Transistor-Level Gate Model Based Statistical Timing Analysis Considering Correlations

Qin Tang, Amir Zjajo, Michel Berkelaar and Nick van der Meijs
Circuits and Systems, Delft University of Technology
e-mail: qintang@ieee.org

***Abstract*—To increase the accuracy of static timing analysis, the traditional nonlinear delay models (NLDMs) are increasingly replaced by the more physical current source models (CSMs). However, the extension of CSMs into statistical models for statistical timing analysis is not easy. In this paper, we propose a novel correlation-preserving statistical timing analysis method based on transistor-level gate models. The correlations among signals and between process variations are fully accounted for. The accuracy and efficiency are obtained from statistical transistor-level gate models, evaluated using a smart Random Differential Equation (RDE)-based solver. The variational waveforms are available, allowing signal integrity checks and circuit optimization. The proposed algorithm is verified with standard cells, simple digital circuits and ISCAS benchmark circuits in a 45nm technology. The results demonstrate the high accuracy and speed of our algorithm.**

I. Introduction

In static timing analysis (STA), the need for accuracy has driven the development of delay models. A long time industry standard is the traditional nonlinear delay model (NLDM) [1], which models gate delay and output slew as a nonlinear function of input slew (S_{in}) and output effective capacitance (C_{eff}). This only represents the signal waveform very crudely, so more recently current source models (CSMs) [2]–[7] have gained attention. Instead of modeling gate delay directly, CSM models every gate with a current source and multiple capacitors, which depend on S_{in} and C_{eff} [2] or input and output voltages [3]–[7]. This CSM representation improves delay calculation accuracy thus there is a level of industry acceptance. However, most CSMs use the assumption that only one input is switching while others are static. As a consequence, some effects such as multiple input simultaneous switching (MISS) are not modeled, leading to large errors [7]. Recently, even higher accuracy is achieved by transistor-level gate models [8]–[10] which can accurately model effects like MISS. Since most CSMs [3]–[7] and transistor-level gate models [8]–[10] have elements dependent on input and output voltages, they are called Voltage-in Voltage-out (ViVo) gate models in this paper.

The down-scaling of technology brings a significant increase in the device and interconnect manufacturing process variations, causing larger spreads in circuit timing uncertainty. To analyze the resulting variation in delay, STA can be performed at multiple corners. Although STA is accurate at every corner, the corner-based method is too pessimistic since it is close to impossible for all process parameters to have extreme values at the same time. Additionally, if the number of process variations is N_p, there are 2^{N_p} process corners—often too many to analyze. Consequently, statistical STA (SSTA) has been developed, which requires statistical gate modeling or statistical gate delay models.

Many published SSTA methods, denoted as function-based SSTA in this paper, model gate delay as a (non)linear function of process variations. The coefficients are stored in look-up tables with entries of S_{in} and C_{eff}. This modeling method is similar to the NLDM concept [1]. Most SSTA methods assume S_{in} and C_{eff} are fixed when calculating gate delay distributions. However, due to process variations in receiver and driver, both S_{in} and C_{eff} are variational. Not considering the statistical S_{in} and C_{eff} can result in 30% delay errors and even worse for bigger circuits [4]. Also, like NLDM, function-based SSTA models can not account for resistive interconnect loads and nonlinear input waveforms. Furthermore, the variational waveforms can not be obtained since only delay and slew variations are available. Additionally, function-based SSTA is entirely based on non-physical or empirical models, which is the major source of inaccuracy [3].

For these reasons, to increase accuracy, CSM models also have been extended for use in SSTA [3]–[6]. In [3], the variational voltages and all elements in CSM are modeled as a stochastic first-order expression in terms of process variations. Then the output voltage is treated as a Markovian process for delay distribution calculation. In [4], the current source value and capacitances in CSM are modeled as a quadratic Hermite function of process variations. Crossing time distributions are calculated by process variation sampling and linear interpolation. A CSM with parametric nonlinear voltage-dependent current source and parametric capacitance is used in [5] and [6]. The voltage in [5] is represented as a time-domain statistical variable and time-domain integration is performed. The gate output voltage distribution in [6] is obtained by Monte Carlo (MC) sampling. However, these methods are just verified on several simple single gates, and the correlations between input and output signals and among process variations are not considered.

