2014. Efficient simulation-based optimization of power grid with on-chip voltage regulator. In *Proceedings of the 2014 19th Asia and South Pacific Design Automation Conference (ASP-DAC'14).* 531–536. DOI:http://dx.doi.org/10.1109/ASPDAC.2014.6742946

Z. Zeng, S. Lai, and P. Li. 2013. IC power delivery: Voltage regulation and conversion, system-level cooptimization and technology implications. *ACM Trans. Des. Autom. Electron. Syst.* 18, 2, Article 29 (April 2013), 21 pages. DOI:http://dx.doi.org/10.1145/2442087.2442100

Z. Zeng, X. Ye, Z. Feng, and P. Li. 2010. Tradeoff analysis and optimization of power delivery networks with on-chip voltage regulation. In *Proceedings of the Design Automation Conference.* 831–836.

X. Zhan, P. Li, and E. Sanchez-Sinencio. 2016. Distributed on-chip regulation: Theoretical stability foundation, over-design reduction and performance optimization. In *Proceedings of the 2016 53nd ACM/EDAC/IEEE Design Automation Conference (DAC'16).* 1–6. DOI:http://dx.doi.org/10.1145/2897937.2898008

R. Zhang, K. Mazumdar, B. H. Meyer, K. Wang, K. Skadron, and M. R. Stan. 2015. Transient voltage noise in charge-recycled power delivery networks for many-layer 3D-IC. In *Proceedings of the 2015 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED'15).* 152–158. DOI:http://dx.doi.org/10.1109/ISLPED.2015.7273506

P. Zhou, A. Paul, C. H. Kim, and S. S. Sapatnekar. 2014. Distributed on-chip switched-capacitor DC-DC converters supporting DVFS in multicore systems. *IEEE Trans. VLSI Syst.* 22, 9 (Sep. 2014), 1954–1967. DOI:http://dx.doi.org/10.1109/TVLSI.2013.2280139