Short Papers

Efficient In-Package Decoupling Capacitor Optimization for I/O Power Integrity

Jun Chen and Lei He

***Abstract*—With high integration density of today's electronic system and reduced noise margins, maintaining high power integrity becomes more challenging for high performance design. Inserting decoupling capacitors is one important and effective solution to improve the power integrity. The existing decoupling capacitor optimization approaches meet constraints on input impedance. In this paper, we show that impedance metric leads to large overdesign and then develop a noise-driven optimization algorithm for decoupling capacitors in packages for power integrity. We use the simulated annealing algorithm to minimize the total cost of decoupling capacitors under the constraints of a worst case noise bound. The key enabler for efficient optimization is an incremental worst case noise computation based on fast Fourier transform over incremental impedance matrix evaluation. Compared to the existing impedance-based approaches, our algorithm reduces the decoupling capacitor cost by 3× and is also more than 10× faster even with explicit noise computation.**

***Index Terms*—Decoupling capacitor, design automation, impedance, power integrity, simultaneous switching noise.**

I. Introduction

Power integrity becomes increasingly important for the performance of integrated circuits with higher integration density and lower noise margins. Compromised power integrity may lead to logic and timing errors. Nowadays, integrated circuit chips operate at very high frequencies and consume a large amount of power. The number input/outputs of I/Os is ever increasing. A large number of I/Os lead to serious simultaneous switching noise (SSN). In this paper, we focus on decoupling capacitor optimization for power integrity of chip I/Os. Our method can be also used for decoupling capacitor optimization in other part of the power delivery system.

For package decoupling purposes, discrete decoupling capacitors are used. Each type of decoupling capacitor has a different equivalent serial capacitance (ESC), equivalent inductance (ESL), and equivalent resistance (ESR) [1]. Consequently, they have different effective frequency ranges and prices. The effectiveness of the decoupling capacitors also depends on its electrical environment and varies with locations. Therefore, the types and locations of the decoupling capacitors have to be optimized for most effective design with minimal cost.

The majority of existing work for in-package or on-board decoupling capacitor optimization is trial-and-error methods, such as [2] and [1], both of which are manual processes. Automatic optimization methods have also been presented. The authors of [3] use the partial element equivalent circuit model and model-order reduction techniques to compute the input impedance and then search for the optimal locations of the decoupling capacitors to minimize the impedance by gradient-based search. In [4], the authors use finite-difference time domain and fast Fourier transform (FFT) to obtain frequency-dependent Poynting vector, and decoupling capacitors are iteratively put at the port with maximum Poynting vector. However, in both papers, the decoupling capacitor value is fixed, and ESL or ESR is not considered.

The most comprehensive work on automatic optimization of package decoupling capacitors is [5]. The authors model the inductive effect of packages with susceptance (inverse of inductance) and extract a resistance–capacitance–susceptance model. Based on this model, a macromodel is built with a model-order reduction technique. Then based on the macromodel, a simulated annealing algorithm is developed to search for the optimal types of decoupling capacitors at given locations to minimize the cost under the constraint of a target impedance at chip I/O ports. Different types of decoupling capacitors with different ESC, ESL, and ESR are considered.

However, the approach is based on impedance metrics, which will lead to significant overdesign. For example, in Fig. 1, we show a case where the noise bound is met but the impedance bound is not. Fig. 1(a) shows that the target impedance is not met in most part of the frequency band. However, the noise bound has been met, as shown in Fig. 1(b). It is clear that the target impedance cannot capture the noise accurately and may cause overdesign.

In this paper, we directly use noise as the metric of SSN and develop an efficient noise model to optimize the location and types of decoupling capacitors. We consider a large number of ports to search for the optimal locations for decoupling capacitors. We assume that the impedance matrix is given and develop an efficient model to compute the new impedance matrix with one decoupling capacitor inserted or removed. The time complexity of our algorithm is $O(n^2)$ compared to $O(n^3)$ in the state-of-the-art existing work [6]. With impedance matrix and precharacterized switching current waveform, we use FFT to compute the noise waveform and obtain the worst case noise. Based on these models, we develop a simulated annealing algorithm to minimize the cost subject to the maximum noise constraint. The algorithm demonstrates good efficiency with a large number of ports. It finished a case with 93 ports in less than 7 min with 5881 iterations, which is more than 10× faster than previous work. We also compare our approach with impedance-based approach and show that impedance is not a good metric for noise, and impedance-based approach leads to overdesign. Compared to our noise-based approach, the impedance-based solution has 3× larger cost.

