# Efficient Optimization of In-Package Decoupling Capacitors for I/O Power Integrity

Jun Chen and Lei He
Electrical Engineering Department,
University of California, Los Angeles

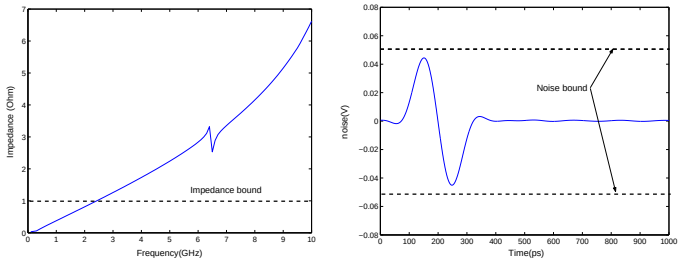
***Abstract*— With high integration density of today's electronic system and reduced noise margins, maintaining high power integrity becomes more challenging for high performance design. Inserting decoupling capacitors is one important and effective solution to improve the power integrity. The existing 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 uses 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 FFT 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.**

## 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, IC chips operate at very high frequencies and consume a large amount of power. The number of I/O's is ever increasing. A large number of I/O's lead to serious simultaneous switching noise (SSN). In this paper, we focus on decoupling capacitor optimization for power integrity of chip I/O's. Our method can be also used for decoupling capacitor optimization in other part of the power delivery system.

For package decoupling purpose, discrete decoupling capacitors are used. Each type of decoupling capacitor has 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.

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 PEEC 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 FDTD and 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.




(a) Impedance (b) Noise waveform

Fig. 1. Impedance and noise waveform

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(RCS) 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 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 can not capture the noise accurately and may cause overdesign.

In this work, 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 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 pre-characterized 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 large number of port. It finished a case with 93 ports in less than 7 minutes with 5881 iterations, which is more than $10\times$ 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 $3\times$ larger cost.

## 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 PCB traces are routed between planes, and package planes are stapled together with vias, and connected to PCB by balls. We assume the locations of chip I/O ports are known and the possible locations for the decoupling capacitors are predefined. We extract the macromodel of the package with the specified ports for I/O's and decoupling capacitors before the optimization process. Specifically, the macromodel we use in this work 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 $Z_d(\omega) = ESR + 1/(\omega ESC) + j\omega ESL$ .

### B. Model current of I/O cells

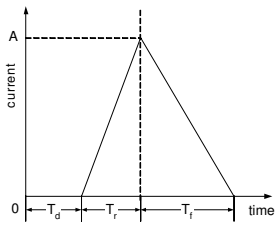



Fig. 2. Switching current model

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 are often specified. Therefore switching current profile is not random. In ideal cases, switching current can be easily obtained by simulation. However, in non-ideal 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 cause the worst case noise or consider all the possible corner cases. In this work, 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 piece-wise linear waveform (triangular waveform) as shown in Fig.2.

## 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 the $n_{io}$ is the number of I/O ports. Another method is presented in [6]. Assuming the macromodel without decoupling capacitors is given in terms of admittance matrix $Y(\omega)$, the impedance with decoupling capacitors is computed as,

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

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 $\omega$. 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 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 approach above always compute all the $Z_{ij}$ even if only a few is 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.

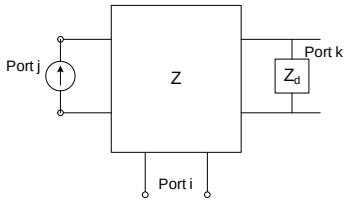



Fig. 3. Inserting one decoupling capacitor

To derive the model we assume 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 an 1A 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 $k$th row of $Z$ and $b_k$ is the $k$th 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 try and error method as will be explained later in section V.

If $n$ decoupling capacitors are changed, the computation in (3) needs to be repeated for $n$ times. When $n << n_p$, it will still be more efficient than (1). The worst case is that $n = n_p$, which means 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,

$$Z_t = \frac{\delta V dd}{I} \tag{5}$$

where, $\delta$ is 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 approaches is pessimistic. In fact, the current is not uniformly distributed in the entire frequency band but generally decreases with frequency increasing. One example is shown in Fig.4. In the contrast the impedance generally decreases with frequency increasing. 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. A 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.

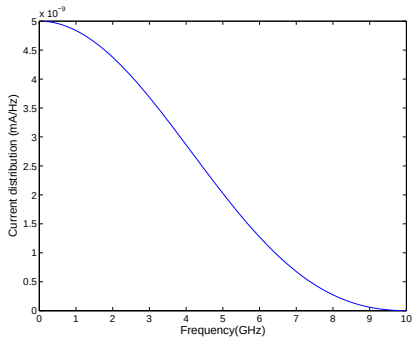



Fig. 4. Spectrum of a switching current

### 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 $k$th 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 Fast Fourier Transform (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 are 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 non-monotonic. In addition, the decoupling capacitors are discrete elements. Mathematical programming methods generally used for on-chip decoupling capacitors is difficult to be applied in such case. With the proposed model, we propose an efficient generic flow for iterative decoupling optimization as shown in Fig.5. 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 criteria 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.5 we only compute the impedance of I/O ports and the complexity of the computation is only $O(n_{IO}^2)$, where $n_{IO}$ is the number of I/O ports or the ports where the impedance needs to be

