Localized Statistical 3D Thermal Analysis Considering Electro-Thermal Coupling

Zuying Luo[†], Jeffrey Fan[§], Sheldon X.-D. Tan[‡]
[†]Department of Electronics, Beijing Normal University, Beijing, China
[§]Department of Electrical and Computer Engineering, Florida International University, Miami, Florida, USA
[‡]Department of Electrical Engineering, University of California, Riverside, USA

***Abstract*— In this paper, we propose a novel method for analyzing fewer hot spots in a chip. The method, called SNSOR (Single-Node Successive Over Relaxation), is based on a novel localized relaxed iterative approach to perform statistical analysis on one hot spot at a time. Based on SNSOR, we propose an approximation method, called ET-SNSOR (Electro-Thermal SNSOR), to deal with the electro-thermal coupling (ETC) effects. ET-SNSOR first uses the iterative method to update correlations from ETC, and then computes standard temperature deviations for hot spots, according to ETC and updated correlations. Experiments show that ET-SNSOR is three orders of magnitude faster than the Monte-Carlo method with small errors (less than 4.76% on maximum). It only takes an average of 0.18 second to analyze one hot spot statistically for a large test case of 1.3M nodes with ETC effects.**

I. Introduction

Process-induced variability becomes a major design concern in the current 65 nm and upcoming 45 nm and 32 nm VLSI technologies. Given the high sensitivity of today's chip performance to even minor temperature fluctuations [1], it is important to characterize the impact of process variations on 3D thermal analysis. However, prior works on 3D thermal analysis have not considered the influence of process variations on temperature distribution [1], [2], [3], [4].

The studies on 3D thermal analysis can be classified into two groups. The first group is related to steady-state work [5], in which inputs are steady-state energy distributions. Because energy distributions do not vary in time domain, steady-state methods only consider thermal conductance and ignore thermal capacitance, which leads to lower time complexity. Later research focuses on dynamic 3D thermal analysis, in which energy distributions vary in time domain. These dynamic methods [3], [4], [6], [7] consider electro-thermal coupling (ETC) effects, in which temperature also has influence on leakage power. The methods include ADI based method [6], green-function based method [7], multi-grid based method [3], [4]. They are much more efficient in dealing with the large problem of transient 3D thermal analysis.

In this work, we first propose a novel single-node SOR method (SNSOR) to efficiently analyze standard temperature deviations (STD) for hot spots in 3D thermal analysis, based on the global SOR method [9]. In SNSOR, only one stimulus is applied at a hot spot, say q, to solve the response vector R_q, where SNSOR can directly compute STDs. SNSOR performs relaxation on nodes from q to its surrounding nodes in the wave-propagating style and the wave stops at some nearby nodes, whose temperature increases are less than a predefined value. SNSOR dynamically reduces the nodes used in the relaxation process, based on their value changes during the process. In addition to SNSOR, we further propose an approximation method, called ET-SNSOR, to analyze STDs of hot spots with ETC consideration. In order to consider the correlation between energy variations and ETC, ET-SNSOR proposes two-stage operations to solve this problem. First, it uses iterative method to update the correlation coefficients among blocks, with which it computes the updated STDs. Second, it uses an analytical method to compensate ETC effects for the temperature variation.

As this work performs an statistical 3D analysis under the power density disturbance and ETC, a large number of experiments have drawn into the following conclusions. (1) Compared with the simulation method, the global ICCG and SOR method [8], [9], SNSOR shows its superiority on efficiency and accuracy. For a large test case of 1.3M nodes, it takes an average of 0.1481 seconds to statistically analyze a hot spot, which means that SNSOR can be used in statistical analysis of hot spots in practical VLSI designs of huge scale without ETC consideration. (2) ETC effect clearly influences correlations of energy distances among blocks. With the increase of ETC effect, iterations managing ETC also increase in both Monte Carlo (MC) method and ET-SNSOR, while errors of ET-SNSOR increase as well. (3) Compared with the simulation method of 5,000 samplings, ET-SNSOR is three orders of magnitudes (2,486X) faster and causes slight accuracy loss (4.76% on maximum) in computing STDs for 450 hot spots. For a large test case of 1.3M nodes under ETC, it takes an average of 0.18 second to statistically analyze a hot spot, which means that ET-SNSOR can be used in statistical analysis of hot spots in practical VLSI design of huge scale with ETC.

