An Algorithm for Optimal Decoupling Capacitor Sizing and Placement for Standard Cell Layouts [*]

Haihua Su
IBM ARL
11501 Burnet Rd.
Austin, TX 78758
haihua@us.ibm.com

Sachin S. Sapatnekar
ECE Dept, Univ of Minnesota
200 Union St. SE
Minneapolis, MN 55455
sachin@ece.umn.edu

Sani R. Nassif
IBM ARL
11501 Burnet Rd.
Austin, TX 78758
nassif@us.ibm.com

ABSTRACT

With technology scaling, the trend for high performance integrated circuits is towards ever higher operating frequency, lower power supply voltages and higher power dissipation. This causes a dramatic increase in the currents being delivered through the on-chip power grid and is recognized in the International Technology Roadmap for Semiconductors as one of the difficult challenges. The addition of decoupling capacitances (decaps) is arguably the most powerful degree of freedom that a designer has for power-grid noise abatement and is becoming more important as technology scales. In this paper, we propose and demonstrate an algorithm for the automated placement and sizing of decaps in ASIC-like circuits. The adjoint sensitivity method is applied to calculate the first-order sensitivity of the power grid noise with respect to every decap. We propose a fast convolution technique based on piecewise linear (PWL) compressions of the original and adjoint waveforms. Experimental results show that power grid noise can be significantly reduced after a judicious optimization of decap placement, with little change of the total chip area.

Categories and Subject Descriptors

B.8.2 [**Performance and Reliability**]: Performance Analysis and Design Aids

General Terms

Algorithms

Keywords

decoupling capacitor, placement, optimization, power grid noise, adjoint sensitivity, ASICs

1. INTRODUCTION AND MOTIVATION

Noise margins have been greatly reduced in modern designs due to the lowering of supply voltages and the presence

Year	L_{eff} nm	f MHz	V_{dd} V	Size mm^2	Power W	Density W/mm^2
2001	65	1684	1.1	310	130	0.42
2002	53	2317	1.0	310	140	0.45
2003	45	3088	1.0	310	150	0.48
2004	37	3990	1.0	310	160	0.52
2005	32	5173	0.9	310	170	0.55
2006	28	5631	0.9	310	180	0.58
2007	25	6739	0.7	310	190	0.61

Table 1: IC technology parameters.

of a larger number of potential noise generators that eat significantly into the noise margins built into a design. The power grid provides the V_{dd} and ground signals throughout the chips, which are among the most important signals to control reliably, since supply voltage variations can lead not only to problems related to spurious transitions in some cases, particularly when dynamic logic is used, but also to delay variations [3] and timing unpredictability. Even if a reliable supply is provided at an input pin of a chip, it can deteriorate significantly within the chip due to the fact that the conductors that transmit these signals throughout the chip are electrically imperfect.

A powerful technique for overcoming this problem is through the use of on-chip decoupling capacitors (decaps) that are intentionally attached to the power grid. To exemplify the role of decaps, let us consider the circuit shown in Fig. 1, which can be thought of as a canonical model of a power grid and loading circuit. In the figure, G_g models the grid conductance, G_d and C_d model a decoupling capacitance, and I_{load} models the time-dependent current waveform of the load, which we model for simplicity as:

$$I_{load} = \begin{cases} 0 & : & t < 0 \\ \mu t & : & t < t_p \\ \mu(2t_p - t) & : & t < 2t_p \\ 0 & : & t > 2t_p \end{cases} \quad (1)$$

We will use data from the International Technology Roadmap for Semiconductors [14], summarized in Table 1, to predict the dependence of the load voltage V_{load} on the various circuit parameters in order to predict trends in power-grid-induced noise with technology scaling. The table shows the projected yearly trends for the effective length L_{eff}, of a transistor, the circuit frequency, f, the supply voltage level, V_{dd}, the chip size, the power dissipation and the density of power dissipation per unit area.
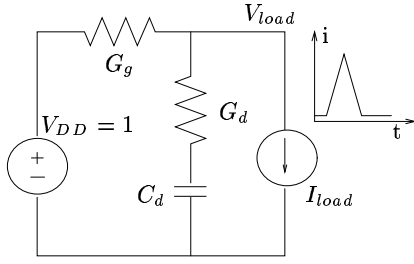
Figure 1: A canonical and approximate circuit representation of a power network.

