Return Loss Optimization of the Microprocessor Package Vertical Interconnect

Arun V. Sathanur[1], Vikram Jandhyala[1],
Kemal Aygün[2], Henning Braunisch[2], and Zhichao Zhang[2]

[1]ACE Lab, Department of Electrical Engineering, University of Washington, Seattle, WA, USA
[2]Intel Corporation, Chandler, AZ, USA
E-mail: arunsv@u.washington.edu

Abstract

A geometry based parametric model of a differential high-speed line traversing a ten-layer microprocessor package is developed. This model is used to undertake a detailed study of the effect of the various geometrical parameters on the return loss performance of such a package. The forward problem is solved using a fast, full-wave electromagnetic (EM) solver. The effect of various types of routing in the intermediate layers is examined closely. A sensitivity analysis based on parametric sweeps is carried out to identify the key variables. A hierarchical response surface based optimization is carried out to arrive at an optimum structure. A global optimizer based on simulated annealing is harnessed to find the optimum of non-linear and, in general, non-convex functions. The optimized structure exhibits excellent return loss characteristics translating into higher channel bandwidth.

1. Introduction

With the advent of multi-core microprocessor architectures, increasing demands are being placed on the input/output (I/O) bandwidths of central processing unit (CPU) package structures. An important component of the high-speed I/O CPU package is the vertical interconnect portion, comprising features such as plated through-holes (PTHs), micro-vias, pads, anti-pads, and lands. Appropriate design of the vertical interconnect is critical for channel impedance matching; degradation of matching results in increased reflections and contributes to reduced signaling speeds and data bandwidth. It is therefore imperative that schemes be developed to optimize the dimensions of the various features of the vertical portion of the package interconnect to achieve maximum possible bandwidth within the existing manufacturing realm before attempting any disruptive designs.

The complexity of the vertical interconnect structure makes the required full-wave electromagnetic analysis even of sub-sections computationally expensive. Traditionally, the emphasis has been on usage of quasi-static models for various portions of a package line and connecting them using a SPICE based analysis tool. Higher frequencies of operation together with non-negligible coupling arising from pitch reduction mean that full-wave analysis is now mandatory for package analysis [1].

However, many commercial full-wave field solvers based on Method of Moments (MoM) and Finite Element Method (FEM) suffer from high main memory requirements and increased time to solution. Accelerated Boundary Element method (BEM) based electromagnetic solvers have been receiving considerable interest, and one such solver [2] is used here as the forward analysis tool in the optimization process.

Even with the use of fast solvers, the large number of parametric variables creates a high dimensionality design space and makes the overall optimization procedure based on using the electromagnetic solver in the loop extremely expensive.

To circumvent these difficulties, the following schema has been developed, as shown in Fig. 1.

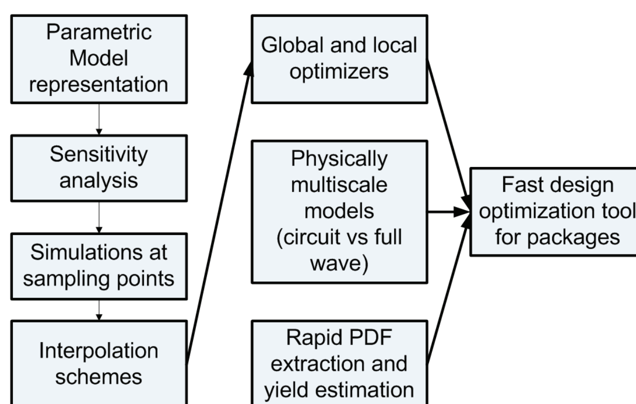


Figure 1: Flowchart for the methodology

Response surface methodology [3] is employed to derive an approximate, closed form relationship between the objective function and multiple geometrical variables of interest. A global optimizer in the form of simulated annealing is used to find the approximate location of this closed form relation. A high order response surface is used to reproduce the global variations while a smaller order response surface is used for refinement around the approximate global optimum. Thus, a multilevel scheme will enable an accurate location of the optimum dimensions of the package vertical interconnect feature sizes while also minimizing the calls to the full-wave EM solver.

Using this approach, differential pair designs exhibiting bandwidths of more than 12 GHz for –10 dB of differential return loss have been synthesized.

2. The parametric package model and the EM solver

To develop an automated framework for parametrics, optimization and statistical analysis of CPU packages, the first step is the development of a parametric package model. A ten-layer differential package line is used as the test structure. It is comprised of three build-up layers on the bottom followed by four core layers and three build-up layers on the top giving rise to a 3-4-3 package structure with a plated through hole

978-1-4244-4476-2/09/$25.00 ©2009 IEEE

Authorized licensed use limited to: University of Washington Libraries. Downloaded on July 13, 2009 at 17:59 from IEEE Xplore. Restrictions apply.