The rest of this paper is organized as follows. In Section II, we first discuss the electrical models for the package system. Then, we present the method to incrementally compute the impedance matrix. We also discuss the noise metric and optimization flow. In Section VI, we use simulated annealing algorithm to optimize the decoupling capacitors. We conclude this paper in Section VII.

II. Electrical Models

A. Package and Decoupling Capacitor Model

Packages for semiconductor chips often consist of multiple signal layers, power planes, and ground planes with dielectric in between. Metal signal traces connecting the chip I/O cells to the printed circuit board (PCB) traces are routed between planes, and package planes are stapled together with vias and connected to PCB by balls. We assume that the locations of chip I/O ports are known, and the possible locations for the decoupling capacitors are predefined. We extract the

Manuscript received May 2, 2006; revised August 15, 2006. This work was supported in part by the National Science Foundation under CAREER Award CCR-0401682 and SRC under Grant 1100, and by a UC MICRO Grant sponsored by Analog Devices, Intel, and RIO Design Automation. This paper was recommended by Guest Editor J. Hu.

The authors are with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA (e-mail: jchen@ee.ucla.edu; lhe@ee.ucla.edu).

Digital Object Identifier 10.1109/TCAD.2006.888262

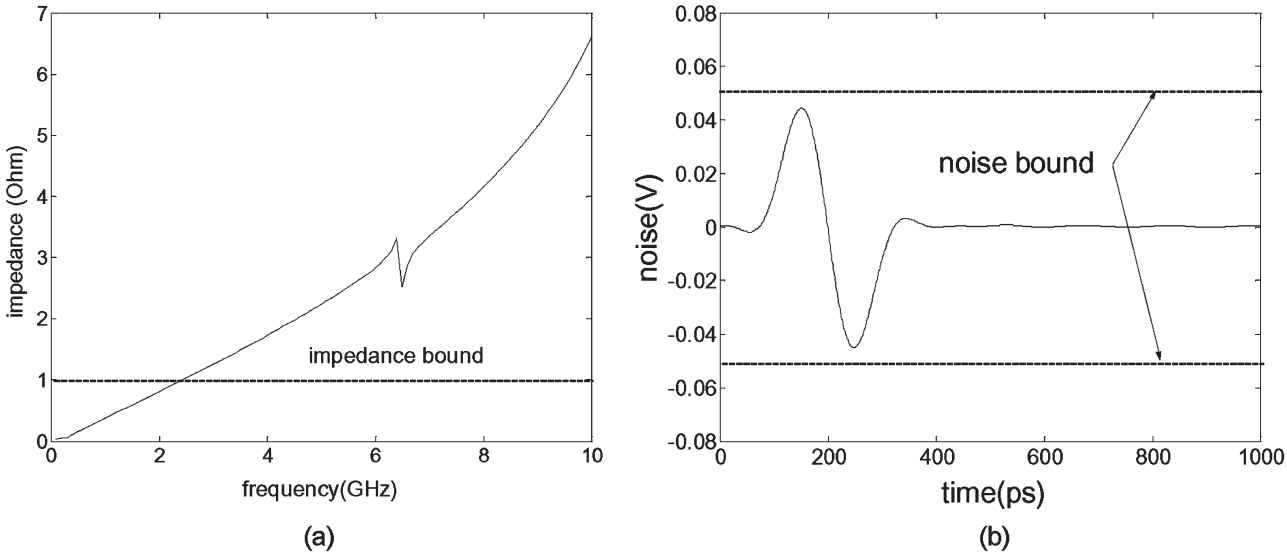

Fig. 1. (a) Impedance. (b) Noise waveform.