For the circuit shown in Fig. 1, we observe that V_{load} normalized by the voltage supply V_{dd} over the time interval from $t = 0$ to $t = t_p$ can be expressed as:

$$V_{load} = 1 - \frac{\mu}{G_g}\left(t - \frac{C_d}{G_g}(1 - e^{-t/\tau})\right) \qquad (2)$$

where

$$\tau = \frac{(G_g + G_d)C_d}{G_g G_d} \qquad (3)$$

The minimum V_{load}, or maximum normalized power-supply-induced noise occurs at $t = t_p$ and the magnitude of the noise is:

$$V_{max} = \frac{\mu}{G_g}\left(t_p - \frac{C_d}{G_g}(1 - e^{-t_p/\tau})\right) \qquad (4)$$

We note that $t_p \propto f^{-1}$, and that power density $P_\square \propto V_{dd}\mu t_p$, implying that $\mu \propto P_\square f/V_{dd}$. Based on the trends in Table 1, f increases by 3.0X through the table, and μ increases by 9.13X. In order to keep V_{max} the same (i.e., keep the same amount of noise as a percentage of V_{dd}), we need to dramatically increase the last term in Eq. (4): $\frac{C_d}{G_g}(1 - e^{-t_p/\tau})$. This means:

- Increasing the decoupling capacitance C_d, which can be done at the cost of small additional area, because the area efficiency of decoupling capacitance is expected to increase as the gate oxide is scaled.

- Increasing the conductance associated with the decoupling capacitance G_d, which can be done by placing the capacitance *closer* to the load.

- Increasing the grid conductance G_g, which will be the most difficult to do because it goes somewhat against the prevailing scaling of interconnect, and the increased restrictions due to the consequent wire congestion emanating from this.

Unless we are able to do all of the above, it is likely that we will find the relative magnitude of power-grid-induced noise more than doubling by 2007.

The first two of these conclusions point a convincing finger towards the use of *appropriately placed* decaps for power grid noise reduction. While the use of decaps is certainly not new[1], the complexity of the problem requires shrewd opti-

---

[1] For example, in a 300MHz CMOS RISC Microprocessor design [4], as much as 160nF of on-chip decoupling capacitance is added to control power-supply noise. In another example [9], the on-chip decoupling capacitance is sized at ten times that of the total active circuit switching capacitance.

mal strategies driven by CAD tools, particularly in standard-cell environments in designs that require quick turn-around times in the face of strong time-to-market pressures.

Previous work [2, 5, 17] on decap allocation and optimization has focused on application in full custom design styles. A decap optimization procedure involving an iterative process of circuit simulation and floor planning is proposed in [5]. A linear programming technique is applied in [17] for allocation of white space for decap use and a heuristic is proposed to insert additional white space into an existing floorplan. Both [2] and [15] propose a sensitivity-based method of placing or optimizing decaps for reducing the voltage droop in the power distribution network; the former method handles the problem in the frequency domain, the latter in the time domain.
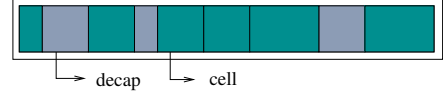

Figure 2: One row of cells in a standard cell layout showing decap locations.

In this work, we investigate the decap optimization and placement issue in the context of row-based standard-cell design typical of Application Specific Integrated Circuits (ASIC) where each row has a fixed height. We consider a chip composed of N rows, with the ith row having M_i cells (blocks). Each of the N rows is filled by cells to some level of ratio $r_i(\leq 100\%)$. Decoupling capacitors can be placed in the empty space, which forms the $(1 - r_i)$ fraction of each row. One such row is illustrated in Fig. 2.

