

STAC: Statistical Timing Analysis with Correlation

Jiayong Le
ECE Department, CMU
Pittsburgh, PA 15213
jiayongl@ece.cmu.edu

Xin Li
ECE Department, CMU
Pittsburgh, PA 15213
xinli@ece.cmu.edu

Lawrence T. Pileggi
ECE Department, CMU
Pittsburgh, PA 15213
pileggi@ece.cmu.edu

Abstract

Current technology trends have led to the growing impact of both inter-die and intra-die process variations on circuit performance. While it is imperative to model parameter variations for sub-100nm technologies to produce an upper bound prediction on timing, it is equally important to consider the correlation of these variations for the bound to be useful. In this paper we present an efficient block-based statistical static timing analysis algorithm that can account for correlations from process parameters and re-converging paths. The algorithm can also accommodate dominant interconnect coupling effects to provide an accurate compilation of statistical timing information. The generality and efficiency for the proposed algorithm is obtained from a novel simplification technique that is derived from the statistical independence theories and principal component analysis (PCA) methods. The technique significantly reduces the cost for mean, variance and covariance computation of a set of correlated random variables.

Categories and Subject Descriptors

B.7.2 [Hardware]: Integrated circuits—Design aids

General Terms

Algorithms, verification

Keywords

Statistical timing, process variation

1. Introduction

With the decrease in feature sizes for nanoscale CMOS technologies, the influence of process variations during manufacturing becomes increasingly important. Process variations can be broadly categorized into inter-die and intra-die variations. Inter-die variations, as they impact a single IC (integrated circuit), are fully correlated. Intra-die variations can have a random component, but also exhibit strong systematic correlations. The increased variability and associated correlations have given a new set of problems for circuit timing analysis. Traditional static timing analysis uses a corner-based methodology for worst-case analysis, which may be overly pessimistic and extremely inaccurate [1].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DAC 2004, June 7–11, 2004, San Diego, California, USA
Copyright 2004 ACM 1-58113-828-8/04/0006...\$5.00.

Many approaches have been proposed recently to deal with the statistical timing analysis problem. Most of the proposed solutions fall into one of two broad categories. The first are path-based algorithms [3,4]. The work in [3] uses the JPFDs (joint probability density functions) to take into account the correlations from both path sharing and global parameters. The recent work in [4] decomposes the correlated process parameters into a set of uncorrelated principle components and computes the circuit delay on this fixed basis. However, while path based algorithms have been shown to capture the correlations with sufficient accuracy for a single path, it becomes difficult to predict statistical circuit delay from path delay. Unlike its deterministic counterpart, statistical timing information on an output timing pin is influenced by all of its input pins. Therefore, although path-based algorithm can select multiple paths, it cannot inherently account for their statistical coupling. Furthermore, path selection often excludes potential interconnect coupling and thus makes it extremely difficult to include the impact of neighboring coupled signals.

The second category of solutions is based on block-based timing analysis algorithms [5,6]. In [5], the delays of the gates and arrival times are modeled as independent discrete random variables. The approach recently proposed in [6] models the delays as independent PDFs (probability density functions) and arrival times as CDFs (cumulative density functions). It is known, however, that including both correlation and arbitrary PDF modeling in a block-based STA approach can create exponential complexity, thereby rendering it impractical. For this reason, most block-based statistical approaches model delays as independent random variables with arbitrary PDFs [5,6]. While the independent delay model assumption renders the approach efficient, we will illustrate in Section 2 that the results can be unrealistic in the presence of inter-die variation and systematic intra-die variation. Moreover, since interconnect coupling can impact the timing as much or more than the variations, it must be considered as part of the statistical timing analysis process. However, the existing block-based and path-based STA approaches that have been recently proposed do not include handling of dominant interconnect coupling.

This paper proposes STAC, a block-based statistical timing analysis algorithm that can capture the correlations among all delay terms. STAC is based on the observation that the gate delay and circuit delay can be approximated as normal random variables without incurring substantial error. This modeling assumption allows the delays to be handled as correlated normal random variables that can be defined with only the 1st and 2nd order moments. From these moments it is straightforward to characterize their joint distributions with arbitrary correlations.