II. 3D Thermal Analysis

Because rampant process variations of transistor length and threshold voltage cause exponential leakage power variation [2], [4], it is necessary to consider energy variation ΔP in 3D thermal analysis. Since conductance variations ΔG can be transformed into the power disturbance, this work only considers ΔP in statistical 3D thermal analysis. And according to the well-known electro-thermal analysis equation $G \bullet T = P$, the following equation is used to solve ΔT.

$$G \bullet \Delta T = \Delta P \tag{1}$$

where ΔT is the temperature respondence for ΔP. Since process variations are rampant in nanometer technology, ΔP is so large that we assign $3\sigma = 60\%$. Similar to reference [10], we define ΔP as normal distribution variations $N(p\mu, p\sigma)$. To consider the spatial correlations, we divide nodes into M blocks in x-y planar and use a simple correlation model: all variations of loaded energies in a block are absolutely correlated while spatial correlations exist between different blocks. But we stress that our method can be used for more complicated correlation models.

Because of the temperature influence on leakage power [2], [4], it is necessary to consider ETC effect in 3D thermal analysis. In this work, we only consider linear-modelling ETC as follows.

$$\Delta P \div P = \alpha \bullet \Delta T \tag{2}$$

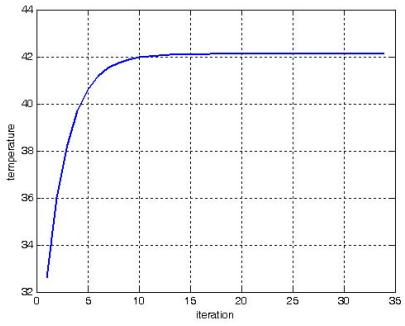

Fig. 1. Temperature evolution with ETC iterations.

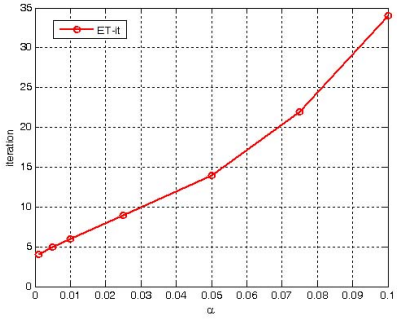

Fig. 2. Iteration number scales with ETC.

where α is the ETC coefficient. As α increases, ETC becomes stronger. In this work, $\alpha \leq 0.005$ means weak ETC, $0.005 < \alpha \leq 0.025$ means medium ETC, and $\alpha > 0.025$ means strong ETC. Thus, equation(1) must be transformed into the following formula after considering ETC.

$$G \bullet \Delta T = \Delta P + (\alpha \bullet P) \bullet \Delta T \tag{3}$$

As a result, we have to use iterations to manage ETC. Shown in Fig. 1, ETC-free temperature is about 32.6 but strong ETC of $\alpha = 0.100$ increases the nodal temperature to about 42.1 after 34 iterations. ETC can increase 29.24% nodal temperature, which means that it is necessary to consider ETC in 3D thermal analysis. Also, shown in Fig.2, as α increases from 0.001 to 0.100, ETC has more influences on 3D thermal analysis and the iteration number increases from 4 to 34.

III. Single-node without ETC

Assume that R_q is the thermal resistor vector of hot spot q. Because the only unit stimulus ($1W$) is put at q, the response vector R_q may be defined as the nodal temperature vector T_q excited by the unit stimulus:

$$T_q = G^{-1} \bullet P_q = R \bullet P_q = R_q \tag{4}$$

where $R = G^{-1}$ is the resistance matrix, P_q is the input stimulus vector in which all elements are zero except for node q that has value 1 in P_q. Although global ICCG [8] and SOR [9] methods and localized RW method [10] can solve R_q directly, we will propose a much more efficient method (SNSOR) in following section.