(PTH) separating the buildup and core layers. The structure primarily models the vertical interconnect portion with short traces running on the top-most and the bottom-most layers to facilitate excitation. A profile view of such a package structure is shown below in Fig. 2.

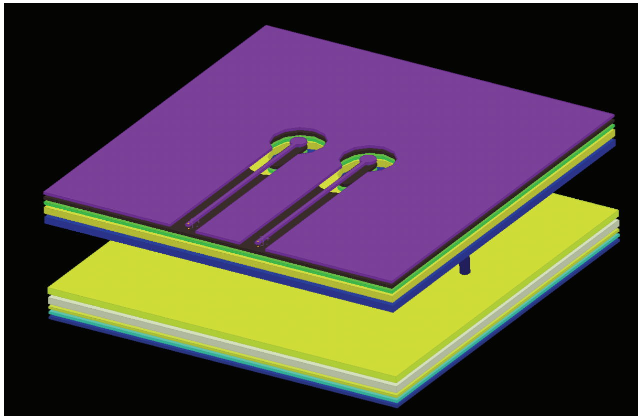


Figure 2: The profile view of the differential package line

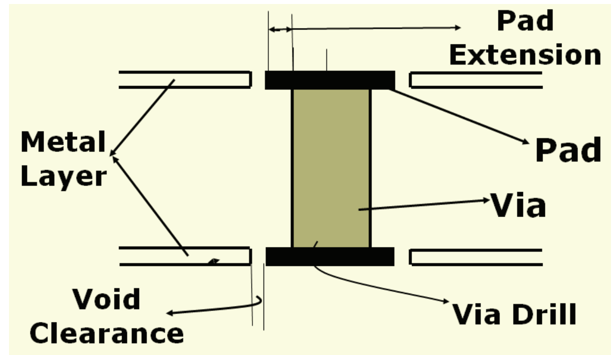
The important aspect of the structure is that the layout and the stack-up are completely parameterized. Every possible geometric parameter and electrical parameter has been parameterized with variables to facilitate parametrics, optimization and statistical study. These include the pad and anti-pad sizes, via diameters, metal and dielectric thicknesses, differential PTH pitch, ground plane sizes and many more.

Such a parametric model, represented in a format similar to Caltech Intermediate Format (CIF), is input to a mesher program which generates the appropriate mesh to be input to the field solver. A full-wave fast Maxwell solver based on QR factorization and compression is employed in this work [2]. The field solver has memory and time that scales nearly linearly with the number of mesh elements generated within the solver and uses fine-grained parallelism to exploit multiple computing cores for speed and scale. All critical physical effects such as radiation, skin and proximity effects, volume and surface current flow, phase and time delay effects, lossy metal and dielectric models are included in the field-based admittance parameter computation. Compared to state-of-the-art FEM and finite difference solvers, as well as dense matrix based MoM solvers, the field solver provides significant speedups in building the response surfaces, while preserving full Maxwell accuracy.

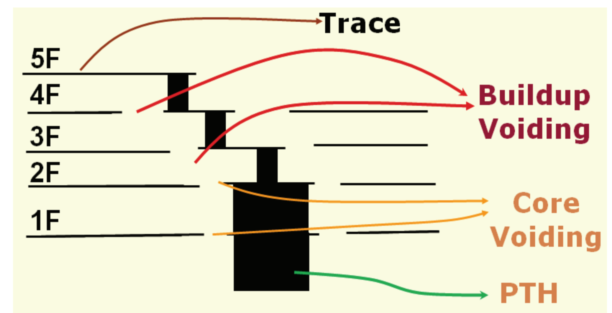
3. Parametric sweeps

A parametric sweep based sensitivity study is undertaken to identify the key parameters to be used for optimization. The nomenclature used in this paper is shown in Fig. 3a and 3b below. Figure 3a shows the generic definitions for the details around a via while Fig. 3b shows the voids around the vias in each layer. 'F' denotes front (or top) and 'B' denotes backside (or bottom) layers while referring to layer names.

Before presenting the results of parametric sweeps, it is interesting to observe the various types of routing of the traces that are possible in the intermediate (build-up) layers. This is shown in Fig. 4. The routing schemes are labeled R-R, R-L, L-R, and L-L based on the routing directions in layers 4B and 3B (and correspondingly symmetrically in layers 3F and 4F) being Right (R) and Left (L).