To build the joint distributions of delays and arrival times requires computing the statistical distribution of the Max and Min of a set of normal random variables. The work in [2] and more recently in [4] use a closed-form formula to approximate the mean, variance and covariance of the Max of two normal random variables based on the

formulas derived in [11]. However, the corresponding formulas for Min, which are equally important for earliest arrival time and time window computation with interconnect coupling, are not available in that paper. In STAC, we developed a PCA-based technique that can efficiently compute the statistical distribution for both Min and Max. The technique dynamically applies PCA to compute the correlations. Therefore, it does not require linear gate delay models and a fixed base as in [4]. The nonlinear gate delay model makes it possible to include large process variations, and the dynamic base significantly reduces the size of the projection matrix (which is only 2×2 in STAC) and improves numerical stability.

Along with process variations, nanoscale feature sizes also cause the dominant portion of wiring capacitance to be the inter-layer neighboring wire capacitance. For this reason, the delay of a gate can be greatly impacted by the switching activity on neighboring wires [8]. Accounting for this *cross-talk* effect, therefore, is a critical part of the statistical timing analysis process. Therefore, we modify the STAC algorithm that is proposed in this paper to demonstrate the impact of analyzing the statistical timing window for circuits with dominant interconnect coupling. The modified algorithm uses the iterative time window alignment technique in [8] to determine the aligned aggressors for a given victim. By constructing new random variables from the boundaries of different time windows, this approach inherently takes into account the correlations between different time window boundaries and reduces the unnecessary pessimism.

The remainder of this paper is organized as follows. Section 2 illustrates the validity of normal approximation with some experiment data. Section 3 describes the basic formulation and implementation of the proposed algorithm. Section 4 presents the extension of the algorithm for statistical timing analysis under interconnects coupling. Results from running STAC on various ISCAS85 benchmark circuits are listed in Section 5. Our conclusions and ideas for future work are discussed in section 6.

2. Normal Delay Distribution Approximations

The process parameter variations can be described as normal distributions [9]. The delays, however, can be theoretically modeled as arbitrary distributions for block-based STA. Arbitrary distributions provide the most generality for statistical analysis. However, they often result in exponential complexity when characterizing the joint probability density function between different random variables. Reference [5] and [6] provide some techniques to handle correlations from re-convergent fanouts, but are apparently not applicable when considering correlations from global process parameters.

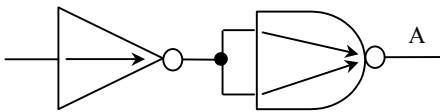


Fig 1. Simple example of signal path

In contrast, we propose that the gate delays and circuit delays can be approximated by normal distributions, even though they are not strictly normal random variables. One significant advantage with treating the delays as nearly normal distributions is that the joint distribution characterization (which includes the correlation between different variables) is greatly simplified. We will

demonstrate through some experimental results that the accuracy penalty incurred by this near normal modeling assumption is negligible compared with the error incurred by the loss of correlation information.

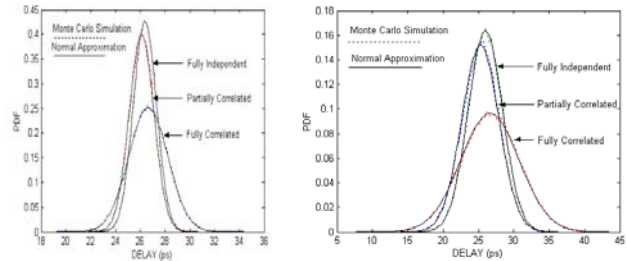


Fig 2 Circuit Delay PDF under (a) typical process variation and (b) large process variation

Based on the central limit theorem, the longer the path the more likely the delay distribution will approach a normal distribution. For this reason, we choose a small circuit with very short path for our analysis and discussions of error. Consider the circuit in Fig. 1, which is comprised of an inverter and a NAND gate, implemented in the TSMC 0.18-micron CMOS technology.

Gate delays can be modeled as linear functions of process parameters under small process variation assumptions [4,9]. However, linear approximations can lead to significant errors as process variations become large [10]. In this example, both gate delays are extracted as quadratic functions of process and environment parameters (L_{GATE} , V_{TH0} , TOX and V_{DD}) via CADENCE circuit simulation. The quadratic gate delay model may lead to non-normal gate delay distribution; however, the corresponding normal approximation can be calculated by directly matching the first two moments of the quadratic function. This approach minimizes non-Normal error even when the second order term in quadratic function is significant. In our example, Monte Carlo simulation is used to generate the PDF of the maximum output delay. The PDF of the corresponding normal approximation is also computed using the mean value and variance of the data generated by Monte Carlo simulation.