To gain even higher accuracy than the above CSM methods,

This research is sponsored by the European Union and the Dutch government as part of the ENIAC MODERN project (ENIAC-120003).

978-3-9810801-8-6/DATE12/ © 2012 EDAA

and to be able to see the important effects such as MISS, in this paper we propose a statistical timing analysis solution based on statistical transistor level gate models to provide statistical information of any required crossing time. The proposed solution has the following features: 1) The variational waveform, which models several varying crossing times, is calculated and propagated through circuits; 2) In the proposed Random Differential Equation (RDE-) based statistical simulation, all input signals are considered together and calculated directly, thus fundamentally addressing MISS in statistical timing analysis; 3) As we use a common format for waveforms and elements in gate models, the correlations among input signals and between input and output signals are preserved during probability density function (*pdf*) computation; 4) Arbitrary distributions of process variations can be handled in *pdf* calculation. The proposed algorithm is verified on some simple circuits and ISCAS85 benchmark circuits considering correlations.

Compared to our previous publications on this topic [10]–[12], in this paper we contribute: 1) optimization on our transistor model for gate modeling; 2) improved algorithms to solve the RDE system in the simulation more efficiently; 3) consideration of correlations between different signals and different process parameters; 4) experiments including circuits, not just single gates.

II. Transistor-level gate modeling

Transistor-level gate models have been introduced for higher accuracy and faster characterization of STA [9], [10]. Since the gate models are constructed at the transistor level, the transistor model is a key issue which needs to have sufficient accuracy, account for the impact of process variations, while still being simple enough to be evaluated efficiently. In this paper, we use the table-based *Statistical Simplified Transistor Model* (SSTM) [10] for gate modeling. Every transistor in the circuit is modeled by a current source i_{ds} and five capacitors as shown in Fig. 1. In SSTM, gate channel capacitances, c_{gs}, c_{gb} and c_{gd}, are modeled as a function of V_{gs} and V_{ds} while junction depletion capacitances, c_{sb} and c_{db}, are represented, for simplicity, by constant values. The current and capacitances in the SSTM are modeled as a linear function of the process variations of interest ξ.

For stage-by-stage timing analysis of large circuits, the input capacitance of every gate is required. In the SSTM-based gate models, the input capacitance of a gate at any input is the sum of the gate capacitances C_g of all the transistors connected to that input, where C_g is the sum of all gate channel capacitances in SSTM. We improved on [10] by characterizing the C_g of every transistor in the library w.r.t. V_g only, based on the following concerns: i) The c_{gs},c_{gb} and c_{gd} in SSTM [10] depend on V_{ds} which is unknown for the previous gate; ii) Evaluating a simple closed-form expression is much more efficient than interpolating from matrices three times.

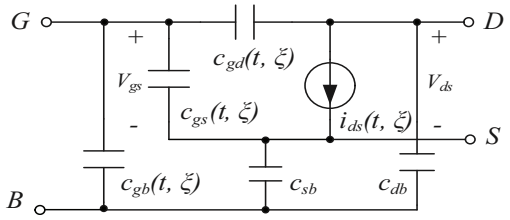

Fig. 1. Statistical Simplified Transistor Model (SSTM)

The gate models are constructed by replacing every transistor in the gate by its corresponding SSTM. To reduce the complexity of the interconnect model after RC extraction, model order reduction techniques can be employed, such as [13]. Every resistance and capacitance in the interconnect model is also modeled as a linear function w.r.t. ξ. It should be noted that the statistical timing analysis method presented in Section III is independent of the type of ViVo-gate models. In other words, it can be used with other ViVo-gate models.

By using efficient threading algorithm and multiple processors, [9] shows that it is practical to use transistor level gate models for multi-million gate STA runs to reach the combination of accuracy and speed.

III. Statistical timing analysis

For optimized ViVo-gate models, like CSMs and transistor-level gate models [3]–[10], nodal analysis (NA) or modified NA (MNA) is used for gate simulation. The NA/MNA equation solution is deterministic with all parameter values fixed. If process variations are included, typically a corner-based method is used as an outer loop. As mentioned in Section I, however, corner-based methods are pessimistic and time-consuming. In this section, we propose a RDE-based statistical simulation method which provides variational waveforms directly after simulating only once. The theoretical derivation in Section III-A is an extension based on our previous work [11].