(a)



(b)

Figure 3: Nomenclature for (a) the package via structure and (b) the package layers and voids (in the top half of the package)

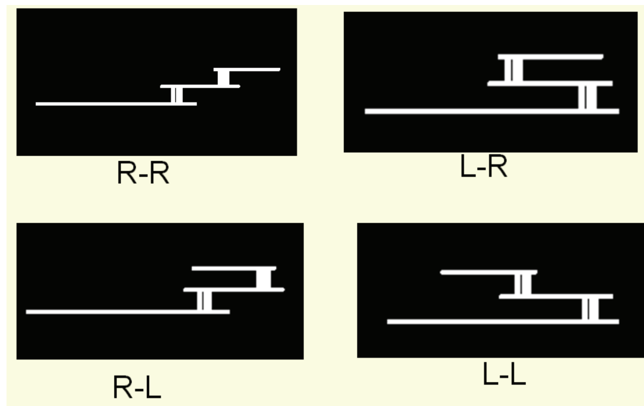


Figure 4: The various routing options in the build-up layers

A nominal design is chosen with an R-R routing (the conventional "waterfall" via stacking), all pad extensions equal to 50 µm, all void clearances equal to 100 µm, the PTH diameter equal to 300 µm, micro-via diameter of 120 µm,

Authorized licensed use limited to: University of Washington Libraries. Downloaded on July 13, 2009 at 17:59 from IEEE Xplore. Restrictions apply.

core thickness equal to 700 µm and a differential pitch equal to 600 µm. The x-y extent of the differential line is fixed at 4000 µm × 4000 µm. The dielectric has a relative dielectric constant ε_r of 4 and a loss tangent of 0.001. The objective function is the differential return loss (DS_{11}) in dB. Around this nominal design various parameters are swept to assess the sensitivities of the objective function with respect to each of the geometry parameters. The core-voiding extension is swept from 50 µm to 500 µm in steps of 50 µm at three different frequencies: 2, 7, and 12 GHz, respectively. The results are illustrated in Fig. 5. As shown the differential return loss improves as the voiding extension is increased. This is because of the better matching obtained due to reduction in parasitic capacitance between the pads and the metal left over after voiding. The gains are seen to saturate for larger voiding.

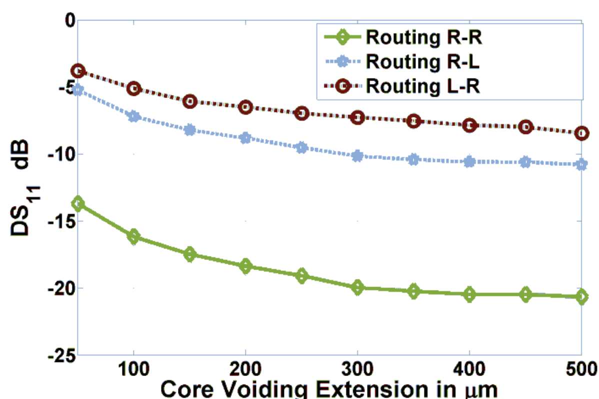


Figure 5: Parametric sweep: Core voiding extension for R-R routing

Another sweep with respect to the core voiding extension is then conducted for the three different routing schemes to assess the impact of the routing strategy on the return loss. The frequency chosen was 5 GHz. This produces interesting results as shown in Fig. 6.

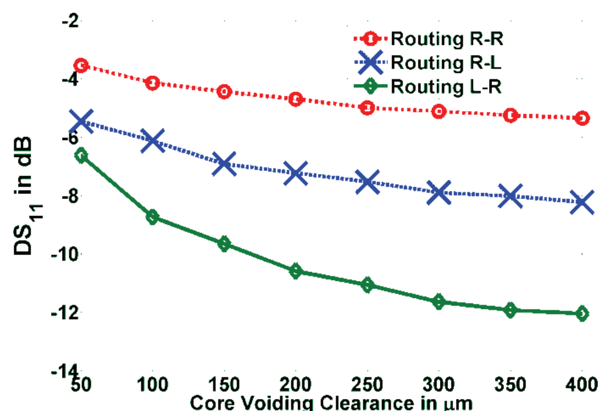


Figure 6: Comparing different routing strategies in terms of differential return loss

It is seen that the L-R routing scheme offers the best performance in terms of return loss. This is consistent with the findings in [3]. This can be explained by looking at Fig. 7 which shows a comparison between two types of routing: L-R and R-R. As shown, there are large parasitic capacitances between the via pads (especially the PTH pads) and the adjacent metal layers. This can result in large impedance mismatches. This effect is more pronounced in the R-R routing as opposed to the L-R routing due to the waterfall nature of the routing. This advantage holds true for both sets of layers (top and bottom) and over the two lines. A comparison with the R-L routing would also reveal that the L-R routing is better for similar reasons. Owing to these advantages, the L-R routing scheme is preferred over the R-R routing in all further optimization studies unless otherwise mentioned.