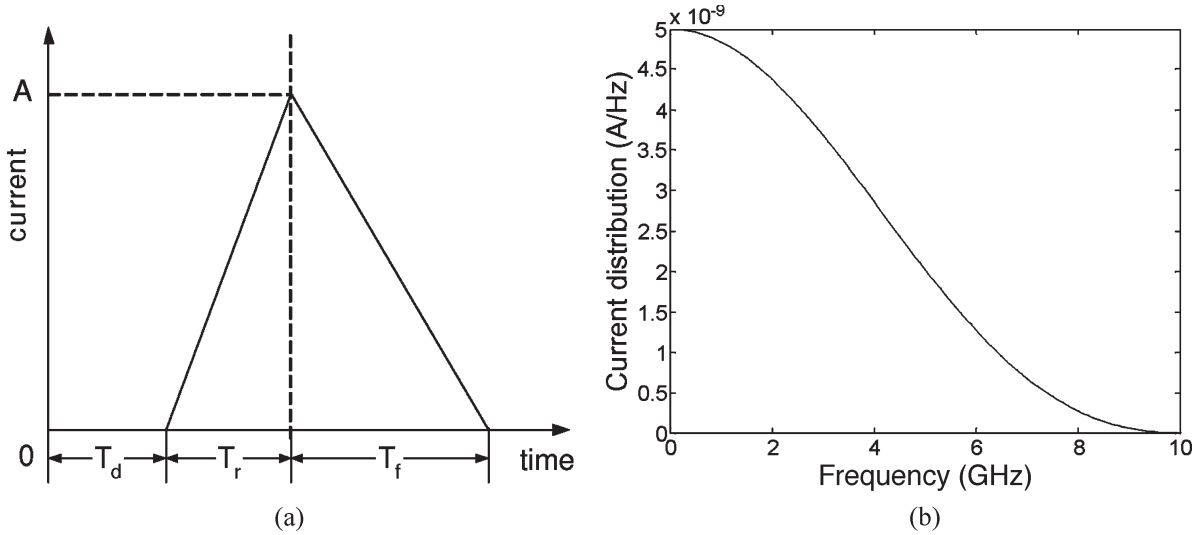

Fig. 2. Switching current model. (a) Time-domain waveform. (b) Spectrum.

macromodel of the package with the specified ports for I/Os and decoupling capacitors before the optimization process. Specifically, the macromodel we use in this paper is the impedance matrix $Z(f_k)$ at a number of sample frequencies f_k. With the macromodel, the efficiency of following optimization process no longer depends on the size of the original circuits but only depends on the number of ports defined. This allows a very complex package to be optimized in a very short time. In this paper, we first extract a detailed RLCK circuit of the package and then use a model-order reduction technique to obtain the impedance matrix. The frequency-dependent impedance Z can also be obtained by other methods, such as full-wave field solvers. The decoupling capacitors for the package are discrete elements. Each type of decoupling capacitors is modeled by ESC, ESL, and ESR. The frequency-dependent impedance at the sample frequencies is $Z_d(\omega) = \text{ESR} + 1/(\omega\text{ESC}) + j\omega\text{ESL}$.

B. Model Current of I/O Cells

Normally, each I/O cell drives a transmission line. When switching, it draws a current from the power delivery system and causes voltage fluctuation (SSN noise). For each type of the I/O drivers, the loading is often specified. Therefore, switching current profile is not random. In ideal cases, switching current can be easily obtained by simulation. However, in nonideal environments with noise and process variations, current can vary around the ideal current waveform. Considering the worst or tolerable design corner cases, designers can obtain the current causing the worst case noise or consider all the possible corner cases. In this paper, we assume that the worst case current profile has been obtained for each driver. Similarly to [7], for simplicity, we model the current waveform as a two-segment piecewise linear waveform (triangular waveform), as shown in Fig. 2(a).

III. Incremental Computation of Impedance

With the insertion or removal of decoupling capacitors, the impedance matrix of the system will change and affect the noise value. Therefore, the impedance matrix has to be updated with changes of decoupling capacitor distribution. In [5], this is done by n_{IO} ac sweeps, where n_{IO} is the number of I/O ports. Another method is presented in [6]. Assuming the macromodel without decoupling capacitors is given

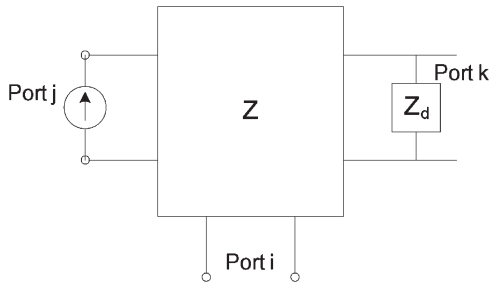

Fig. 3. Inserting one decoupling capacitor.

in terms of admittance matrix $Y(\omega)$, the impedance with decoupling capacitors is computed as

$$Z(\omega) = \left(Y(\omega) + \tilde{Y}(\omega)\right)^{-1} \tag{1}$$

where $\tilde{Y}(\omega)$ is a diagonal matrix with $\tilde{Y}_{ii}$ equal to the admittance of the decoupling capacitor at port i at frequency ω. Both of these methods need at least one matrix inversion, on which the computation time of this operation mainly depends. Because Y is a macromodel, it is usually a dense matrix, and the time complexity of the matrix inversion is roughly $O(n_p^3)$, where n_p is the number of ports including the I/O ports and the ports for the decoupling capacitors.