In this work, we first use the global SOR method to compute the nominal temperature vector T_0 according to P_0 with the equation $G \bullet T_0 = P_0$. Based on the obtained T_0, we can find fewer hot spots and then, use R_q to directly compute the standard temperature deviation $t\sigma_q$ for node q. Since ΔP is spatially correlated, spatial correlations must be considered in solving $t\sigma_q$.

$$(t\sigma_q)^2 = \sum_{bi=1}^{M} \left\{ t\sigma_{q,bi} \sum_{bj=1}^{M} (C_{bi,bj} \bullet t\sigma_{q,bj}) \right\} \tag{5}$$

where M is the number of blocks, $t\sigma_{q,bi}$ is the contribution of all nodes in block bi. And $C_{bi,bj}$ is the spatial correlation coefficient between blocks bi and bj. Because all energy variations are absolutely correlated among nodes in a block, we can use the following equation to compute $t\sigma_{q,bi}$.

$$t\sigma_{q,bi} = \sum_{p \in N_{bi}} (r_{p,q} \bullet p\sigma_p) \tag{6}$$

where N_{bi} is the node set of block bi and node p belongs to N_{bi}. $p\sigma_p$ is the standard energy deviation of node p and $r_{p,q}$ is the resistance between nodes p, q. And $r_{p,q}$ is a p^{th} element in R_q.

IV. SNSOR: single-node SOR method

After obtaining the nominal temperature vector T_0, we will mark all nodes with large temperature increase as hot spots. The SNSOR algorithm is listed below to obtain the resistance vectors with which EQ(5) - EQ(6) are used to directly compute the standard temperature deviations.

1) Initialize the test case. Temperature of each node is pre-defined as $T_{initial} = 20°C$. Node q of the only 1W stimulus is stored into relaxation set A as an edge node.
2) If all nodes in A have been relaxed, turn to step 7. Else, sequentially take a node p from A and go to the next step.
3) If node p is a relaxation-ending node, return to step 2. Else, go to the next step.
4) Relax p and then do next according to following three cases. Case one, if p is an edge node and $\delta t_p^{(k+1)} = \mid t_p^{(k)} - t_p^{(k+1)} \mid$ is less than an pre-defined value $\varepsilon_2 \ll 1.0$, return to step 2. Case two, if p is an edge node and $\delta t_p^{(k+1)} \geq \varepsilon_2$, mark p as an internal node and go to step 5. Case three, if p is an internal node, go to step 6.
5) For the new-marked internal node, mark all un-relaxed neighbor nodes as edge nodes and push these new edge ones into A. Then, return to step 2.
6) if $\delta t_p^{(k+1)}$ is less than an pre-defined value $\varepsilon_3 \ll 1.0$(in this work, $\varepsilon_3 = 0.1\varepsilon_2$), mark p as a relaxation-ending node and return to step 2.
7) If errors of all nodes are less than a pre-defined value ε_1, end the algorithm and output the solution. Else, relax A again and return to step 2.

To demonstrate the effectiveness of SNSOR, we use SNSOR to calculate response vectors and observe the convergence and the number of the relaxation nodes of the method. Fig. 3 shows the convergence rate for a test-case with 544.5k nodes. As shown in the left-hand side of Fig. 3, the number of SNSOR iterations is 29 and the nodal temperature quickly converges to the final value in the 18th iteration.

In Step4 of the above SNSOR algorithm, as node p is marked as an edge node, p can't be further marked as an internal node in Step4 if the resistance $r_{p,q}$ is small enough. Thus, the relaxation wave will stop at the edge node, which limits relaxation only to some nodes around q. Shown in the right hand side of Fig. 3, SOR nodes from this wave-ending technique are named as the candidate SOR nodes. Because the number of the candidate SOR nodes speedily converges to a

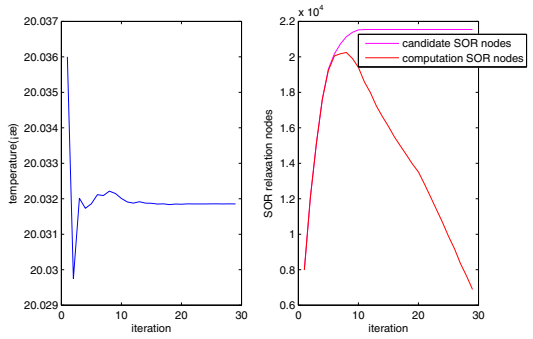

Fig. 3. Convergence rate of SNSOR.

number (21,545) far less than all nodes (544,500) as SNSOR reaches 12^{th} iteration, the wave-ending technique in Step4 can limit SNSOR into a small region (3.97%).