Our approach is designed to be applicable subsequent to the placement phase for the design, where cells have already been assigned to rows. Since placement is designed to optimally place cells in order to achieve compactness for the layout and to control the wire length, timing and congestion, we use that result as the starting point for decap optimization, and perturb that solution in a minimal way in solving the *decap placement* problem. Because of this minimal perturbation, the following global and detailed routing results are expected to be slightly affected.

Specifically, we propose to use the empty spaces that may be available within each row (when $r_i < 1$) to place decaps. In doing so, the exact position of each cell in that row is considered to be flexible although the order and the *relative* positions are fixed. Different placement of cells can lead to different widths and location of decaps, and consequently different impacts on the power supply noise, and the problem that we wish to tackle is that of finding the optimal cell placement which results in the minimization of a metric for the power supply noise. Note that since typical values of r_i are close to 1, the major attributes of the original cell placement will be, for the most part, unaffected by our procedure.

The contributions of this work are as follows. We propose a nonlinear programmng based decap optimization scheme for standard-cell designs. A fast convolution technique based on piecewise linear (PWL) compressions of waveforms for the adjoint sensitivity computation is presented.

2. POWER SUPPLY NOISE METRIC AND ITS SENSITIVITY ANALYSIS

2.1 Modeling and analysis

For the ASIC row-based standard-cell design style outlined above, it is common to use a predefined mesh-like power distribution network. As in [5] and [6], we model the network as follows:

- The power distribution network (grid) is modeled as a resistive mesh

- The cells are modeled as time-varying current sources connected between power and ground. Each current source waveform is obtained from other tools that determine the worst-case patterns. Various work on worst-case current estimations can be found in [1, 11], etc.

- The decoupling capacitors are modeled as single lumped capacitors connected between power and ground.

- The top-level metal is connected to a package modeled as an inductance connected to an ideal constant voltage source.

The behavior of such a circuit is described by a first order differential equation formulates using modified nodal analysis (MNA) [13]. After the transient analysis of the circuit, the voltage waveform at every node is known. Given that the treatment for nodes on the ground grid is completely symmetric, we restrict our discussion to the V_{dd} nodes for which we define the *droop* at node n to be simply $V_{dd} - V_n(t)$, where $V_n(.)$ signifies the voltage at node n.
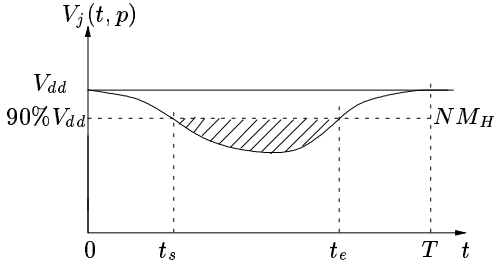

Figure 3: Illustration of the voltage droop at a given node in the V_{dd} power grid. The area of the shaded region corresponds to the integral z at that node.

An efficient metric to estimate power-grid-induced noise at a node is the integral of the voltage droop below a user specified noise ceiling [7]:

$$z_j(p) = \int_0^T \max\{NM_H - v_j(t,p), 0\}dt$$
$$= \int_{t_s}^{t_e} \{NM_H - v_j(t,p)\}dt, \qquad (5)$$

where p represents the tunable circuit parameters which, in our case, are the *widths* of the decoupling capacitors[2]. The voltage droop integral beyond the expressed by Eq. (5)

[2]We choose the width since the height of the decoupling capacitors is constrained to be the same as the height of the functional cells in the same row, as illustrated in Fig. 2.

represents the shaded area in Fig. 3. We define the measure of goodness for the whole circuit as the sum of the individual node metrics:

$$Z = \sum_{j=1}^K z_j(p), \qquad (6)$$

where K is the number of nodes. This metric penalizes more harshly transients that exceed the imposed noise ceiling by a large amount for a long time, and has empirically been seen to be more effective in practice than one that penalizes merely the maximum noise violation. Intuitively, this can be explained by the fact the the metric incorporates, in a sense, both the voltage and time axes together, as well as spatial considerations through the summation over all nodes in the circuit.