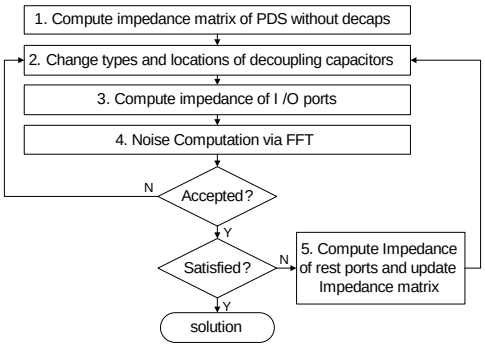



Fig. 5. Efficient decoupling capacitor optimization flow

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 |

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 impedance of the rest ports (ports for decoupling capacitors) are computed further in step 6 of Fig.5 and the complexity is $O(n_p^2 - n_{IO}^2)$. 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 $n_{IO} \cong n_p$, the complexity is roughly $O(n_{IO}^2)$. 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 the Vdd is 2.5V and require the noise to be less than 15% of Vdd, which is 0.35V.

### 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

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

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 |

where $\alpha$ and $\beta$ are weights for the noise and cost respectively. $\alpha$ is chosen to be much larger than $\beta$ 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 1cm×2cm 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.4. Since cells close to each other have similar impedance and strongly couple to each other, we partition the 30 I/O cells into 3 groups and define 3 I/O ports. Each cell is connected to the closest I/O port and each of the ports is connected with 10 I/O cells. Note 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. Totally, 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 later in this paper in Fig.7, the resonance peaks has 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.6, 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.

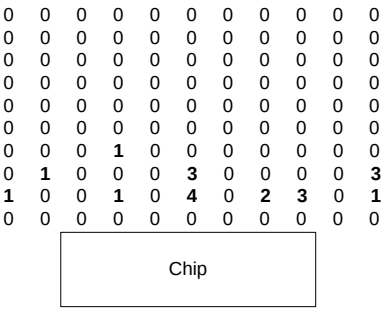



Fig. 6. Optimal distribution of decoupling capacitors from noise driven approach

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 less than 0.35V and the total peak current of 10 I/Os connected to one port is 500mA, the target impedance for each port is 0.7Ω. We can see that though the decoupling

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 |

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.7 the input impedance at one port is compared to the impedance at the same port from the noise based approach. We can see 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 the target impedance can not be reached but the noise is already well below the noise bound. This shows using impedance as a noise metric will lead to large overdesign.

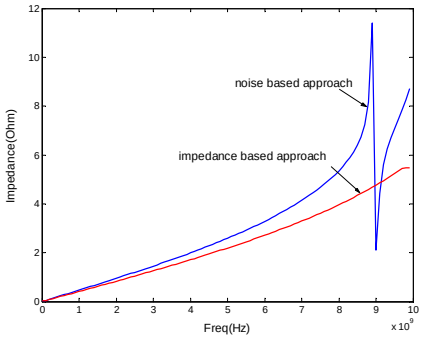



Fig. 7. Comparison of impedances after optimization

### D. Runtime

We implement the algorithms in Matlab and conduct experiments on a 2.8GHz 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 method 1 and 2, we can see that the incremental computation of impedance is 11× faster than the matrix inversion based approach. Comparing method 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.

## VII. DISCUSSION AND CONCLUSION

We studied the optimization of decoupling capacitors for package power integrity. Traditionally, impedance is used as 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 can not consider all kinds of current waveform and has the difficulty to determine the effective frequency range. Too pessimistic 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 work 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 the worst case current is given in this work. We will develop the methodology to determine the worst case current waveform for I/O drivers.

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 |

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

## REFERENCES

[1] X. Yang et al. , "The optimal value selection of decoupling capacitors based on FDFD combined with optimization," in *EPEP*, 2002.
[2] Y. Chen et al., "Optimum placement of decoupling capacitors on packages and printed circuit boards under the guidance of electromagnetic field simulation," in *ECTC*, 1996.
[3] A. Kamo et al., "An optimization method for placement of decoupling capacitors on printed circuit board," in *EPEP*, 2000.
[4] I. Hattori et al., "A searching method for optimal locations of decoupling capacitors based on electromagnetic field analysis by FDTD method," in *EPEP*, 2002.
[5] H. Zheng et al., "On-package decoupling optimization with package macromodels," in *CICC*, 2003.
[6] J. Zhao and O. P. Mandhana, "A fast evaluation of power delivery system input impedance of printed circuit boards with decoupling capacitors," in *EPEP*, 2004.
[7] H. Su et al., "Optimal decoupling capacitor sizing and placement for standard-cell layout designs," *TCAD*, vol. 22, pp. 428–436, 2003.
[8] G. H. Golub and C. F. V. Loan, *Matrix Computations*. John Hopkins University Press, 1983.
[9] L. Smith et al., "Power distribution system design methodology and capacitor selection for modern cmos technology," *IEEE Transactions on Advanced Packaging*, vol. 22, pp. 284–291, 1993.