A. RDE-based statistical simulation

The deterministic NA equation can be written in the compact format:

$$F(\dot{x}, x, t, p) = 0 \qquad p = p_0, x(t_0) = x_0 \tag{1}$$

where x denotes node voltages, $\dot{x}$ is its time derivative and p_0 is the nominal value of the process parameter vector p. Denote $x_s(t)$ as the solution of (1) which satisfies:

$$F_s = F(\dot{x_s}, x_s, t, p_0) = 0 \qquad x(t_0) = x_0 \tag{2}$$

Since all process parameters have their nominal values p_0, $x_s(t)$ is deterministic, which means it can be solved similar to ViVo-gate model based STA methods, like [8]–[10]. However, if process variations are considered the solution becomes statistical.

If we take into account process variations, $p = p_0 + \xi$ where ξ is the process variation vector which includes both global and local variations. Consequently, (1) becomes a random differential equation (RDE):

$$F_x = F(\dot{x}, x, t, \xi) = 0 \qquad \xi = p - p_0, x(t_0) = x_0 + \delta_0 \tag{3}$$

where δ_0 denotes the initial condition variation caused by process variations. It is worth noticing that the main difficulty to solve (3) is the high nonlinearity w.r.t. the random variables

ξ and the large number of process variations including local variations. If the number of local variables in the problem is very large, techniques exist to reduce it considerably [9]. In order to make (3) manageable, it is linearized[1] by a truncated Taylor expansion around x_s and p_0:

$$\begin{aligned} F_x &\approx F_s + \frac{\partial F_s}{\partial \dot{x_s}}(t)(\dot{x}(t) - \dot{x_s}(t)) \\ &+ \frac{\partial F_s}{\partial x_s}(t)(x(t) - x_s(t)) + \frac{\partial F_s}{\partial p_0}(t)\xi = 0 \end{aligned} \quad (4)$$

where F_s is defined in (2).

To simplify the notation, the variation of state variable x is denoted by y, thus $x(t)$ can be rewritten as $x(t) = x_s(t) + y(t)$. Inserting this and (2) into (4) and replacing the matrices $\partial F_s/\partial \dot{x_s}$, $\partial F_s/\partial x_s$ and $\partial F_s/\partial p_0$ with $\mathbf{C(x_s)}$, $-\mathbf{E(x_s)}$ and $-\mathbf{F(x_s)}$, respectively, we obtain:

$$\mathbf{C}(x_s)\dot{y}(t) = \mathbf{E}(x_s)y(t) + \mathbf{F}(x_s)\xi \qquad y(t_0) = y_0 = \delta_0 \quad (5)$$

$\mathbf{C}$, $\mathbf{E}$ and $\mathbf{F}$ are $N_v \times N_v$, $N_v \times N_v$ and $N_v \times N_p$ matrices respectively, where N_v is the number of unknown nodes and N_p is the number of process variations. Consequently, the nonlinear equation (3) is converted to a linear RDE in y with x_s-dependent coefficient matrices. $x_s(t)$ can be solved by well-known deterministic STA methods like in [8]–[10].

Unfortunately, the variation of state variable $y(t)$ can not be calculated directly from (5) since ξ is a random variable. According to the Random Differential Equation (RDE) theorem [14], (5) has a unique mean square solution which can be represented as (6).

$$\begin{aligned} y(t) &= \Phi(t, t_0)y_0 + \Theta(t)\xi & (6) \\ &= \mathbf{\Psi}(t)\xi & (7) \end{aligned}$$

where $\Phi(t, t_0)$ is the homogeneous solution of (5) satisfying

$$\mathbf{C}(x_s)\dot{\Phi}(t, t_0) = \mathbf{E}(x_s)\Phi(t, t_0) \quad (8)$$

and $\Theta(t)$ is an integral in the range $[t_0, t]$, which depends on Φ, $\mathbf{C}$ and each column of $\mathbf{F}$ [11]. If the initial condition x_0 is deterministic, then y_0 is zero. Since the voltage variation can be considered as zero when the signal is not switching for delay calculation, the initial condition for our problem is deterministic. Even if the initial condition y_0 is statistical due to process variations, it can also be represented as a first-order function w.r.t. ξ. Therefore, $y(t)$ is rewritten as $\mathbf{\Psi}(t)\xi$ in (7) where $\mathbf{\Psi}(t)$ is a $N_v \times N_p$ matrix.