The approach above is good for computing impedance when simultaneously inserting or removing a large number of decoupling capacitors. However, in an iterative optimization process, we normally add or remove one or a small number of decoupling capacitors each time. In this case, matrix inversion is not necessary for impedance computation. In addition, the approaches above always compute all the Z_{ij} even if only a few are needed. In fact, we only need to compute Z_{ij} when necessary in iterative algorithm. We propose an efficient incremental method to compute each impedance element Z_{ij} separately with much less overall complexity.

To derive the model, we assume that, at a certain frequency, the impedance matrix before inserting the decoupling capacitor is Z, and we insert one decoupling capacitor at port k, as shown in Fig. 3. We need to solve the new impedance $\hat{Z}$. $\hat{Z}_{ij}$, which is the transfer impedance from port j to port i, is equal to the voltage at i when applying a 1-A current source at port j. With two-port network theory, the current running through the decoupling capacitor can be derived as $Z_{kj}/(Z_{kk} + Z_d)$, where Z_d is the impedance of the decoupling capacitor. Replacing the capacitor with a current source of the same current, according to the superposition principle

$$\hat{Z}_{ij} = Z_{ij} - \frac{Z_{ik}Z_{kj}}{Z_{kk} + Z_d}. \tag{2}$$

The overall impedance matrix with the decoupling capacitor added at port k at a given frequency is

$$\hat{Z} = Z - \frac{b_k a_k}{Z_{kk} + Z_d} \tag{3}$$

where a_k is the kth row of Z, and b_k is the kth column of Z. This is a rank-one updating [8]. The complexity of this process is $O(n_p^2)$. Similarly, the overall impedance matrix with the decoupling capacitor removed from port k at a given frequency is

$$\hat{Z} = Z - \frac{b_k a_k}{Z_{kk} - Z_d}. \tag{4}$$

Compared to (1), this method is obviously more efficient and scalable with the number of ports, when only one decoupling capacitor is added or removed. This is especially suitable for iterative optimization process or trial-and-error process, in which one or a small number of decoupling capacitors are changed, and the impedance matrix is needed to be reevaluated in each iteration. Another advantage of this method is that to obtain certain ports' impedance, we only need to selectively compute them with (2) without computing the impedance of other ports. This again is good for trial-and-error methods as will be explained in Section V. If n decoupling capacitors are changed, the computation in (3) needs to be repeated for n times. When $n \ll n_p$, it will still be more efficient than (1). The worst case is that $n = n_p$, which means that the distribution of decoupling capacitors changes at all the ports, and the complexity becomes $O(n_p^3)$, same as [6]. Fortunately, this case will never happen in one iteration.

IV. Noise Metric

A. Impedance Metric

Traditionally, for the integrity of power delivery system, the impedance at given ports is required to be lower than a computed target impedance in the entire frequency bandwidth of interest. According to [9], the target impedance can be computed as follows:

$$Z_t = \frac{\delta \mathrm{Vdd}}{I} \tag{5}$$

where δ is the tolerable variation of Vdd and I is the switching current at the given ports. However, the impedance is not directly proportional to the noise, and this kind of approach is pessimistic. In fact, the current is not uniformly distributed in the entire frequency band but generally decreases with increasing frequency. One example is shown in Fig. 2(b). In contrast, the impedance generally decreases with increasing frequency. Also, different frequency components have different amplitude and phase and may cancel each other. The impedance needs not to be very small in the entire frequency band. Large impedance at a lower frequency may cause large time-domain noise but may not cause problem at a higher frequency. One case has been shown in Fig. 1.

B. Time-Domain Metric

In this paper, we directly consider the noise in the power delivery system at each port of interest. For the noise at port i induced by the switching activity at port j, the noise component at the kth frequency sampling point can be easily computed as

$$V_{ij}(f_k) = Z_{ij}(f_k)I_j(f_k). \tag{6}$$

We then use FFT to compute the time-domain waveform. The time complexity of FFT is $O(n \log n)$, where n is the number of the sampling points. At a given port, we consider both the noise induced by the I/O cells connected to the port and the noise induced by the switching activity of I/O cells connected at other ports. Because the switching of the I/O cells is random and the system is linear, the worst case noise at one port is the sum of the maximum noises induced by all the cells.