Results are next presented for the core-voiding extension sweep at 10 GHz at two different PTH diameter (drill) values. It is possible to use PTH drill sizes of 100 µm or 300 µm. It is interesting to note the effect of a smaller PTH on the return loss, shown in Fig. 8.

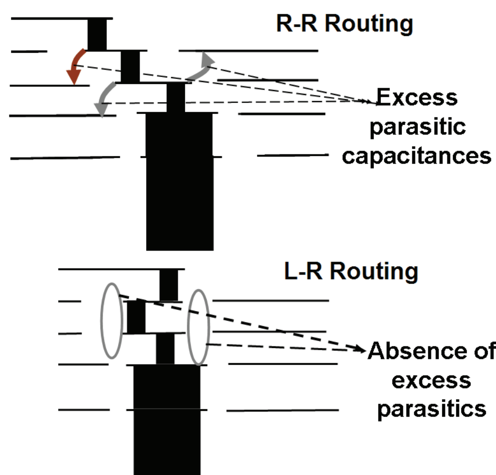


Figure 7: Two types of routing examined for parasitics

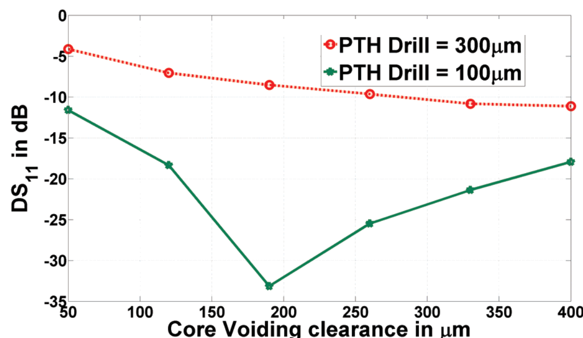


Figure 8: Effect of the PTH drill size on the differential return loss at 10 GHz

For a smaller PTH drill, the return loss exhibits an optimum at a voiding clearance of about 175 µm. Further, this optimum is seen to occur at smaller voiding extensions for L-R routing as opposed to R-R routing which exhibits a

Authorized licensed use limited to: University of Washington Libraries. Downloaded on July 13, 2009 at 17:59 from IEEE Xplore. Restrictions apply.

relatively monotonous response for all practical voiding extensions (a voiding extension of 400 μm is considered the limiting case, for a higher voiding would leave very little metal for power distribution). This effect can be explained as follows. For a larger PTH drill (two 300 μm PTHs separated at the edges by 300 μm), the interaction between the two PTHs forming the differential pair induces large mismatches in the impedance while a smaller PTH size (two 100 μm PTHs separated at the edges by 500 μm) would almost nullify this interaction. Then other effects start to dominate. In particular, referring back to Fig. 7, it can be seen that there are two competing factors as far as return loss is concerned. When the voiding below the trace is increased, the ground plane size starts diminishing and therefore this starts to destroy the transmission line guidance increasing the return loss. On the other hand this also facilitates the decrease in parasitic capacitances and decreases the return loss. In the case of R-R routing, the very nature of routing allows for the layer next to the voiding to act as continuity to the ground plane destroyed in the immediately upper layer resulting in signal guidance with a smaller mismatch in the characteristic impedance. However, in case of L-R routing, this is not possible as seen from the scenarios depicted in Fig. 7. Hence the competition between the two factors results in an optimum at a smaller voiding extension as compared to the R-R routing scheme. Next we look at the sweeps with respect to core voiding for two different values of buildup voiding at 5 GHz. These results are depicted in Fig. 9. It can be seen that while one is monotonous, the other (at 201 μm) exhibits an optimum.

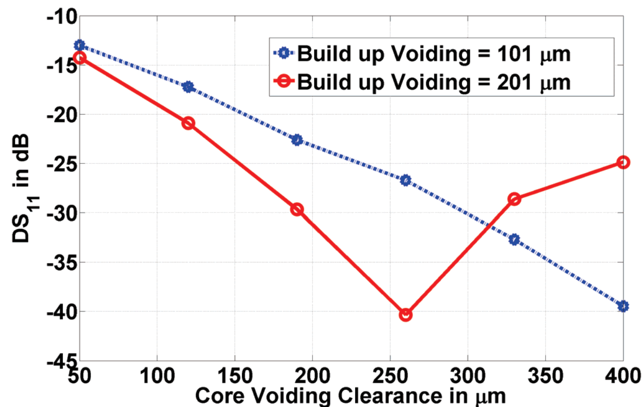


Figure 9: Core voiding sweeps for two different build-up voiding sizes at 5 GHz

All these results suggest the existence of complex interaction terms between the various geometrical parameters. Obviously, it is impossible to derive closed-form analytical expressions relating an objective function such as differential return loss with the various geometrical parameters. Also it is impossible to utilize brute-force search techniques with a full-wave electromagnetic solver given the prohibitive amount of simulation time this is going to consume. Hence techniques which offer approximate closed form relations utilizing a smaller number of simulations are looked at next. Before considering these techniques, results from more parametric sweeps are presented in Figs. 10–13.