We obtain $\mathbf{\Psi}(t)$ by substituting (7) into (5):

$$\mathbf{C}(x_s)\dot{\mathbf{\Psi}}(t) = \mathbf{E}(x_s)\mathbf{\Psi}(t) + \mathbf{F}(x_s) \quad (9)$$

After solving x_s and $\mathbf{\Psi}(t)$, $x(t)$ can be obtained based on $x(t) = x_s(t) + y(t)$ and $y = \mathbf{\Psi}(t)\xi$ in (7):

$$x(t) = x_s(t) + \mathbf{\Psi}(t)\xi \quad (10)$$

Equation (10) is used to calculate the time-varying moments of the output voltage. The first two central moments and covariance are expressed in (11)-(13), where the correlation coefficient ρ between every two process variations are included in the $E\{\xi\xi^T\}$ calculation. Since both input and output voltages have the same model w.r.t. ξ, the correlations among input and output voltages and the correlations among process variations can be easily considered during moment calculation. For more efficient memory consumption, a smaller number of normal voltage (x_s) points and their corresponding coefficients at the node of interest can be saved and propagated.

$$\begin{aligned} E\{x(t)\} &= x_s(t) & (11) \\ Var\{x(t)\} &= \mathbf{\Psi}(t)E\{\xi\xi^T\}\mathbf{\Psi}^T(t) & (12) \\ Cov\{x(t_a), x(t_b)\} &= \mathbf{\Psi}(t_a)E\{\xi\xi^T\}\mathbf{\Psi}^T(t_b) & (13) \end{aligned}$$

B. Analysis flow

The Delay distribution analysis procedure is shown in Algorithm 1. The implementation details of steps 1-6 are presented below.

Step 1. Initial condition x_0 of every gate is obtained from the data characterized in library according to the switching of nominal input signals (rising, falling or static).

Step 2. The nominal waveform $x_s(t)$ is computed by a method as commonly in CSM-based STA. In our simulation, instead of Newton-Raphson iterations, Broyden's method is used at each integration step, as it is a better fit for our table-based representation of the SSTM. Additionally, for higher efficiency, we choose linear interpolation based on triangulation [15].

Step 3. At every time point, once x_s is known, $\mathbf{C}$, $\mathbf{E}$ and $\mathbf{F}$ are updated and function (9) can be solved to obtain $\mathbf{\Psi}$. However, the high dimensionality of $\mathbf{\Psi}$ and $\mathbf{F}$ poses an additional difficulty, which is solved in Step 4.

Algorithm 1 Delay distribution calculation flow

Initialization :
SSTM-based gate models {section II}
Input waveform data (variational or deterministic)
The number k: the k^{th} node output which needs to propagate
Analysis
1. Initial condition x_0
2. STA: solve (1) for nominal value $x_s(t)$
3. Update matrices $\mathbf{C}$, $\mathbf{E}$, $\mathbf{F}$ based on $x_s(t)$
4. Solve (9) for $\mathbf{\Psi}$ by following iterations:
for $j = 1$ to N_p **do**
  if $F_j(x_s) \neq \mathbf{0}$ **then**
    solve $\mathbf{C}(x_s)\dot{\Psi}_j(t) = \mathbf{E}(x_s)\Psi_j(t) + F_j(x_s)$
  else
    $\Psi_j(t) = \mathbf{0}$ {$\mathbf{0}$: empty vector}
  end if
end for
Ψ_j and F_j are the j^{th} column of $\mathbf{\Psi}$ and $\mathbf{F}$, respectively
5. Save the output data for propagation: $x_{sk}(t)$ and $\Psi_{(k)}(t)$
6. Compute the delay distribution {section III-C}

[1]Higher accuracy can be obtained by using higher-order models or piecewise linear simulation method at the cost of complexity.