V. Efficient Flow for Decoupling Capacitor Optimization

With imperfect decoupling capacitors and large inductive effects, the solution space of in-package decoupling capacitor optimization problem is strongly nonmonotonic. In addition, the decoupling capacitors are discrete elements. Mathematical programming methods generally used for on-chip decoupling capacitors are difficult to be applied in such case. With the proposed model, we propose an efficient generic

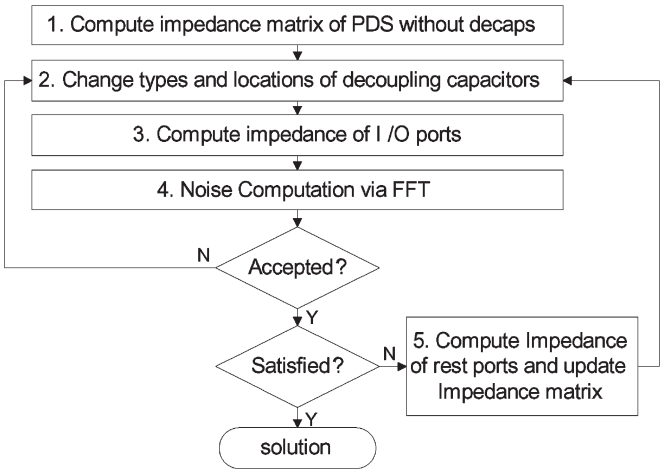

Fig. 4. Efficient decoupling capacitor optimization flow.

TABLE I
DECOUPLING CAPACITORS [5]

Type	1	2	3	4
ESC(nF)	50	100	50	100
ESR(Ω)	0.06	0.06	0.03	0.03
ESL(pH)	100	100	40	40
Price	1	2	2	4

flow for iterative decoupling optimization, as shown in Fig. 4. In this flow, we compute the SSN at the I/O drivers to determine whether the altered solution should be kept. If kept, the impedance matrix is further completely computed and updated. The iteration terminates when certain criterion is met. Since the SSN only depends on the impedance of I/O ports, we only need to compute these impedance elements to decide whether to accept the new solution. Therefore, in Step 3 of Fig. 4, we only compute the impedance of I/O ports, and the complexity of the computation is only $O(n_{\rm IO}^2)$, where $n_{\rm IO}$ is the number of I/O ports or the ports where the impedance needs to be controlled. The square in the formula is because we consider the coupling between these ports. If we accept the new solution based on the new SSN computed from impedance at I/O ports, the impedances of the rest ports (ports for decoupling capacitors) are computed further in step 6 of Fig. 4, and the complexity is $O(n_p^2 - n_{\rm IO}^2)$. If the number of I/O ports is much smaller than the total number of ports, this separation of computation can further save significant computation power. When $n_{\rm IO} \cong n_p$, the complexity is roughly $O(n_{\rm IO}^2)$. This flow is suitable for both automatic optimization and manual optimization. In addition, even without the noise computation via FFT, this flow can also greatly speed up impedance-based iterative optimization process.

VI. NOISE-DRIVEN OPTIMIZATION

A. Settings

In this section, we use the developed impedance and noise models to minimize the cost of the decoupling capacitors in a package under the constraint of noise in the power delivery system. The package is often cut into different domains for different supply voltages. We optimize each voltage domain separately. Similar to [5], we also try to minimize the total decoupling capacitor cost. We consider different types of decoupling capacitors with different prices. We assume the same set of decoupling capacitors as in [5], which are summarized in Table I. We assume that Vdd is 2.5 V and require the noise to be less than 15% of Vdd, which is 0.35 V.

TABLE II
WORST CASE NOISE AT PORTS

port	1	2	3
before optimization	2.52V	2.49V	2.48V
after optimization	0.344V	0.343V	0.344V

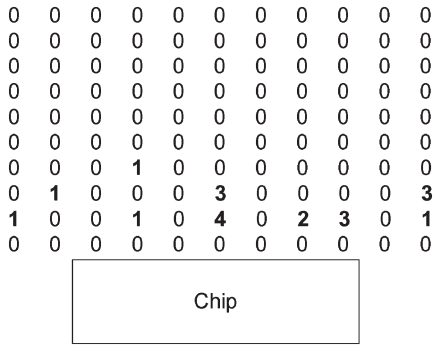

Fig. 5. Optimal distribution of decoupling capacitors from noise-driven approach.