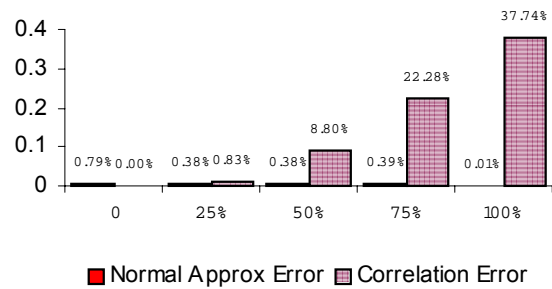


Fig 3 delay error compare for 99% in CDF

In order to validate the delay distribution under different correlation models, the process parameters are broken into global parameters and local parameters. All of the delay variables share the same global parameters, while each variable has its own local parameters. If global parameters are zero, the delays are totally independent. If local parameters are zero, the delays are totally correlated. In all other cases, the delays are partially correlated. Figure 2 shows the corresponding PDF for the circuit delay. In figure 2-a, variations for different parameters are set as $\Delta L_{GATE} = \pm 17.5\%$, $\Delta V_{TH0} = \pm 5\%$,

$\Delta TOX = \pm 5\%$ and $\Delta V_{DD} = \pm 5\%$, as applied in [9]. In figure 2-b, the variation of ΔL_{GATE} is increased to $\pm 50\%$. As shown in the plots, the PDF of circuit delay changes significantly when the correlation changes. In contrast, the normal approximation matches quite well with the Monte Carlo simulation even when ΔL_{GATE} reaches $\pm 50\%$.

Figure 3 compares the errors at the 99% point in the CDF caused by the normal approximation with that incurred by ignoring correlations between the delay variables. The x-axis values represent the ratio of global process parameters over total process parameters. When the nominal delay value is comparable with the magnitude of the variations, errors can be measured as the delay difference over total delay ($\Delta D/D$) [5,6]. However, large nominal delays for some circuits can render this error estimation overly optimistic. Therefore, the relative error in this paper is measured as the delay difference over 6σ ($\Delta \text{delay}/6\sigma$) of the delay distribution. When the global parameters (inter-die and systematic intra-die variation) dominate, the error caused by ignoring the correlation is significant, and clearly the independent delay assumption is not practical. However, note that the error from the normal approximation is negligible, regardless of the correlation. This suggests that the normal approximation is valid, even for such a small example. As the size of the logic path increases, the delay distribution will become more normal based on the central limit theorem.

3. Statistical Timing Analysis Methodology

STAC employs a block-based statistical timing analysis approach, which is performed in a breadth-first manner. The circuit is partitioned into different levels through topological sorting. The arrival time at the primary input is propagated through the gates at each level until it reaches the primary output.

3.1 Atomic Operation

A key function in statistical static timing analysis is the propagation of arrival times through the gates. In order to compute the statistical arrival time distribution at the gate output, three atomic operations are required: Sum, Max and Min. Mean, variance and covariance for the resulting random variables at the gate output are then computed with the normal distribution approximation. The remainder of this section focuses on the computation procedures for Sum, Max and Min of two random variables. Multi-variable cases are easily broken down into multiple two-variable operations.

Mean Value Table :

| | | | | |
|-----------|-----------|-----------|-----------|-----|
| x | y | z | w | ... |
| $mean(x)$ | $mean(y)$ | $mean(z)$ | $mean(w)$ | ... |

Covariance Matrix :

| | | | | |
|-------|------------|------------|------------|------------|
| | x | y | z | w |
| x | $var(x)$ | $cov(x,y)$ | $cov(x,z)$ | $cov(x,w)$ |
| y | $cov(x,y)$ | $var(y)$ | $cov(y,z)$ | $cov(y,w)$ |
| z | $cov(x,z)$ | $cov(y,z)$ | $var(z)$ | $cov(z,w)$ |
| | | | | |

Fig 4 Updated mean and covariance tables

3.1.1 Sum of two normal random variables

Computing the mean, variance and covariance for the sum of two normal random variables is straightforward. Assuming x, y, z, w are

four normal random variables and $z = x + y$, the mean, variance and covariance for variable z can be computed as

$$\begin{aligned} mean(z) &= mean(x) + mean(y) \\ Var(z) &= Var(x) + 2Cov(x,y) + Var(y) \\ Cov(z,w) &= Cov(x,w) + Cov(y,w) \end{aligned} \quad (1)$$

Assuming that x, y and w are normal random variables that have already been computed, the values on the right hand side of equation (1) exist and can be summarized in the form of a table. Figure 4 shows the updated mean and covariance table after the computation is completed.