Step 4. Based on moment matching, (9) is split into N_p ordinary differential equations (ODEs):

$$\mathbf{C}(x_s)\dot{\Psi}_j(t) = \mathbf{E}(x_s)\Psi_j(t) + F_j(x_s) \quad j = 1 : N_p \qquad (14)$$

where F_j and Ψ_j are the j^{th} column of $\mathbf{F}$ and $\mathbf{\Psi}$, respectively. After using a numerical integration method, due to x_s-dependent coefficients $\mathbf{C}$, $\mathbf{E}$ and F_j, (14) becomes a linear algebraic equation (LAE), which means that the LAE can be solved fast without the necessity of root-finding iterations. Only LU decomposition, and forward and backward substitution are needed to solve the LAE. Additionally, the same coefficients $\mathbf{C}$ and $\mathbf{E}$ of N_p ODEs in (14) requires LU decomposition only once to solve these N_p ODEs.

Step 5. The k^{th} node voltage, which needs to be stored and propagated (denoted as $v(t)$), can be expressed as:

$$v(t) = x_{sk}(t) + \Psi_{(k)}(t)\xi \qquad (15)$$

where $x_{sk}(t)$ and $\Psi_{(k)}(t)$ are the k^{th} element of $x_s(t)$ and the k^{th} row of $\mathbf{\Psi}(t)$ respectively.

C. Computing the delay distribution

For timing analysis, the problem of interest is to compute the moments of arrival time, gate delay or in general crossing time. The crossing time t_η is defined as the first time for voltages to cross the threshold voltage $V_\eta = \eta\% \cdot V_{dd}$. The *cdf* of crossing time is calculated when the nominal voltage is in transition. For a rising transition this is expressed as:

$$\begin{aligned} F_n &= P(t_\eta \le t_n) = 1 - P(t_\eta > t_n) = 1 - G_n && (16)\\ G_n &= P(v_1 \le V_\eta \cap v_2 \le V_\eta \cap \cdots \cap v_n \le V_\eta) && (17)\\ &= P(v_n \le V_\eta | v_{n-1} \le V_\eta, \ldots, v_1 \le V_\eta) \cdot G_{n-1} && (18)\\ &= P(v_n \le V_\eta | v_{n-1} \le V_\eta) \cdot G_{n-1} (n = 2 : N) && (19)\\ &= \frac{P(v_n \le V_\eta \cap v_{n-1} \le V_\eta)}{P(v_{n-1} \le V_\eta))} \cdot G_{n-1} && (20) \end{aligned}$$

where v_i is the voltage of interest at time t_i and F_n denotes the *cdf* of crossing time at time t_n. Equation (18) is rewritten in (19) since the voltages are modeled as Markovian processes [3], [12]. Based on (16) to (20) an iteration method is used to calculate the *cdf* of the corresponding crossing time with initial condition G_1=1. Given the moments and covariances calculated in the RDE-based statistical simulator in (11)-(13), the joint probability and single probability in (20) can be obtained easily.

The relationship between the *cdf* ($F(t)$) and the discretized *pdf* ($f(t)$) in our algorithm is illustrated in Fig. 2. To simplify the calculations, the *cdf*s and *pdf*s have these properties: i) $F = 1$ if $F \ge F_{max}$ and $F = 0$ if $F \le F_{min}$. The time t_{start} and t_{end} correspond to F_{min} and F_{max} shown in Fig. 2, respectively; ii) $f(t)$ is calculated during the period $[t_{start}, t_{end}]$, hence the *pdf* has values only on the definite interval $[t_{start}, t_{end}]$. Let $t_{n'} = (t_{n-1} + t_n)/2$, then the discretized *pdf* is approximated by the following method: $f(t_{n'}) = \int_{t_{n-1}}^{t_n} f(t)dt = F(t_n) - F(t_{n-1})$ where $f(t_1) = 0$.

The effective *cdf* is defined as the *cdf* within $[t_{start}, t_{end}]$. If the simulation uses a non-uniform time step algorithm, the effective *cdf* needs to be uniformly sampled for *pdf* computation. After uniformly sampling and interpolating from the effective *cdf* with N_s samples, the $N_s \times 1$ time and *cdf* vectors are obtained and denoted as T_1 and cdf_u, respectively. These vectors are used to calculate the *pdf* vector Ω with element $\Omega_k = cdfu_k - cdfu_{k-1}$ ($\Omega_1 = 0$, $k = 2 : N_s$).