If an internal node is marked as a relaxation-ending node in Step6, the node will not be relaxed in later SOR iterations according to the operation in Step3. More internal nodes will be marked as relaxation-ending ones in later SOR iterations. Thus, SNSOR is much more efficient than the global SOR method. Shown in the right hand side of Fig. 3, the SOR nodes, which are not relaxation-ending nodes, are named as the computation SOR nodes. With SNSOR iteration increasing, the number of the computation SOR nodes reach to the maximum: 20,245, at the 8^{th} iteration and then it goes down to 6893 in the end. Compared with RW method [10] that always walks through paths from the source node to terminal nodes, SNSOR only walks partial distances after excluding relaxation-ending nodes in later SOR iterations, which means much more efficiency.

V. ET-SNSOR METHOD FOR ETC

Based on SNSOR, we further propose an approximation method, called ET-SNSOR, to analyze temperature variations of hot spots under energy variance and ETC. Since ETC has influence on thermal analysis, we have to use the iterative method to compute nominal temperature T_0 according to P_0. Then, based on T_0, we can find fewer hot spots and use ET-SNSOR to computer standard temperature deviations for them. In order to break the coupling relation between energy variations and ETC, ET-SNSOR includes the following two-stage operations.

A. Correlation updating

Besides initial correlations, ETC creates additional correlations. In this work, we consider initial correlation, $C_{bi,bj}$, as $C^0_{bi,bj}$ and then use iterative method to update the correlation $C^{k+1}_{bi,bj}$ from $C^k_{bi,bj}$ as follows:

$$C^{k+1}_{bi,bj} = \frac{\alpha \bullet p\mu_q \bullet t\sigma_{q,bi} + C^k_{bi,bj} \bullet p\sigma_q}{\alpha \bullet p\mu_q \bullet t\sigma^k_q + p\sigma_q} \tag{7}$$

where $t\sigma^k_q$ is the standard temperature deviation of node q calculated from the k^{th} updated correlations $C^k_{bi,bj}$ while $t\sigma^0_q$ is the standard temperature deviation obtained from $C^0_{bi,bj}$ without considering ETC. With $C^{k+1}_{bi,bj}$, we can compute the $t\sigma^{k+1}_q$ as the following:

$$(t\sigma^{k+1}_q)^2 = \sum_{bi=1}^{M} \left\{ t\sigma_{q,bi} \sum_{bj=1}^{M} (C_{bi,bj^{k+1}} \bullet t\sigma_{q,bj}) \right\} \tag{8}$$

The iteration will end as $\delta t\sigma^{k+1}_q = \mid t\sigma^{k+1}_q - t\sigma^k_q \mid$ is less than a pre-defined value $\varepsilon_4 \ll 1.0$. Also, $t\sigma^{k+1}_q$ is considered as the internal value $t\sigma^C_q$ for the last correlations.

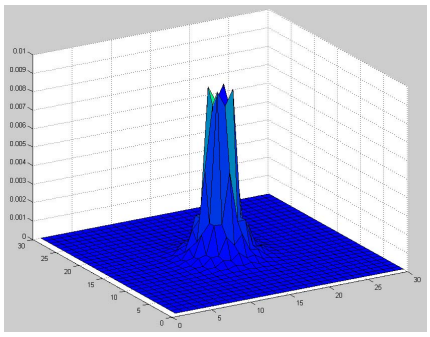

Fig. 4. Correlations updated for ETC.

To show ETC's influence on correlations, we design an example. At the beginning, there is no correlation. Then, EQ (7-8) are used to iteratively update the correlations among blocks. At the end, all correlations between the block to which node q belongs and the other blocks are shown in Fig. 4. It is the big surprise that ETC of $\alpha = 0.025$ causes remarkably visible correlations on the correlation-free base. In fact, ETC has increasing influences on correlation updating.