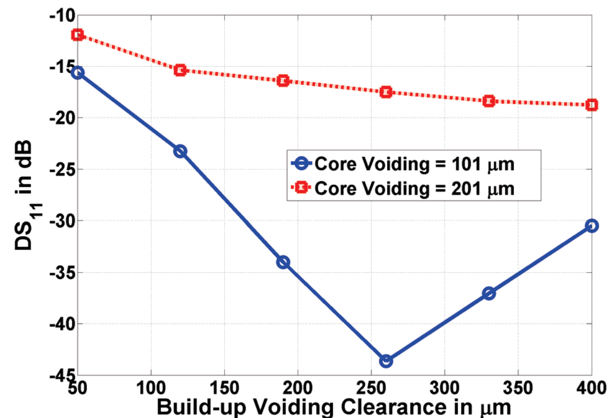


Figure 10: Build-up voiding sweeps for two different core voiding sizes at 5 GHz

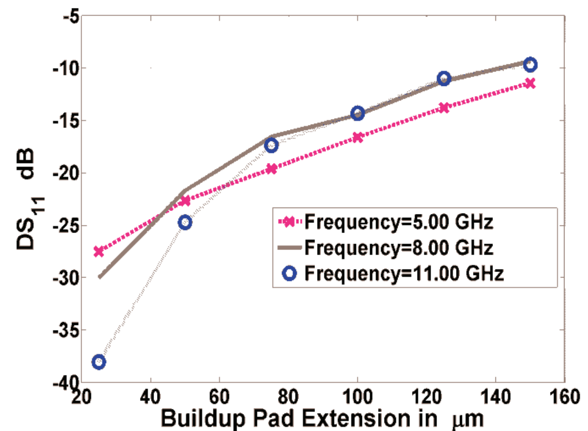


Figure 11: Differential return loss sweeps with respect to the build-up pad extension

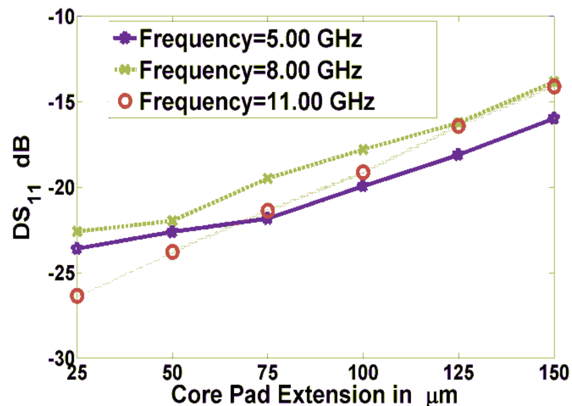


Figure 12: Differential return loss sweeps with respect to the core pad extension

Authorized licensed use limited to: University of Washington Libraries. Downloaded on July 13, 2009 at 17:59 from IEEE Xplore. Restrictions apply.

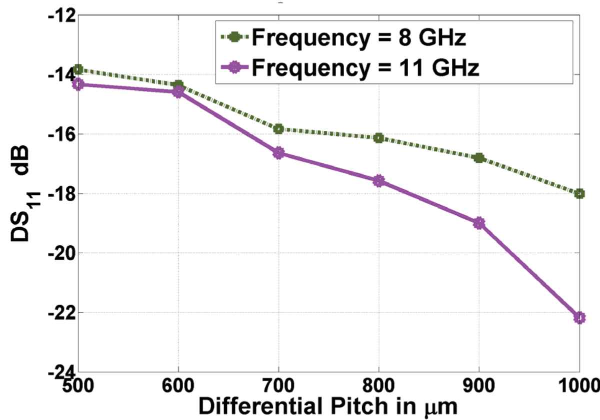


Figure 13: Differential return loss sweeps with respect to the differential pitch

The results confirm intuitive observations that the differential return loss worsens as the pad sizes are increased and improves with the increase of differential pitch owing to reduced interaction between the two PTH vias.

4. Response Surface Methodology (RSM)

This section provides a brief overview of RSM, which is a multivariate interpolation scheme based on minimizing the mean-squared error between the predicted and the observed values of the objective function. Let y be an objective function that depends on a vector of input variables $[x_1, x_2, \ldots, x_n]$. If the objective is to fit a 2nd order response surface for y in terms of the x_i^s, then y can be written as

$$y = a_0 + a^T x + x^T \mathbf{B} x \quad (1)$$

where a_0 is a constant, a is a vector of coefficients and $\mathbf{B}$ is a symmetric matrix of coefficients. As an illustration, a second order response surface for y in terms of two variables x_1 and x_2 would take the form

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1^2 + a_4 x_2^2 + a_5 x_1 x_2 \quad (2)$$

Experiments are designed by assigning each variable a pre-determined sample value depending on the sampling scheme used. An mth order response surface in n variables would thus need $(m+1)^n$ simulations for a full-factorial design.