3.1.2 Max/Min of two normal random variables

The challenge is the computation of Max and Min of two normal random variables, since the result is no longer a strict normal random variable. The works in [2] and [4] motivate the approximation of the Max as a normal distribution, and apply the closed-form formula from [11] to approximate the mean, variance and covariance of the Max of two normal random variables. However, since we also require the Min as well, in STAC, we developed a PCA-based technique that can work for both Max and Min. Since the computation for Max and Min are almost identical, the rest of this section focuses on the derivation of Max computation.

3.1.2.1 Mean value computation

The Max operation includes two random variables, $Max(x,y)$. By subtracting x from both variables, we can move it to the outside of the operator, as shown in equation (2):

$$Max(x,y) = x + Max(0, y-x) \quad (2)$$

Since x and y are both normal random variables, $y-x$ is also a normal random variable. It follows that the mean value can be computed by:

$$Mean(Max(x,y)) = Mean(x) + Mean(Max(0, y-x)) \quad (3)$$

In (3), $Mean(x)$ has already been computed. $Mean(Max(0,y-x))$ is a one-dimensional function, which can be further simplified as

$$\begin{aligned} &Mean(Max(0, y-x)) \\ &= \mu + \sigma Mean\left(Max\left(-\frac{\mu}{\sigma}, \frac{(y-x)-\mu}{\sigma}\right)\right) \end{aligned} \quad (4)$$

Where μ and σ are the mean and standard deviation of random variable $y-x$. Equation (4) is a one-dimensional function with standard normal random variables, and is therefore easily computed via table look-up.

3.1.2.2 Variance and covariance computation

By moving one random variable to the outside of the Max operator, the variance of $Max(x,y)$ can also be rewritten as

$$\begin{aligned} &Var(Max(x,y)) \\ &= Var(x) + Cov(x, Max(0, y-x)) + Var(Max(0, y-x)) \end{aligned} \quad (5)$$

In equation (5), the first term $Var(x)$ has already been computed. The third term $Var(Max(0,y-x))$ can be further simplified into equation (6) where μ and σ are the mean and standard deviation of random variable $y-x$. Like equation (4), Equation (6) can also be easily computed via one-dimensional look-up tables.

$$\begin{aligned} \text{Var}(\text{Max}(0, y-x)) &= \mu^2 + 2\mu\sigma \cdot \text{Mean}\left(\text{Max}\left(-\frac{\mu}{\sigma}, \frac{(y-x)-\mu}{\sigma}\right)\right) \\ &+ \sigma^2 \cdot \text{Mean}\left(\text{Max}^2\left(-\frac{\mu}{\sigma}, \frac{(y-x)-\mu}{\sigma}\right)\right) - \text{Mean}^2(\text{Max}(0, y-x)) \end{aligned} \quad (6)$$

The second term in equation (5), $\text{Cov}(x, \text{Max}(0, y-x))$, however, is still a two-dimensional function. Therefore STAC applies a PCA-based (Principle Component Analysis) technique to simplify its computation. Before explaining this PCA-based technique, we first review two important theorems [7].

Theorem 1: If normal random variables x_1 and x_2 are uncorrelated, they are also independent.

Theorem 2: If x_1 and x_2 are statistically independent and $f(x_1)$ and $g(x_2)$ are continuous, then $f(x_1)$ and $g(x_2)$ are statistically independent.

Based on the definition of covariance, $\text{Cov}(x, \text{Max}(0, y-x))$ can be expanded as

$$\begin{aligned} \text{Cov}(x, \text{Max}(0, y-x)) \\ = \text{Mean}\left(\left(x - \mu_x\right)\left(\text{Max}\left(-\mu_{y-x}, y-x - \mu_{y-x}\right)\right)\right) \end{aligned} \quad (7)$$

In the equation (7), $x-\mu_x$ and $(y-x)-\mu_{y-x}$ are correlated normal random variables. With PCA, they can be decomposed into two orthogonal standard normal random variables as shown in equation (8)

$$\begin{bmatrix} x - \mu_x \\ (y-x) - \mu_{y-x} \end{bmatrix} \stackrel{\text{PCA}}{=} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \quad (8)$$