2.2 Integral sensitivity computation

Adjoint sensitivity analysis is a standard technique for circuit optimization where the sensitivity of one performance function with respect to many parameter values is required [8, 10, 13]. For our problem, the use of this method is a natural choice since we are interested in the sensitivity of the scalar objective function (Eq. (6)) with respect to the widths of all decaps in the network.

An adjoint network with the same topology as the original network is constructed, with all the voltage sources in the original network shorted and current sources open. For noise functions of the form given in Eq. (5), the adjoint network will include a current source of value $-u(t - t_s) + u(t - t_e)$ applied at node j if $z_j \neq 0$. We set the initial conditions to the adjoint circuit to zero and analyze it backward in time. We use the same time step h as the original circuit, thus allowing us to reuse the previously computed LU factorization for $(G + C/h)^{-1}$. Consequently, the extra simulation cost is reduced to one forward/backward solve for each time step of the adjoint circuit. Obviously, a smaller time step results in a higher accuracy for both the original and adjoint waveforms, and consequently higher accuracy in the sensitivities at the expense of a longer runtime. We find that in order to insure the accuracy of adjoint sensitivities, using 500-1000 steps per clock cycle (i.e. $h = 0.002T_{period}$ or $0.001T_{period}$) is sufficient.

The sensitivity of the objective function with respect to all of the decoupling capacitors in the circuit can be computed from the following convolution [8, 10]:

$$\frac{\partial Z}{\partial C} = \int_0^T \psi_C(T - t)\dot{v}_C(t)dt, \qquad (7)$$

where $\psi_C(\tau)$ is the waveform across the capacitor C in the adjoint circuit.

In our context, we cannot use this approach directly, and must tailor it to control the storage required by the direct application of this method. Specifically, a significant complication arises in the case of very large networks where the total amount of data to be stored is proportional to the number of nodes multiplied by the number of time steps, and could reach 10^9 bytes or more for large networks with millions of nodes. In order to alleviate the problem, we store the waveforms of the original and adjoint network using a compressed piecewise linear form. This results in a situation of the type illustrated in Fig. 4, where the time points on the original and adjoint waveforms are not aligned. However, since we know that waveforms are divided by linear segments, the
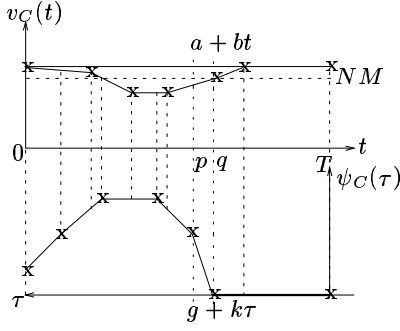
Figure 4: Compressed piecewise-linear waveforms for the original and adjoint networks.

convolution (Eq. (7)) of the waveforms $\psi_C(\tau) = g + k\tau$ and $v_C(t) = a + bt$ over the time interval $[p, q]$ can be expressed as:

$$\int_p^q (g + k(T - t)) \ \frac{d(a + bt)}{dt} \ dt$$

$$= \int_p^q (g + k(T - t)) \ b \ dt$$

$$= b(q - p) \left(g - kT - k \left[\frac{q - p}{2} \right] \right) \tag{8}$$

The complexity of the convolution calculation over $[0, T]$ is $O(N + M)$, where N and M are the number of linear segments on the original and adjoint waveforms.