Objective function values are obtained by simulations for each of the vector of variables and equations written out. The equations are linear in the unknown co-efficient and thus the over-determined system can be solved in the least-squares sense to obtain an estimate of the coefficients. A linear response surface is given by

$$y = a_0 + \sum_{i=1}^{n} a_i x_i \quad (3)$$

The solution for the coefficients in a linear response surface is given formally by

$$a = \left(\left(\mathbf{X}^{\mathrm{T}} \mathbf{X} \right)^{-1} \mathbf{X}^{\mathrm{T}} \right) v \quad (4)$$

where a is the vector of coefficients, which is the unknown in this case, $\mathbf{X}$ is the matrix of size $m \times (n+1)$, m is the number of simulations and n is the number of input variables. Also the first column has all its entries equal to unity and the remaining entries are formed from the values of the input variables x_i corresponding to the particular objective function evaluation. The vector v is comprised of the observed values of the objective function. The solution process, however, is performed using techniques such as singular value decomposition (SVD). The higher order response surfaces are constructed similarly by introducing new variables for the higher order and cross terms. Figures 14 and 15 show some results for the response surfaces for the differential return loss with respect to one and two variables respectively. Simple visualization can provide the optimum for dimensions of one and two. For dimensions greater than two an optimizer routine is needed to find the optimum.

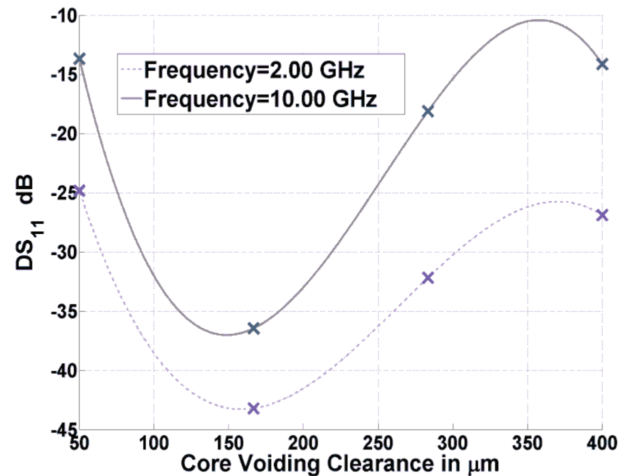


Figure 14: Differential return loss, third order response surface in core voiding

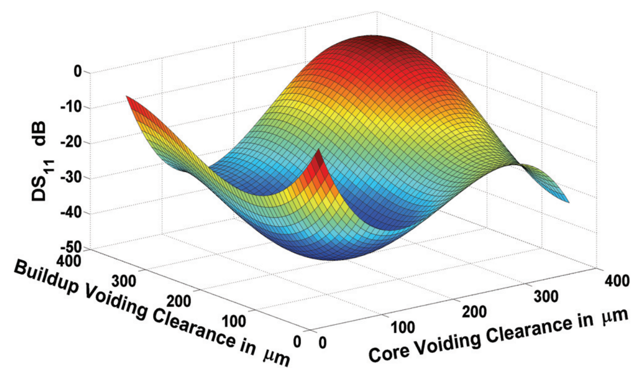


Figure 15: Third order response surface in two variables: Buildup and core voiding extensions – at 10 GHz

5. Optimization

Once the response surface is constructed, the goal is to find the combination of input parameters that results in the optimum of the objective function (which itself might be the weighted sum of a number of performance measures). In this case, the optimizer operates on the approximate closed-form

Authorized licensed use limited to: University of Washington Libraries. Downloaded on July 13, 2009 at 17:59 from IEEE Xplore. Restrictions apply.

expression derived with the use of response surface. Traditional gradient based optimizers find only a local optimum while the goal here is to look for the global optimum. For convex functions, it can be shown that a local optimum is the global optimum. However, in this case, the expression cannot be guaranteed to be convex. Hence, Simulated Annealing, a physics based global optimizer, is utilized to find the global minimum of the approximate surface [5].

This has been very successfully used to solve the problems of analog synthesis [6] and VLSI floor-planning [7]. It is a global optimizer since it allows escape out of local minima by accepting uphill moves in the optimization process. The algorithm has a slow convergence rate. However, in this work, it is used only on the closed form expression derived from the RSM which results in a few seconds of run time. The algorithm is briefly outlined in Fig. 16. It makes use of a numerical quantity called the temperature (T) which controls the acceptance of uphill moves. Further details can be found in any of the references in [5]–[7].

