On the Accuracy of Monte Carlo Yield Estimators

Alp Arslan Bayrakci
Department of Computer Engineering
Gebze Institute of Technology

***Abstract*—Since the statistical timing analysis of VLSI circuits became a necessity, the accuracy of the proposed statistical static timing analysis (SSTA) methods is checked with the Monte Carlo (MC) estimation methods, which are called golden. However, the MC methods can have very different levels of accuracy depending on the model used beneath the method. In this paper, we build and compare three different MC yield estimators to see the effect of the models used beneath the estimator on the accuracy.**

I. Introduction

With the shrinking sizes of VLSI technology, the variation of device parameters like gate length and threshold voltage result in significant deviations from the prescribed timing specifications of the integrated circuits (IC). As a result, estimating the timing yield of the produced ICs is a challenging problem. SSTA methods that are developed to estimate yield considering the parameter variations, are still unable to overcome the accuracy and efficiency issues and there is no widespread acceptance for any of the SSTA algorithms. This is because SSTA methods inherently make assumptions and approximations which decrease their accuracy.

The best timing yield estimation method in terms of accuracy is the Monte Carlo (MC) estimation with *big enough* number of sample points. It is called *golden* and universally used to test the accuracy of the proposed SSTA methods in the literature. In this sense, the accuracy of MC is very crucial. However, there is an ambiguity about MC methods. All MC statistical timing analysis methods can not be called golden in terms of accuracy. The most accurate timing yield estimations can be gathered by MC based on transistor level (TL) SPICE simulations (TL-MC). But, because of the unbearable cost of TL simulations, most of the SSTA proposals in the literature perform MC based on block (gate) level static timing analysis (BL-MC) for checking the accuracy of the proposed SSTA method. The proposals in the literature that use TL-MC to check the accuracy either use a small number of sample points [1] or perform experiments on very small test circuits [2].

In this paper, we investigate the relationship between the chosen method beneath the MC yield estimators and the accuracy of the resultant yield estimate. For this purpose, we produce two BL-MC methods; one based on linear (BL-MC-Lin) and the other on quadratic models (BL-MC-Quad). We also perform TL-MC as a reference genuine golden method.

II. MC Based Timing Yield Estimation

Loss is the fraction of the circuits that fail to satisfy the timing constraint. Then, *Yield*, the fraction of the circuits that fulfill the timing constraint is simply given by $1-Loss$.

This research is supported by The Scientific and Technological Research Council of Turkey (TUBITAK) under the project number 112E237.

The random variables that represent the statistical variations in the circuit are collected into an n-dimensional vector X, with a joint probability density function (PDF) denoted by $f(X)$. An indicator random variable $I^M(X)$ can be defined as follows

$$I^M(X) = \begin{cases} 1 & \text{if } d_C^M(X) > T_c \\ 0 & \text{if } d_C^M(X) \leq T_c \end{cases} \quad (1)$$

where $d_C^M(X)$ is the circuit delay computed by method M and T_c is the maximum acceptable delay or timing constraint. This indicator variable indicates whether the delay of the circuit meets the timing constraint for a given realization of the random variables in X. Then, MC estimator for loss is written as in (2).

$$Loss^M = (1/N)\sum_{i=1}^{N} I^M(X_i) \quad (2)$$

where M is the method used to compute $d_C^M(X)$, X_i's are the drawn samples according to $f(X)$ and T_c is the timing constraint.

Without loss of generality, we assume in this paper two different sources of variability: transistor gate length L and threshold voltage V_t. If these two are considered as the random transistor parameters, then the delay of a gate r can be represented as

$$d_r^M(L_r, V_{tr}, h_r, InS_r) \quad (3)$$

where L_r is the gate length and V_{tr} is the threshold voltage value for the transistors in gate r, h_r is the fanout and InS_r is the input slope for gate r. d_r^M is the delay of gate r computed by method M. Actually, we also consider high to low, low to high delays and high to low, low to high output slopes separately which makes a total of four different equations for each cell in the standard cell library. Circuit delay in (1) is computed by using the delays of the individual gates, shown by (3).

MC yield estimators use the same equation shown in (2) to estimate *Loss* but they differ by the method M they prefer to compute the gate delay shown in (3). In this paper, we will investigate three different MC yield estimators:

1) Transistor Level Monte Carlo (TL-MC): This MC estimator computes gate delays and circuit delay for each sample point by performing the industry standard, precise transistor level (TL) SPICE simulations. Therefore the method M can be replaced with TL, i.e. it estimates $Loss^{TL}$. In this paper, we use 50,000 sample points with TL-MC estimator so that the resultant $Loss^{TL}$ estimate is assumed as *actual loss* value. This actual loss is used to check the accuracy of other MC estimators.

978-1-4799-0524-9/13/$31.00 ©2013 IEEE

2) Block Level Monte Carlo Based on Linear Models (BL-MC-Lin): This MC estimator computes gate delays and thus the circuit delay for each sample point using linear gate delay model in (4). For each cell in the standard cell library, we compute the coefficients (a_i) of (4) by the least squares fitting.

$$d_r^{Lin}(X) = \sum_{i=1}^{4} a_i P_i \tag{4}$$

where P_i is L_r, V_{tr}, h_r and InS_r for i = 1, 2, 3, 4 respectively. The linear delay model is used by several SSTA algorithms in the literature [3]. In BL-MC-Lin estimator the method M in (3) is replaced with Lin and the loss it estimates by (2) can be written as $Loss^{Lin}$.