Once the sensitivities of Z with respect to all of the decoupling capacitor values are computed, the sensitivities to the *width* of each capacitor can be calculated using the chain rule, as in [15]:

$$\frac{\partial Z}{\partial w} = \frac{\partial Z}{\partial C} \times \frac{\partial C}{\partial w} \tag{9}$$

Given that we calculate the decoupling capacitance from:

$$C = \frac{\varepsilon_{ox}}{T_{ox}} \times w \times h, \tag{10}$$

where T_{ox} and ε_{ox} are the thickness and permittivity of the gate oxide, and h is the fixed height of the decap, it is easily verified that Eq. (9) becomes:

$$\frac{\partial Z}{\partial w} = \frac{\partial Z}{\partial C} \times \frac{\varepsilon_{ox}}{T_{ox}} \times h \tag{11}$$

3. OPTIMIZATION AND PLACEMENT

3.1 Problem formulation

The problem of decoupling capacitor optimization is now formulated as:

Minimize	$Z(w_j)$	$j = 1 \cdots N_{decap}$
Subject to	$\sum_{k \in row_i} w_k \leq (1 - r_i)W_{chip}$	$i = 1 \cdots N_{row}$
and	$0 \leq w_j \leq w_{max}$	$j = 1 \cdots N_{decap}$

The scalar objective Z, defined in Eq. (6), is a function of all the decap widths and N_{decap} is the total number of decaps in the chip. The first constraint states that the total decap width in a row cannot exceed the total amount of empty space in that row, and W_{chip} and N_{row} denote, respectively, the width of the chip and the number of rows in the chip.

The second constraint restricts the decap widths within a realistic range. An upper bound w_{max} for a cell in row i is easily seen to be $(1 - r_i)W_{chip}$, which is the largest empty space in row i; while the lower bound of each decap width is zero.
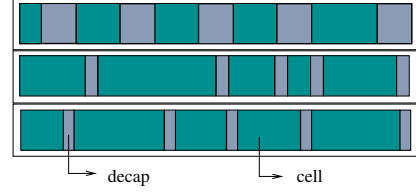

Figure 5: Illustration of the initial equal distribution of decaps.

Eq. (12) represents a linearly constrained nonlinear optimization problem. The objective function Z can be obtained after the transient analysis of the power grid circuit, and its sensitivity with respect to all the variables w_j can be calculated using the adjoint method discussed in Section 2.2. We choose to use a standard quadratic programming (QP) solver [18] for solving large nonlinear optimization problems. We start the optimization with an initial guess that uniformly distributes the vacant space in each row to each decoupling capacitor in each row, as illustrated in Fig. 5. It can be seen that initially there is one decap next to each cell. The initial chip width is chosen to be the maximum width occupied by cells and decaps among all rows.

3.2 Optimization and placement scheme

The optimization procedure invokes the QP optimizer, and the set of steps that are repeated during each iteration of the optimizer can be summarized as follows:

- Perform the transient simulation of the original power grid circuit and store piecewise linear waveforms of all decaps.

- Check all nodal voltages for those that fall below the noise margin, identify hot spots and compute the objective function Z.

- Set up the sources corresponding to these failure nodes for the adjoint circuit.

- Perform the transient simulation of the adjoint circuit and store piecewise linear waveforms of all decaps.

- Compute the sensitivities $\frac{\partial Z}{\partial C_j}$ by convolution and use the chain rule to obtain $\frac{\partial Z}{\partial w_j}$.

- Compute the constraint function and its Jacobian.

- Feed all the information into a QP solver and update the vector of widths, $\vec{w}$, according to the values returned by the solver.

- According to the updated $\vec{w}$, reposition all of the cells and decaps in the row from left to right.

4. EXPERIMENTAL RESULTS

The proposed decap optimization and placement scheme has been integrated into a linear circuit simulator written in C++ and the QP solver is applied. All experimental results are performed on a 1.8GHz Pentium IV machine under the Redhat Linux operating system.