B. ETC compensation

In the description above, we only consider ETC's influence on the correlation and obtain $t\sigma^C_q$. In fact, ETC also has influence on $t\sigma_q$ itself. Since it is a positive feedback, we can use the following analytic equation to compensate the influence and obtain the final result $t\sigma_q$.

$$t\sigma_q = \frac{p\sigma_q}{p\sigma_q - \alpha \bullet p\mu_q \bullet t\sigma^C_q} \bullet t\sigma^C_q \tag{9}$$

In order to validate the accuracy of ET-SNSOR, we compare the standard temperature deviation originated from the MC method of 5,000 samples as the normalized comparison benchmark, *i.e.* 100%. Then, we compare the ETC-free value $t\sigma^0_q$ and ETC-compensated value $t\sigma_q$ from equation(9). Comparison results are shown in Fig.5. As a result, accuracy of ETC-free and ET-SNSOR decreases with increasing ETC. As ETC scales from $\alpha = 0.001$ to $\alpha = 0.050$, the ratios of $t\sigma^0_q$ range from 99.64% to 80.19%. Meanwhile, the ratios of ET-SNSOR will range from 99.95% to 95.31%, which means that ET-SNSOR is accurate enough. In the following experiments, we will assign $\alpha = 0.025$ for all test cases.

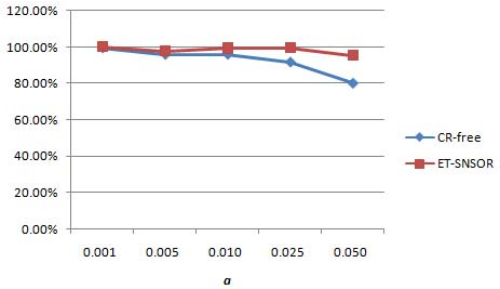

Fig. 5. ET-SNSOR accuracy validation.

TABLE I
METHOD COMPARISON FOR 450 HOT SPOTS W/O ETC

nodes	MC	Global SOR			Global ICCG			SN-SOR		
	τ_{sim}	τ_{std}(S)	speedup	Max. Err.	τ_{std}(S)	speedup	Max. Err.	τ_{std}(S)	speedup	Max. Err.
40500	1872.95	61.88	30.27	0.737%	153.45	12.21	0.725%	23.64	79.22	2.240%
112500	4790.20	171.61	27.91	2.895%	599.89	7.99	2.909%	31.16	153.75	3.153%
220500	9226.56	333.31	27.68	1.857%	1595.25	5.78	1.855%	36.30	254.20	1.984%
364500	14884.50	552.05	26.96	1.387%	2361.64	6.30	1.394%	40.88	364.15	1.810%
544500	27005.34	983.33	27.46	2.066%	3503.19	7.71	2.083%	45.70	590.89	2.185%
760500	NA	1179.33	Base	base	4948.89	0.24	0.021%	53.53	22.03	0.308%
1012500	NA	1551.52	Base	base	6699.00	0.23	0.011%	59.44	26.10	0.149%
1300500	NA	1996.95	Base	base	8499.94	0.23	0.023%	66.66	29.96	0.126%

VI. EXPERIMENTAL RESULTS

The hardware platform is D930(2XCPU-3.0GHz) with 1GB memory, while the software is written in standard C. In 3D thermal analysis, each die is divided into five levels in z direction. Each test case is divided into 900 blocks in X-Y planar. The 450 blocks are weak ones, in which each node is linked with $N(p\mu, p\sigma)$ energy disturbance, while 450 blocks are strong ones in which each node is linked with N((2-10)$p\mu$, (2-10)$p\sigma$) energy disturbance. Meanwhile, weak and strong blocks are arranged into a chessboard array in which each weak block has four strong blocks as neighbors. Thus, centers of 450 strong blocks are hot spots that locate in the same z-direction level far away from the heat-sink. This work mainly focuses on temperature variations of these spots. In order to show the effectiveness of SNSOR and ET-SNSOR, we compare them with the MC simulation method, the global SOR solving method, and the global ICCG solving method. The simulation method of 5,000 samples takes τ_{sim} runtime to obtain standard deviations of all nodal temperatures for accuracy comparison. In this work, it takes τ_{mean} runtime for the global SOR method to solve for the nominal temperature vector T_0. Then based on T_0, for all solving methods, including Global ICCG, Global SOR, SNSOR, and ET-SNSOR, EQ(5)-EQ(9) are used to compute standard temperature deviations (STD) for 450 hot spots. The runtime of 450 STDs solving is named as τ_{std}.