B. Simulated Annealing Algorithm

We use the simulated annealing algorithm to optimize the types and locations of the decoupling capacitors so that the total cost is minimized and the noise in the power/ground plane is smaller than a given bound. The objective function is defined as

$$F(p_i, c_j) = \alpha \sum_{i \in \rm IO} p_i + \beta \sum_j c_j \tag{7}$$

where α and β are weights for the noise and cost, respectively. α is chosen to be much larger than β so that the noise constraint can be achieved. p_i is the penalty function for violation of the noise constraint.

C. Results

In this case, we assume 1×2 cm rectangular cut of a package with a power plane and a ground plane. I/O cells are located at one edge of the structure. We assume that there are 30 I/O cells. Each of them will draw the current shown in Fig. 2(b). Since cells close to each other have similar impedance and strongly couple to each other, we partition the 30 I/O cells into three groups and define three I/O ports. Each cell is connected to the closest I/O port, and each of the ports is connected with ten I/O cells. Note that for higher accuracy, more ports can be defined if necessary. We allow the decoupling capacitors to be distributed across the plane and therefore define 90 uniformly distributed ports on the package. In total, there are 93 ports in our macromodel.

Our noise-based algorithm found a valid solution where all the ports meet the noise constraint. The worst case noise of each port is listed in Table II. As we will show in Fig. 6, the resonance peaks have been effectively pushed to higher frequency, and impedance at low frequency has been reduced. Correspondingly, the large oscillation noise has been largely reduced. The total cost of the decoupling capacitors is 20. In Fig. 5, we show the distribution of the decoupling capacitors in a uniform grid. In this figure, the numbers stand for the type of decoupling capacitor, and "0" means no decoupling capacitor.

We further compare our results with an impedance-based approach. In this approach, for the objective function, we substitute the noise with the maximum impedance and replace the noise bound with the target impedance. Because we require the noise to be less than 0.35 V, and the total peak current of ten I/Os connected to one port is 500 mA,

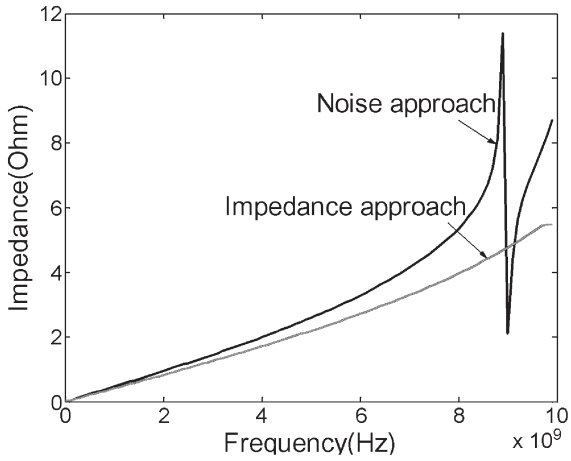

Fig. 6. Comparison of impedances after optimization.

TABLE III
IMPEDANCE AND NOISE AT PORTS

port	1	2	3	bound
maximum impedance	5.31Ω	5.59Ω	7.12Ω	0.7Ω
worst-case noise	0.256V	0.302V	0.284V	0.35V

TABLE IV
RUNTIME

approach	1	2	3
ports	93	93	20
iterations	5881	5403	1920
run time(s)	389.5	4156.1	2916.0
avg. run time(s)	0.0662	0.7692	1.519

the target impedance for each port is 0.7 Ω. We can see that the decoupling capacitors still concentrate around the chip but spread more across the planes than noise-driven approach. The total cost is 72, which is more than 3× larger than the results of noise-driven approach. In Fig. 6, the input impedance at one port is compared to the impedance at the same port from the noise-based approach. We can see that with more decoupling capacitors, the impedance-based approach further reduces the impedance. However, such low impedance is not necessary for the target noise bound. In Table III, we summarize the maximum impedance and the worst case noise at each port. We can see that the target impedance cannot be reached, but the noise is already well below the noise bound. This shows that using impedance as a noise metric will lead to large overdesign.