Chp		Num bad nds	Num of nds	V_m (V)	Z ($V \times ns$)	Num of rows	Num of dcps	CPU time (min)
1	B	105	974	0.193	0.121	53	1964	0.9
	A	0		0.176	0.000			
2	B	80	861	0.230	0.366	85	3288	15.2
	A	63		0.196	0.063			
3	B	100	828	0.222	0.649	132	3664	12.5
	A	70		0.201	0.200			

B = Before optimization; A = After optimization

Table 2: Optimization results

Table 2 lists the decap optimization results for three industrial ASIC designs, which are referred to as Chip1, Chip2 and Chip3. Each of them is a $0.18\mu m$ CMOS design operating under a supply voltage of $1.8V$. The occupancy ratio r_i for each row of these chips is around 80%. Initially, decaps are uniformly distributed across each row. In Table 2, the second column shows the number of nodes with noise violations (i.e., nodes j with a nonzero value of z_j) before and after optimization; the total number of nodes in the power grid are shown in the third column. Although the power grid size of each chip is small, we emphasize that in a hierarchical design style the whole chip is divided into smaller functional modules and the decap optimization of each module can be performed individually because the noise suppression effect of decaps is very localized. The next two columns compare the worst-case voltage droop and the sum of integral area Z (i.e., the objective function) before and after optimization. Of the three examples, the worst-case (chip 3) noise (Z) reduction is around one-third of the initial value, which corresponds to the uniform distribution of decaps. Column 6 shows the total number of rows in the chip. The total number of decoupling capacitors placed in the whole chip is listed in column 7. Finally, the last column lists the total amount of CPU time to run each example. For each of these three chips, the worst case voltage droops and sums of the integral area are both reduced successfully.

It should be noted that decoupling capacitor is not the only method for noise reduction, and that other techniques such as wire widening, or increasing the density of the power grid, can be applied to further improve the power grid performance. Therefore, these results that holistically reduce the degree of noise violation by decap placement can be complemented with other techniques to obtain a solution that satisfies the noise constraints imposed on the design.

The V_{dd} and ground contour of chip2 is shown in Fig. 6 and Fig. 7. The small ovals in each figure represent VDD or GND c4 locations. In both figures, each gray-scaled color corresponds a voltage droop range and the number written in each color sample shows the lowest voltage droop in that range. Darker colors mean larger voltage droops. It can be seen that the voltage range in the V_{dd} plane is $1.610 - 1.8V$ and the hot spot is located on the right side of the chip. Similarly, the voltage range in the ground plane is $0 - 0.230V$, and the hot spot is located on the left side of the chip. The
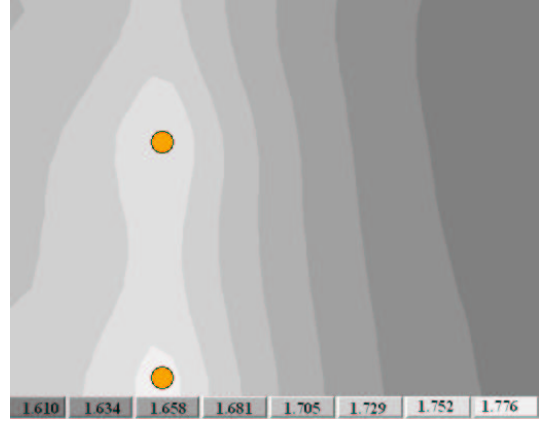

Figure 6: The original voltage droop contour of the Vdd plane.
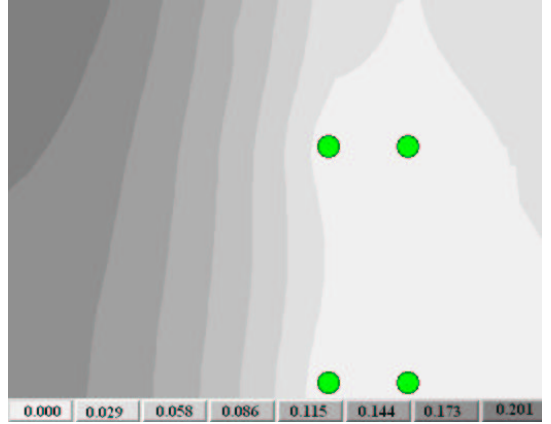