Note that α_1 and α_2 are both independent standard normal random variables, and their covariance matrix is the identity matrix. Therefore, the covariance matrix of $x-\mu_x$ and $(y-x)-\mu_{y-x}$ can be expressed as

$$\begin{bmatrix} \text{Var}(x - \mu_x) & \text{Cov}(x - \mu_x, y - x - \mu_{y-x}) \\ \text{Cov}(x - \mu_x, y - x - \mu_{y-x}) & \text{Var}(y - x - \mu_{y-x}) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}^T \quad (9)$$

The covariance matrix of $x-\mu_x$ and $(y-x)-\mu_{y-x}$ is the same as the covariance matrix of x and $y-x$, which can be computed from the covariance matrix of x and y . Given that the covariance matrix is positive definite, one can compute the projection matrix P in equation (8) from equation (9). Since the base used in PCA is not fixed and the size of the matrix P is only 2×2 , this computation maintains high numerical stability and low cost, which is often difficult to achieve for PCA with large matrices and fixed base [4,7]. Then, the random variable $x-\mu_x$ can be decomposed as the linear combination of two orthogonal random variables via Gram-Schmidt Orthogonalization:

$$\begin{aligned} x - \mu_x &= \frac{p_{11}p_{21} + p_{12}p_{22}}{p_{21}^2 + p_{22}^2} (y - x - \mu_{y-x}) + w_{\perp} \\ &= k \cdot (y - x - \mu_{y-x}) + (x - \mu_x)_{\perp} \end{aligned} \quad (10)$$

In equation (10), k is the projection factor of random variable $x-\mu_x$ on $(y-x)-\mu_{y-x}$. $(x-\mu_x)_{\perp}$ is uncorrelated with $(y-x)-\mu_{y-x}$. Based on theorem 1, it is also statistically independent with $(y-x)-\mu_{y-x}$. Following theorem 2 and knowing that $(x-\mu_x)_{\perp}$ and $(y-x)-\mu_{y-x}$ are both normal random variables, $(x-\mu_x)_{\perp}$ and $\text{Max}(-\mu_{y-x}, (y-x)-\mu_{y-x})$ are also statistically independent. Substituting (10) into (7), the

covariance between x and $y-x$ can be further simplified as a one-dimensional function (11) and also computed via a look-up table:

$$\begin{aligned} \text{Cov}(x, \text{Max}(0, y-x)) \\ = k \cdot \text{Mean}\left(\left(y-x - \mu_{y-x}\right)\left(\text{Max}\left(-\mu_{y-x}, y-x - \mu_{y-x}\right)\right)\right) \\ = k\sigma^2 \text{Mean}\left(\left(\frac{y-x - \mu_{y-x}}{\sigma}\right)\left(\text{Max}\left(-\mu_{y-x}, \frac{y-x - \mu_{y-x}}{\sigma}\right)\right)\right) \end{aligned} \quad (11)$$

The same technique can be directly used to compute the covariance between any random variable w and $\text{Max}(x,y)$. Due to the limited space available in this paper, the derivation is not included.

Table 1: Built-in table for max / min computation.

| | $f(-\mu, \phi)$ | $\phi \cdot f(-\mu, \phi)$ | $f^2(-\mu, \phi)$ |
|----------------|-----------------|----------------------------|-------------------|
| $\mu < -5$ | $-\mu / 0$ | $0 / 1$ | $\mu^2 / 1$ |
| $\mu > 5$ | $0 / -\mu$ | $1 / 0$ | $1 / \mu^2$ |
| $-5 < \mu < 5$ | Table-value | Table-value | Table-value |

Substituting Max operators with Min operators in the above equations yields the Min computation for our approach. Table 1 lists the look-up tables used in the mean, variance and covariance computation for the Max and Min of two normal random variables, where ϕ is a standard normal random variable and f represents the max or min function.

3.2 Arrival Time Propagation

Once the atomic operations are available, they can be applied to compute the distribution of the arrival times at each node and level of the circuit until all primary outputs are reached.