3) Block Level Monte Carlo Based on Quadratic Models (BL-MC-Quad): This MC estimator computes gate delays and thus the circuit delay for each sample point using quadratic gate delay model in (5).

$$d_r^{Quad}(X) = \sum_{i=1}^{4} b_i P_i + \sum_{i=1}^{4}\sum_{j=1}^{4} c_{ij} P_i P_j \tag{5}$$

where P_i is as explained above, a_i's b_i's and c_{ij}'s are the coefficients computed by least squares fitting. The method M in (3) is replaced with $Quad$ and the loss it estimates by (2) can be written as $Loss^{Quad}$.

III. Comparison of MC Estimators

The results are presented on the ISCAS'85 benchmark suite. Two random transistor parameters, namely the transistor gate length L and the threshold voltage V_t, are considered. Both inter and intra-die variations with spatial correlations are considered. In this model, a total $3\sigma/\mu$ ratio of 15% is assumed for both of the random parameters [4].

First of all, 50,000 sample points are generated according to the inter and intra die variation model. Using these sample points and the MC estimator in (2), three MC loss estimations are performed: $Loss^{TL}$ by TL-MC, $Loss^{Lin}$ by BL-MC-Lin and $Loss^{Quad}$ by BL-MC-Quad estimator. We have chosen a timing constraint T_c for each benchmark circuit such that the actual loss is a reasonable value around 20% for each circuit.

TABLE I. Comparison of Loss Estimations (%) Computed by Three Different Methods: TL-MC, BL-MC-Lin, BL-MC-Quad.

	$Loss^{TL}$ (Actual Loss) (%)	$Loss^{Lin}$ (%)	$Loss^{Quad}$ (%)	Accuracy Degradation
c432	19.9	1.2	18.2	11.0
c499	20.1	1.3	15.5	4.1
c880	20.1	13.6	17.8	2.8
c1355	19.9	3.4	17.5	6.9
c1908	20.0	9.6	21.5	7.3
c2670	20.0	3.3	19.4	28.5
c3540	19.9	11.1	19.3	14.7
c5315	20.1	10.5	20.7	17.7
c7552	20.0	6.4	20.9	15.4

Table I shows us that the delay model used inside the MC estimator affects the accuracy of the resultant loss estimate dramatically. $Loss^{Lin}$ is 13.3 far and $Loss^{Quad}$ 1.67 far from the actual loss on the average. The last column in Table I refers how much worse or farther is $Loss^{Lin}$ loss estimate than the $Loss^{Quad}$ loss estimate. It can be seen that MC estimator based on linear models (BL-MC-Lin) is about 12 times on the average and 28.5 times at most farther to the actual loss than the MC estimator based on quadratic delay models (BL-MC-Quad). With the results presented in the table, it is clear that the MC estimator with the linear delay models inside, gives misleading results. Several SSTA methods in the literature prefer linear delay models for efficiency and if they also use same models for the 'golden' MC method used to check accuracy, this results in a misleading accuracy outcome.

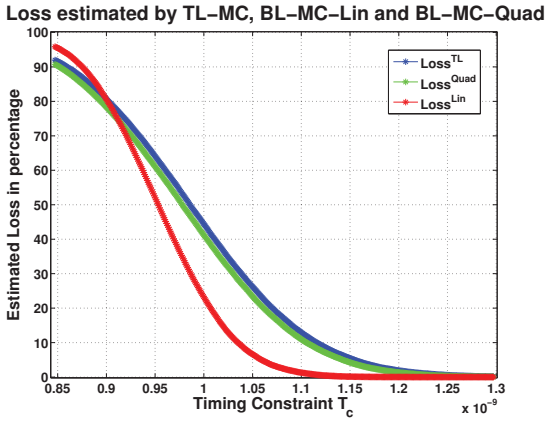

Fig. 1. Three losses: $Loss^{TL}$ computed by TL-MC, $Loss^{Lin}$ computed by BL-MC-Lin and $Loss^{Quad}$ computed by BL-MC-Quad. x-axis represents the timing constraint that is iterated to get different loss results. Circuit is c1355.

Instead of setting the timing constraint T_c to a fixed value, we have also iterated it, starting from the values causing more than 90% loss to the values causing 0% loss. The resultant loss estimations corresponding to different MC estimators are plotted in Figure 1. $Loss^{Lin}$ estimate is much worse than $Loss^{Quad}$ almost everywhere on the plot.

In conclusion, the accuracy, i.e. the unbiasedness of an MC estimator depends strongly on the underlying delay model. The resultant error of the MC estimator cannot be decreased by increasing the number of samples (N) because the error is due to the bias of the estimator. This brings the necessity to quantify the accuracy of the MC method before using it as a 'golden' method to check the accuracy. The quantification and details of the MC methods that are used to verify must be included in the papers proposing a new SSTA method.

References

[1] J.-J. Liou, A. Krstic, L.-C. Wang, and K.-T. Cheng, "False-path-aware statistical timing analysis and efficient path selection for delay testing and timing validation," in *Proceedings of the 39th annual Design Automation Conference*, ser. DAC '02, 2002, pp. 566–569.

[2] M. Merrett, P. Asenov, Y. Wang, M. Zwolinski, D. Reid, C. Millar, S. Roy, Z. Liu, S. Furber, and A. Asenov, "Modelling circuit performance variations due to statistical variability: Monte carlo static timing analysis," in *Design, Automation Test in Europe Conference Exhibition (DATE), 2011*, 2011, pp. 1–4.

[3] D. Blaauw, K. Chopra, A. Srivastava, and L. Scheffer, "Statistical timing analysis: From basic principles to state of the art," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, pp. 589–607, April 2008.

[4] A. A. Bayrakci, A. Demir, and S. Tasiran, "Fast monte carlo estimation of timing yield with importance sampling and transistor-level circuit simulation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 29, pp. 1328–1341, September 2010.