Figure 7: The original voltage droop contour of the ground plane.

result of the optimal cell and decap placement for chip2 is shown in Fig. 8. We observe that this placement is consistent with the hot spots of the chip, i.e., larger decaps are allocated closer to the two sides of the chip. After optimization, the voltage droop in the V_{dd} plane is in the range of $0 - 0.196V$ and that of the ground plane is in the range of $0 - 0.191V$. The optimization process has judiciously balanced the power grid voltage droop on the whole chip.

The noise reduction trend with respect to the cell occupancy ratio r_i for chip2 is shown graphically in Fig. 9. This experiment is performed by removing some cells from each row of the chip to achieve the desired occupancy ratio. For each case, around 10 percent of the total grid nodes are beyond the noise margin. A chip with lower occupancy ratios provides more empty space for decoupling capacitors and consequently is easier to optimize. Therefore, in Fig. 9, the noise reduction is more efficient for cases with lower occupancy ratios than for those with higher ones.

5. CONCLUSION

This paper has presented an on-chip decoupling capacitor sizing and placement scheme aimed at making the best use of empty spaces in the row-based standard-cell design of
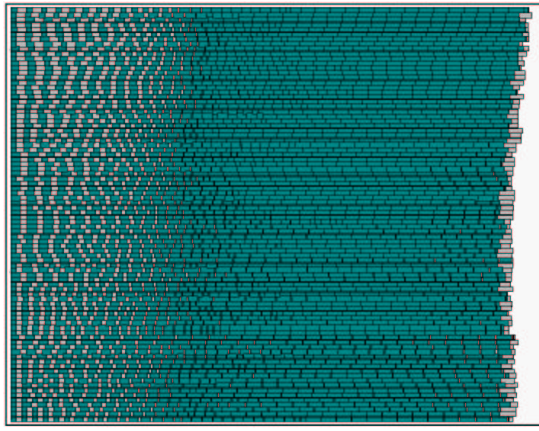
Figure 8: Results of the decap placement algorithm on chip2. The dark regions represent the standard cells, and the light regions are the decaps.
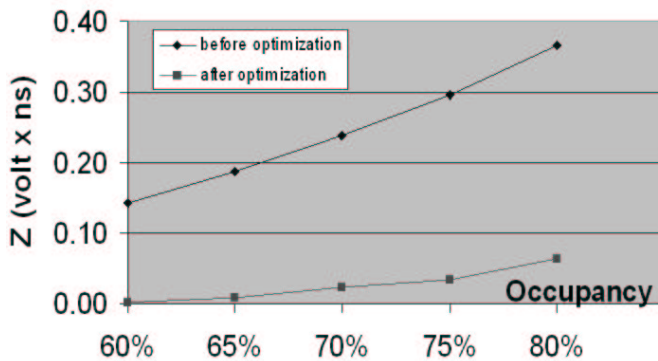

Figure 9: Variation of the noise metric with the occupancy ratio.

ASICs. The problem of decap insertion and placement has been motivated for current and future technologies, and the problem has been formulated as a constrained non-linear optimization problem that is successfully solved using the gradient-based QP solver. For a pre-designed power distribution network, the location and size of each decap is updated iteratively such that the total transient noise in the power grid is minimized, and the technique is demonstrated on several industrial designs.

6. REFERENCES

[1] G. Bai, S. Bobba, and I. N. Hajj. Emerging Pwer Mnagement Tols for Processor Design. In *Proc. International Symposium on Low Power Electronics and Design*, pages 143–148, Monterey, CA, August 1998.