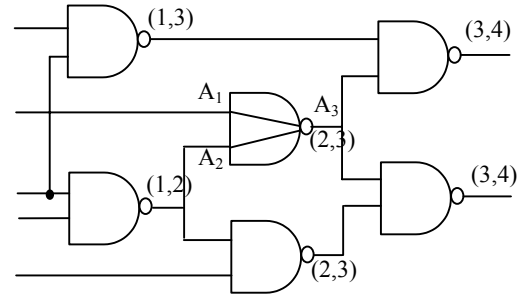


Fig 5 Timing Graph for ISCAS85 C17

Figure 5 shows the timing graph of the ISCAS85 benchmark circuit C17. As shown in the figure, once arrival times for A1 and A2 are computed, A3 can be computed as $A3 = \text{Max}(A1+D1, A2+D2)$, which requires two SUM operations and one MAX operation. Three extra variables are inserted to the mean and variance tables. During the computation, the mean value table and covariance matrix keeps growing, which may increase memory consumption. However, once the gate output arrival time (A3) is computed, its delay (D1, D2) and temporary variables (A1+D1, A2+D2) will no longer be used and can be deleted from the table. When the output arrival times for all gates in the same level have been computed, the input arrival time at that level (A1, A2...) can also be deleted. This dynamic memory allocation heuristic is very useful in the reduction of memory consumption, which becomes important with large circuits.

In the proposed algorithm, each gate in a circuit graph is assigned two level numbers via depth-first search, as shown in figure 5. The first level (compute_level) represents the order in which the gate

can be computed, which is decided by its latest computed input gate. The second level (del_level) represents the order in which the output arrival time for the gate can be deleted, which is decided from its latest computed fan-out gate. The entire algorithm for the proposed statistical timing analysis is described in the following pseudo code.

```

Sort circuit timing graph; //Set compute_level and del_level
Setup delay variables for gate; //2n variables for n-input gate
Compute mean and covariance tables for delay variables;
For level = 1: max_level,
  For each gate, if compute_level == level,
    Compute and insert output arrival time into mean and
    covariance tables;
    Delete gate delay variables and all temporary variables from
    mean and covariance tables;
  End
  For each gate, if del_level == level,
    Delete output arrival time from mean/covariance tables;
  End
End;
Save mean and covariance tables for primary output.

```

4. STAC with Coupling

In nominal static timing analysis, the problem of delay computation in the presence of crosstalk can be formulated as computing the earliest and latest arrival time among all possible waveforms of aligned aggressors. An iterative algorithm to compute the time window for a given circuit was proposed in [8]. In each loop, the early and late arrival times at the primary inputs are propagated to the primary outputs taking into account the influence of aggressor gates. The resulting timing window of each net is compared with its aggressors to decide the aligned aggressors. The aggressors whose time windows are not overlapped with the victim net will be set as unaligned aggressors in the next loop to shrink the time window. As shown in figure 6, the time windows for two nets are overlapped if and only if $t1 < t4$ and $t3 < t2$ where $t1, t3$ are the early arrival times and $t2, t4$ are the late arrival times.

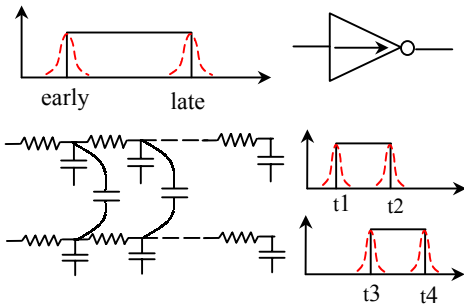


Fig 6 Statistical Timing with

Following this iterative time window alignment procedure, STAC can be extended to consider the impact of variations and coupling effects concurrently. In the scenario of statistical timing analysis, the earliest arrival time and latest arrival time for a time window of a given net become random variables, as shown in figure 6. The overlap of two time windows can no longer be simply determined by the condition $t1 < t4$ and $t3 < t2$. On the other hand, since $t1, t2, t3, t4$ are all random variables, we can construct new random variables $(t4-t1)$ and $(t2-t3)$ and redefine the overlap condition as follows:

$$\begin{aligned} \mu_{t4-t1} + 3\sigma_{t4-t1} &> 0 \\ \mu_{t2-t3} + 3\sigma_{t2-t3} &> 0 \end{aligned} \quad (12)$$

By using the 3σ values to determine the overlap of two time windows, the proposed method prevents the over-shrink of the time windows and preserves the earliest and latest arrival times. Furthermore, the correlations between different arrival times are inherently incorporated into the new random variables, which remove any unnecessary pessimism in the time window alignment. The mean value and the standard deviation for new random variable t_i-t_j can be computed from the existing mean and covariance tables:

$$\begin{aligned} \mu_{t_i-t_j} &= \text{Mean}(t_i) - \text{Mean}(t_j) \\ \sigma_{t_i-t_j} &= \sqrt{\text{Var}(t_i) - 2\text{Cov}(t_i, t_j) + \text{Var}(t_j)} \end{aligned} \quad (13)$$