6. The hierarchical methodology

With the simulation results at a small number of sampling points, it is impossible to reconstruct the entire landscape of the objective function very accurately, as the number of dimensions increase. This is shown in Table I, which enumerates the error in the response surface by simulating 40 random points in the parameter space and evaluating from the response surface. It is clear that a third order response surface is necessary to ensure a better accuracy.

It is to be noted that an exact replication of the landscape is not necessary. A more important requirement is the reproduction of the location of the optimum rather than the exact shape of the terrain. This is illustrated in Fig. 17 which compares the results from a fine parametric sweep and a third order response surface. It is seen that the location of the optimum is predicted accurately by the response surface while the actual value is in considerable error (about 9%).

Also, in this case the true optimum happened to be close to 200 µm which is also a sampling point. This may not happen always especially with higher dimensions where the exploration volume will be higher. To circumvent this, a hierarchical strategy has been proposed. In this method, first, a higher order response surface (3rd order in this work) is constructed over the global range of parameters. The optimum of this is found and a lower order response surface is constructed around this by shrinking the interval size for all the variables. This allows for a more accurate reproduction of the location of the optimum.

7. Optimization results

Four variables were used in the optimization process. They included build-up and core pad extensions and the voiding extensions. All the four core-voiding extensions , all the four core pad extensions, all the six build-up voiding extensions and all the six build-up pad extensions were set to one variable each.

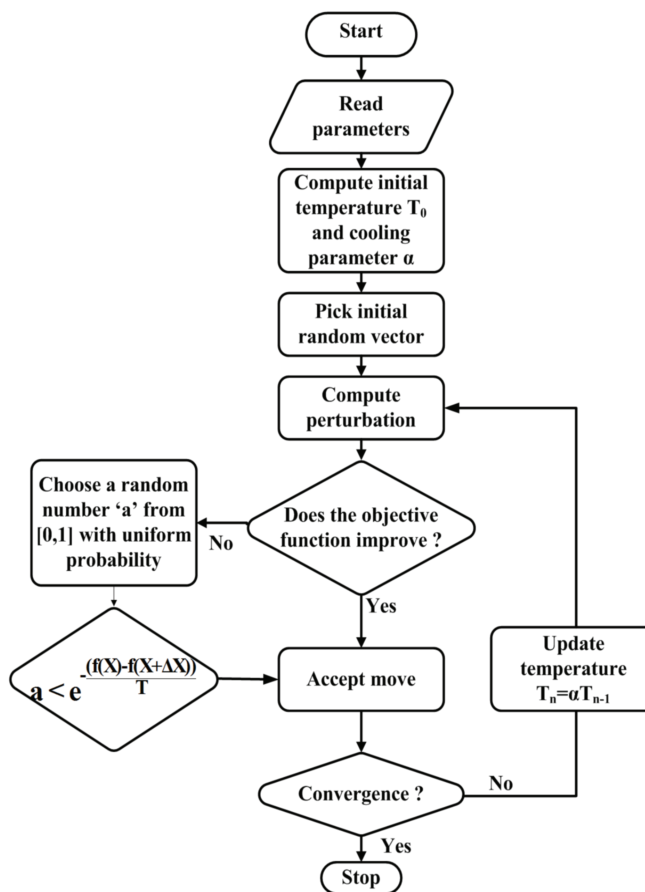


Figure 16: Simulated Annealing optimization flow

Table I: Response surface error performance

Number of variables	Variables	Average error for 2nd order RS	Average error for 3rd order RS
1	Build-up Pad	4.3%	2.8%
1	Core Pad	1.9%	1%
1	Build-up Voiding	14.5%	2.9%
1	Core Voiding	22.3%	9.5%
2	Build-up and Core Pad	6.3%	3.1%
2	Build-up and Core Void	18%	11.6%
4	All pads and voids	–	13.6%

Authorized licensed use limited to: University of Washington Libraries. Downloaded on July 13, 2009 at 17:59 from IEEE Xplore. Restrictions apply.