TABLE II
SUPERIORITY OF ET-SNSOR METHOD.

nodes	MC	ET-SNSOR			
	$T_{std}(S)$	$T_{std}(S)$	Speedup	Avr. Err.	Max. Err.
40500	15034	47.02	319.8	1.69%	4.50%
112500	40729	48.09	846.9	0.96%	4.28%
220500	77993	50.91	1532.1	0.87%	3.51%
364500	127609	51.33	2486.1	1.13%	4.76%
544500	NA	57.14	NA	NA	NA
760500	NA	61.41	NA	NA	NA
1012500	NA	70.77	NA	NA	NA
1300500	NA	79.28	NA	NA	NA

Efficiency and accuracy comparisons are shown in Table I for ETC-free test cases. Compared with the simulation method, SNSOR is three orders of magnitude (1,101X) faster and causes slight accuracy loss (2.218% on maximum) for STDs of 450 hot spots. Compared with the global SOR method [9], SNSOR is 32X times faster and causes slight accuracy loss (0.541% on maximum). SNSOR takes only 67 seconds to solve 450 hot spots in the largest case of 1.3M nodes, which means it takes an average of about 0.15 seconds to solve a spot.

Compared with the MC method, we validate the efficiency and accuracy of ET-SNSOR with test cases under ETC. Shown in Table.2, ET-SNSOR can speedup three orders of magnitudes (2,486X) with the slightly accuracy loss of 4.76% on maximum. It takes only 79 seconds to solve 450 hot spots in the largest case of 1.3M nodes under ETC, which means it takes an average of about 0.18 seconds to solve a spot. Compared with SNSOR, ET-SNSOR only takes slight additional runtime (less than 0.03 seconds per node) to take ETC into consideration.

VII. CONCLUSIONS

This work is the extended study on statistical 3D thermal analysis considering spatially correlated energy disturbances and electro-thermal coupling (ETC) effects. We have shown that ETC effect has a clear impact on the deterministic and statistical 3D thermal analysis. We propose an efficient single-node statistical thermal analysis method, SNSOR, to solve thermal resistance vectors in order to compute standard temperature deviations of fewer hot spots. To take ETC into consideration, we further propose an ET-SNSOR method based on the proposed SNSOR. Extensive experiments demonstrate that SNSOR and ET-SNSOR are both efficient and accurate, such that they can be used in statistical analysis of hot spots in practical VLSI design in huge scale without or with the consideration of ETC effects.

ACKNOWLEDGMENT

This work was supported in part by National Science Foundation of China (No.60876025) and National High-Tech Research and Development (863) Program of China (No.2008AA01Z147, No. 2007AA01Z109).

REFERENCES

[1] Sundaresan K and Mahapatra N R, "An Analysis of Timing Violations Due to Spatially Distributed Thermal Effects in Global Wires," in *Proc. DAC*, 2007, pp. 515–520.
[2] Zhan Y, et al, "Electro-thermal Analysis and Optimization Techniques for Nano-scale Integrated Circuits," in *Proc. ASP-DAC*, 2006, pp. 219–222.
[3] Yang Y H, et al, "Adaptive multi-domain thermal modeling and analysis for integrated circuit synthesis and design," in *Proc. ICCAD*, 2006, pp. 575–582.
[4] Lin S C and Banerjee K, "An Electrothermally-Aware Full-Chip Substrate Temperature Gradient Evaluation Methodology for Leakage Dominant Technologies with Implications for Power Estimation and Hot-Spot Management," in *Proc. ICCAD*, 2006, 568–574.
[5] Tsai C and Kang S, " Cell-level placement for improving substrate thermal distribution," in *IEEE Trans. on Computer-Aided Design*, 2000, 19(2), pp. 253–266.
[6] Wang T Y, et al, "3D Thermal-ADI: An Efficient Chip-Level Transient Thermal Simulator," in *Proc. ISPD*, 2003, pp. 10–17.
[7] Zhan Y, and Sapatnekar S S, "A high efficiency full-chip thermal simulation algorithm," in *Proc. ICCAD*, 2004, pp. 634–637.
[8] Wu X.-H, et al, "Area Minimization of Power Distribution Network Using Efficient Nonlinear Programming Techniques," *IEEE Trans. on CAD*, 2004, 23(7), pp. 1086–1094.
[9] Zhong Y, Wong D.-F , "Fast Algorithms for IR Drop Analysis in Large Power Grid," in *Proc. ICCAD*, 2005, pp. 351–357.
[10] Li P , "Variational analysis of large power grids by exploring statistical sampling sharing and spatial locality," in *Proc. ICCAD*, 2005, pp. 644–650.