5. Results

The proposed block-based statistical timing analysis approach has been implemented in C++ and we present results for various ISCAS85 benchmark circuits. These ISCAS85 circuit examples have been mapped using a library implemented with TSMC 0.18-micron technology. The quadratic delay models with different process parameters for different gates are extracted from Cadence simulations. The amplitudes for the process variation used in model extraction are $\Delta L_{\text{GATE}} = \pm 17.5\%$, $\Delta V_{\text{TH0}} = \pm 5\%$, $\Delta \text{TOX} = \pm 5\%$ and $\Delta V_{\text{DD}} = \pm 5\%$, as described in reference [9]. All process parameters are further decomposed into global process parameter components and local process parameter components to model the correlations between different delay random variables.

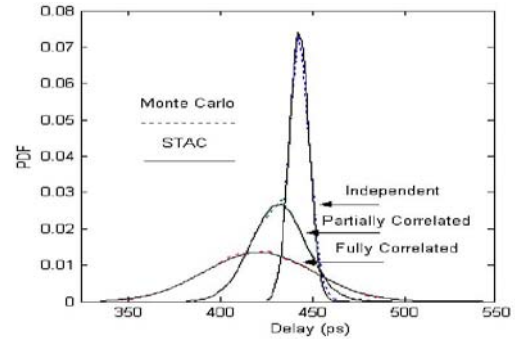


Fig 7 Delay Distribution for ISCAS85

5.1 Accuracy and efficiency

Figure 7 shows circuit delay distribution for the circuit C432 with different correlations. If gate delays are modeled as independent variables, the circuit delay distribution is characterized by a sharp PDF. When gate delays are correlated, the PDF of the circuit delay shows significantly larger variation. The difference between the delays with different strength of correlation is significant. However, the normal approximation results produced using STAC precisely approximate the delay distribution for all correlations.

Table 2 compares the error in latest arrival time incurred from ignoring the correlations versus that due to the normal approximation assumption in STAC. For the partial correlation columns in the table, the ratio of global process parameter variation

over total process parameter variation ($\Delta_{\text{Global}}/\Delta_{\text{Total}}$) is 50%. For the full correlation columns, the ratio is 100%. The comparisons are made against 10,000-run Monte Carlo simulation. The error is measured as the difference in delay value over 6σ of the delay distribution ($\Delta_{\text{delay}}/6\sigma$). As can be seen from the table, the error from the STAC approximation is almost negligible compared with the error from ignoring the correlations.

Table 2 Comparisons on 99% point of latest arrival time CDF

| 99% Max Delay | Partial Correlation | | Full Correlation | |
|---------------|---------------------|-------|------------------|-------|
| | Ind. | Stac | Ind. | Stac |
| C432 | 11.80% | 2.74% | 18.34% | 0.58% |
| C499 | 13.33% | 2.12% | 21.03% | 0.75% |
| C880 | 18.04% | 0.52% | 27.25% | 0.64% |
| C1355 | 13.85% | 0.23% | 22.06% | 0.27% |
| C1908 | 21.68% | 0.58% | 28.29% | 0.13% |
| C2670 | 20.13% | 0.24% | 30.24% | 1.96% |
| C3540 | 13.15% | 0.23% | 22.69% | 0.46% |
| C6288 | 27.63% | 0.91% | 4.86% | 0.66% |
| C7552 | 25.27% | 1.10% | 32.52% | 1.39% |

Table 3 Runtime for ISCAS85 Benchmark Circuits (Seconds)

| Circuit | C432 | C499 | C880 | C1355 | C1908 |
|---------|-------|--------|--------|--------|--------|
| STAC | 0.162 | 0.217 | 0.987 | 1.368 | 2.975 |
| MC | 97.35 | 142.42 | 419.03 | 814.8 | 1848 |
| Circuit | C2670 | C3540 | C5315 | C6288 | C7552 |
| STAC | 9.725 | 15.53 | 44.068 | 48.711 | 91.084 |
| MC | 3640 | 8206 | 16921 | 23227 | 37223 |

The run times for different benchmarks in Tables 2 are listed in table 3. In our implementation, several heuristics are applied for runtime efficiency. First, the covariance matrix is stored as a static matrix to reduce access time. Secondly, instead of physically deleting an unused variable, the variable is marked as being deleted to prevent large operation on the covariance matrix.