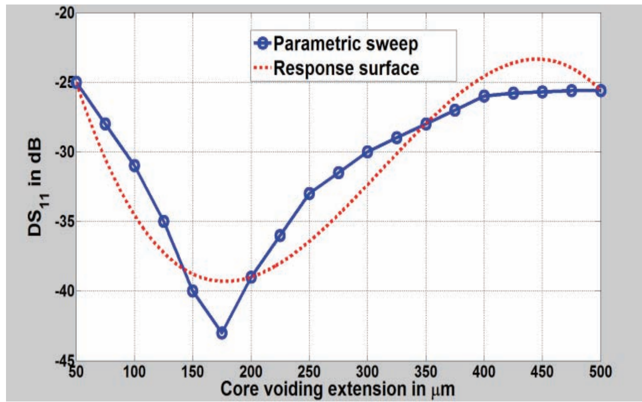


Figure 17: Comparing parametric sweep and RS

The ranges fixed for the voiding extensions were 100 to 300 µm and the pad extensions were allowed to move in the range of 20 to 80 µm. The frequency response of the differential return loss for these package structures has been observed to show an ”LC” type resonance depending on the various physical dimensions which affect the PTH inductance and the parasitic capacitance. Thus, if the optimization is carried out at a single frequency, such as 10 GHz, the optimizer routine shifts the resonance to 10 GHz while allowing a less than satisfactory performance at, say, 5–7 GHz. Hence the optimization was carried with a weighted objective function at three different frequencies (3, 7 and 11 GHz) to ensure a good overall response. The simulations were carried out on a 2.4 GHz dual processor, quad-core machine and it took a little under 27 hours to complete. The optimum was found at a core voiding of 285 µm, a build-up voiding of 160µm, a core pad extension of 50 µm and a build-up pad extension of 70 µm (all rounded to the nearest 5 µm). Fig. 18 shows the improvement in the differential return loss over a starting nominal structure with R-R routing.

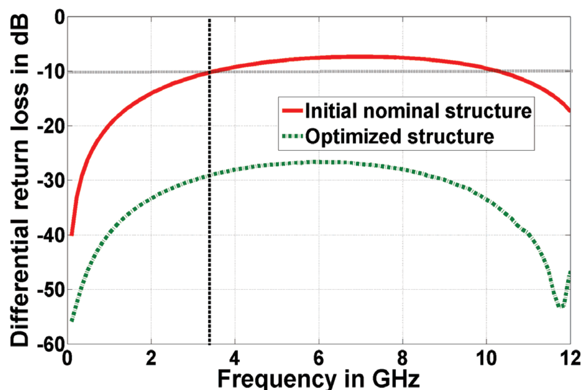


Figure 18: Comparison of a nominal and an optimal structure with respect to the differential return loss

The initial structure had all pad sizes equal to 50 µm and all void sizes equal to 100 µm. Clearly, the differential bandwidth (measured at –10 dB of differential return loss) exceeds 12 GHz for the optimal structure while it is a little less than 4 GHz for the initial nominal structure.

8. Conclusions

This paper presents a detailed parametric study of a realistic ten-layer package structure. The effect of routing strategy was examined in detail. A bare-bone optimizer based on hierarchical response surface methodology and simulated annealing has been used to optimize the design with respect to four important design parameters. It is shown that the optimizer is capable of producing design guidelines with respect to improved differential return loss. Future work includes incorporation of the solder ball in the package design, including differential cross-talk in the optimization and trade-off and techniques to reduce the number of simulations in addition to expanding on the number of variables in the design space to explore. The same frame-work will be utilized to account for the spread in the performance measures occurring due to the variations in physical dimensions and electrical properties, arising out of manufacturing variations.

9. References

1. J. Mao, G. Fitzgerald, A. Kuo, and S. Wane, "Coupled Analysis of Quasi-static and Full-Wave Solution towards IC, Package and Board Co-design," in *Proc. IEEE Electrical Performance of Electronic Packaging (EPEP)*, Atlanta, GA, Oct. 29–31, 2007, pp. 111–114.
2. D. Gope and V. Jandhyala, “Efficient solution of EFIE via low-rank compression of multilevel predetermined interactions,” *IEEE Trans. Antennas Propagat.*, vol. 53, no. 10, pp. 3324–3333, Oct. 2005.
3. W. T. Beyene, S. Hao, J. Feng, and C. Yuan, “Design and analysis methodologies of a 6.4 Gb/s memory interconnect system using conventional packaging and board technologies,” in *Proc. Electronic Components and Technol. Conf. (ECTC)*, Las Vegas, NV, June 1–4, 2004, pp. 1406–1411.
4. R. H. Myers and D. C. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 2nd ed., Wiley-Interscience, New York, Feb. 2005.
5. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 598, pp. 671–680, May 1983.
6. C. Sechen and A. Sangiovanni-Vincentelli, "The TimberWolf placement and routing package," *IEEE J. Solid-State Circuits*, vol. 20, no. 2, pp. 510–522, Apr. 1985.
7. E. S. Ochotta, R. A. Rutenbar, and L. R. Carley, "Synthesis of high-performance analog circuits in ASTRX/OBLX," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, no. 3, pp. 273–294, Mar. 1996.

Authorized licensed use limited to: University of Washington Libraries. Downloaded on July 13, 2009 at 17:59 from IEEE Xplore. Restrictions apply.