D. Runtime

We implement the algorithms in Matlab and conduct experiments on a 2.8-GHz Xeon system. For comparison, we also implemented the method of (1). The runtime of different methods is shown in Table IV.[1] In the table, Method 1 is the proposed method using the proposed incremental impedance computation and FFT for noise computation. Method 2 uses the impedance computation method from [6] and FFT for noise computation. Method 3 is from [5]. By comparing methods 1 and 2, we can see that the incremental computation of impedance is 11× faster than the matrix inversion-based approach. Comparing Methods 1 and 3, our method is significantly faster than Method 3 even considering the speed difference of the computing platforms and with more ports. We can see that the models and algorithm can handle a large number of ports and can be readily used for optimization of real designs.

[1]The runtime of Method 3 in Table IV is taken from [5]. The computation platform is 1-GHz Pentium III, and the computing language is unknown.

VII. DISCUSSION AND CONCLUSION

We studied the optimization of decoupling capacitors for package power integrity. Traditionally, impedance is used as a noise metric. However, this approach is also based on certain assumption of current waveform to determine the effective frequency range over which the impedance bound is applied. Commonly, the frequency range is determined from signal rising time. It still cannot consider all kinds of current waveform and has the difficulty to determine the effective frequency range. Too pessimistic an estimation of frequency range can lead to large overdesign. To obtain a tight frequency range and avoid large overdesign, the worst case current profile should be determined and noise metric should be used.

In this paper, we used time-domain noise as the metric to guide the optimization. To do this, we developed an efficient worst case noise model. We first developed an efficient method to compute the port impedance incrementally with changes in decoupling capacitor configuration. The complexity of the method is only $O(n^2)$ compared to previous work's $O(n^3)$ complexity. Based on the impedance, we then computed the noise with FFT. We further developed a simulated annealing algorithm to minimize the cost of the decoupling capacitors under the constraints of worst case noise. Experiments showed that our algorithm demonstrates good efficiency with large number of ports. Compared to previous work, we gained more than 10× speedup. The cost of the solution from our noise-based approach is 3× smaller than the cost from the solution of the impedance-based approach. In this paper, we mainly consider the SSN at I/O drivers, but the SSN in the entire package can be easily considered with additional probing ports. We assumed that the worst case current is given in this paper. We will develop the methodology to determine the worst case current waveform for I/O drivers.

REFERENCES

[1] X. Yang, Q. Chen, and C. Chen, "The optimal value selection of decoupling capacitors based on FDFD combined with optimization," in *Proc. IEEE Top. Meeting Elect. Perform. Electron. Packag.*, 2002, pp. 191–194.
[2] Y. Chen, Z. Chen, and J. Fang, "Optimum placement of decoupling capacitors on packages and printed circuit boards under the guidance of electromagnetic field simulation," in *Proc. Electron. Compon. Technol. Conf.*, 1996, pp. 756–760.
[3] A. Kamo, T. Watanabe, and H. Asai, "An optimization method for placement of decoupling capacitors on printed circuit board," in *Proc. IEEE Top. Meeting Elect. Perform. Electron. Packag.*, 2000, pp. 73–76.
[4] I. Hattori, A. Kamo, T. Watanabe, and H. Asai, "A searching method for optimal locations of decoupling capacitors based on electromagnetic field analysis by FDTD method," in *Proc. IEEE Top. Meeting Elect. Perform. Electron. Packag.*, 2002, pp. 159–162.
[5] H. Zheng, B. Krauter, and L. Pileggi, "On-package decoupling optimization with package macromodels," in *Proc. Custom Integr. Circuits Conf.*, 2003, pp. 723–726.
[6] J. Zhao and O. P. Mandhana, "A fast evaluation of power delivery system input impedance of printed circuit boards with decoupling capacitors," in *Proc. IEEE Top. Meeting Elect. Perform. Electron. Packag.*, 2004, pp. 111–114.
[7] H. Su, S. S. Sapatnekar, and S. R. Nassif, "Optimal decoupling capacitor sizing and placement for standard-cell layout designs," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 22, no. 4, pp. 428–436, Apr. 2003.
[8] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1983.
[9] L. Smith, R. Anderson, D. Forehand, T. Pelc, and T. Roy, "Power distribution system design methodology and capacitor selection for modern CMOS technology," *IEEE Trans. Adv. Packag.*, vol. 22, no. 3, pp. 284–291, Aug. 1993.