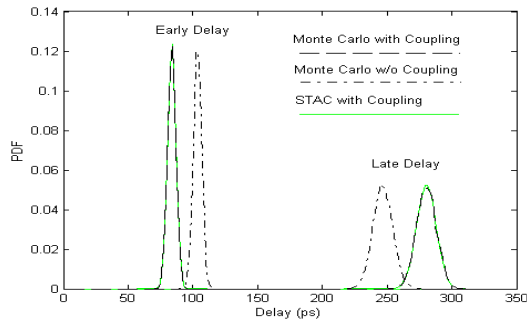


Fig 8 PDF of the time window for ISCAS85 C432

5.2 Impact of coupling

Using the C432 benchmark circuit we illustrate the accuracy of the proposed method for statistical timing analysis with coupling. In our experiment, the maximum number of aggressors for each net is limited to three, which in practice has been found to be sufficient

for most circuits [8]. For this example, the coupling analysis iterations converge very fast, in two iterations. Figure 8 shows the resulting time window at one of the primary outputs (node 370). While taking into account the interconnect coupling, the time window expands. As is shown in Figure 8, the results from the proposed algorithm match very well with Monte Carlo simulation in the presence of interconnect coupling. The Monte Carlo with coupling essentially overlaps the STAC with coupling.

6. Conclusion and Future Work

This paper presents a new block-based statistical timing analysis algorithm that can handle correlations from process parameters and re-converging paths without loss of generality. The algorithm can also take into account the interconnect coupling during timing analysis. The proposed algorithm models delays and arrival times in the circuit as normal random variables. An efficient technique is developed to compute the mean, variance and covariance of the Max and Min of two normal random variables. Results from different ISCAS85 benchmarks are presented which show high accuracy and efficiency.

As future work we plan to test the algorithm on industrial examples with extracted coupling and inter/intra die variation information.

7. Acknowledgments

This work was supported in part by National Science Foundation, Grant No. CCR-0205523, under ITR: Scalable Molecular Electronics and by the MARCO/DARPA Center for Circuits, Systems and Software, C2S2.

8. REFERENCES

- [1] M. Orshansky, L. Milor, P. Chen, K. Keutzer, and C. Hu, "Impact of Systematic Spatial Intra-Chip Gate Length Variability on Performance of High-Speed Digital Circuits," ICCAD-2000, pp. 62-67, November 2000.
- [2] S. Tsukiyama, M. Tanaka, and M. Fukui, "A New Statistical Static Timing Analyzer Considering Correlation Between Delays," in Proc. TAU, pp. 27-33, Dec 2000.
- [3] J. A. G. Jess and K. Kalafala et al, "Statistical timing for parametric yield prediction of digital integrated circuits", Design Automation Conference (DAC), pp. 932-937, June 2003.
- [4] H. Chang and S. S. Sapatnekar, "Statistical Timing Analysis Considering Spatial Correlations using a Single Pert-like Transerval", ICCAD 2003, pp. 621-625, Nov 2003.
- [5] Aseem Agarwal, David Blaauw, Vladimir Zolotov and Sarma B. K. Vrudhula, "Statistical Timing Analysis Using Bounds", DATE 2003, pp. 10062-10067.
- [6] Anirudh Devgan and Chandramouli Kashyap, "Block-based Static Timing Analysis with Uncertainty", ICCAD 2003, November 2003.
- [7] Aapo Hyvarinen, "Survey on Independent Component Analysis", Helsinki University of Technology, Lab of Computer and Information Science, Finland.
- [8] R. Arunachalam, K. Rajagopal and L. Pileggi, "TACO: Timing Analysis with Coupling" Proceedings of the Design Automation Conference, pp. 266-269, June 2000.
- [9] Sani R. Nassif, "Modeling and Analysis of Manufacturing Variations", IEEE 2001 Custom Integrated Circuits Conference, pp. 223-228
- [10] Min Cao, "Static Timing Analysis in Presence of Process Variations", Ph.D dissertation, ECE Department, Carnegie Mellon University, 2002.
- [11] C.E. Clark, "The Greatest of a Finite Set of Random Variables", Operations Research, vol. 9 pp. 85-91, 1961.