[2] G. Bai, S. Bobba, and I. N. Hajj. Simulation and Optimization of the Power Distribution Network in VLSI Circuits. In *Proc. International Conference on Computer-Aided Design*, pages 481–486, San Jose, CA, November 2000.

[3] G. Bai, S. Bobba, and T. N. Hajj. Static Timing Analysis Including Power Supply Noise Effect on Propagation Delay in VLSI Circuits. In *Proc. Design Automation Conference*, pages 295–300, Las Vegas, NV, June 2001.

[4] W. J. Bowhill, R. L. Allmon, and S. L. Bell. A 300MHz 64b Quad-Issue CMOS RISC Microprocessor. In *Proc. International Solid-State Circuits Conference*, pages 182–183, Piscataway, NJ, February 1995.

[5] H. H. Chen and D. D. Ling. Power Supply Noise Analysis Methodology for Deep-Submicron VLSI Chip Design. In *Proc. Design Automation Conference*, pages 638–643, Anaheim, CA, June 1997.

[6] H. H. Chen and J. S. Neely. Interconnect and Circuit Modeling Techniques for Full-Chip Power Supply Noise Analysis. *IEEE Transactions on Components, Packaging, and Manufacturing Technology, Part B*, 21(3):209–215, August 1998.

[7] A. R. Conn, R. A. Haring, and C. Visweswariah. Noise Considerations in Circuit Optimization. In *Proc. International Conference on Computer-Aided Design*, pages 220–227, San Jose, CA, November 1998.

[8] S. W. Director and R. A. Rohrer. The Generalized Adjoint Network and Network Sensitivities. *IEEE Transactions on Circuit Theory*, 16(3):318–323, August 1969.

[9] D. W. Dobberpuhl, R. T. Witek, and R. Allmon. A 200-MHz 64-b Dual-Issue CMOS Microprocessor. *IEEE Journal of Solid-State Circuits*, 27(11):1555–1567, November 1992.

[10] P. Feldmann, T. V. Nguyen, S. W. Director, and R. A. Rohrer. Sensitivity Computation in Piecewise Approximation Circuit Simulation. *IEEE Transactions on Computer-Aided Design of ICs and Systems*, 10(2):171–183, February 1991.

[11] H. Kriplani, F. Najm, and I. Hajj. Maximum Current Estimation in CMOS Circuits. In *Proc. Design Automation Conference*, pages 2–7, Anaheim, CA, June 1992.

[12] S. R. Nassif and J. N. Kozhaya. Fast Power Grid Simulation. In *Proc. Design Automation Conference*, pages 156–161, Los Angeles, CA, June 2000.

[13] L. T. Pillage, R. A. Rohrer, and C. Visweswariah. *Electronic and System Simulation Methods*. McGraw-Hill, New York, NY, 1995.

[14] Semiconductor Industry Association, $http://public.itrs.net/Files/2001ITRS/Home.html$. *The International Technology Roadmap for Semiconductors*, 2001.

[15] H. Su, K. H. Gala, and S. S. Sapatnekar. Fast Analysis and Optimization of Power/Ground Networks. In *Proc. International Conference on Computer-Aided Design*, pages 477–480, San Jose, CA, November 2000.

[16] M. Zhao, R. V. Panda, S. S. Sapatnekar, T. Edwards, R. Chaudhry, and D. Blaauw. Hierarchical Analysis of Power Distribution Networks. In *Proc. Design Automation Conference*, pages 481–486, Los Angeles, CA, June 2000.

[17] S. Zhao, K. Roy, and C.-K. Koh. Decoupling Capacitance Allocation for Power Supply Noise Suppression. In *Proc. International Symposium on Physical Design*, pages 66–71, Napa, CA, April 2001.

[18] C. Zhu, R. H. Byrd, and J. Nocedal. *LBFGS-B: Fortran subroutines for large-scale bound constrained optimization*. EECS Department, Northwestern University, 1994.