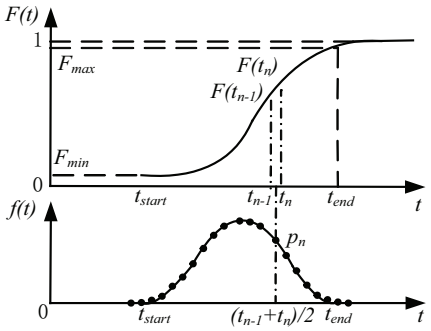

Fig. 2. Cumulative distribution function (*cdf*) and discretized probability density (*pdf*) function.

The last step is to calculate the moments of crossing time (mean μ, standard deviation σ and skewness γ). Denoting T_1^T as the transposition of the column vector T_1, the calculation method can be formulated as following:

$$\mu = T_1^T \Omega \qquad \sigma = T_2^T \Omega - \mu^2 \qquad (21)$$

$$\gamma = (\Gamma - 3\mu\sigma^2 - \mu^3)/(\sigma^3) \quad (\Gamma = T_3 \Omega^T) \qquad (22)$$

The relationships between the elements of T_2 and T_3 with T_1 are $T_2(k) = T_1^2(k)$ and $T_3(k) = T_1^3(k)$ where $k = 1 : N_s$.

The calculation method for a falling transition is similar to the above methods with the only difference in (17) where v_i is replaced by $V_{dd} - v_i$. If the waveform is non-monotonic and crosses V_η multiple times, the method above is used to iteratively find all crossing times.

D. Complexity analysis

As shown in Algorithm 1, the majority of the runtime is consumed in step 2 to calculate the nominal value x_s and in step 4 to compute the sensitivities $\mathbf{\Psi}$. Therefore, $T_{SSTA} \approx T_{STA} + T_{\mathbf{\Psi}}$, where T_{SSTA} is the runtime of the whole statistical timing analysis algorithm, T_{STA} is the runtime of step 2 and $T_{\mathbf{\Psi}}$ is the time of step 3-4. Step 2 can be solved by ViVo-gate model based STA procedures [3]–[6], [8]–[10], and its complexity depends on the gate models used. For the proposed SSTM-based gate models, the method proposed in [8], [9] also can be used. Compared to traditional ViVo-gate model based STA, our statistical timing analysis method requires extra runtime $T_{\mathbf{\Psi}}$.

Step 4 has complexity $O(N_p)$. There are 5-7 most important process parameters such as length and threshold voltage. Fortunately, the local variations can be collapsed into a much smaller number of variations (or even one variation [9]) after using Principle Component Analysis-like methods [16]. In our method, after using numerical integration the solving procedure of the N_p ODEs in (14) requires LU decomposition only once. Furthermore, to calculate $\mathbf{\Psi}$, no root-finding iterations are necessary. Therefore, compared to T_{STA}, $T_{\mathbf{\Psi}}$

is approximately proportional to $\frac{N_p}{N_{iter}} \cdot T_{STA}$ if the average number of iterations at each integration step in step 2 is N_{iter}. This is more efficient than a corner-based method.

IV. Correlations of Variational Waveforms

During statistical timing analysis, the correlation of signals caused by process variations and path re-convergence should be considered and efficiently simulated. Fig. 3 indicates the delay standard deviation (σ) and delay skewness (γ) of a NAND2 with respect to different correlation coefficient (ρ) of input arrival times, with different nominal arrival time difference (dt). ρ changes from 0 to 0.9. Input signals are simple ramps with the same arrival time variance ($\sigma_t = 10ps$), but different arrival time means (μ_t). It should be noted that, when μ_ts are far away from each other ($dt = 6\sigma_t$), the correlation has significantly less impact on delay distribution and therefore can be ignored.

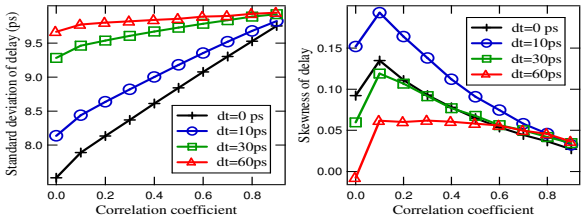

Fig. 3. The importance of correlations

If more than one input switches in a multi-input gate, the 50% crossing time $cdfs$ of the switching inputs can be calculated and the corresponding effective cdf range mentioned in section III-C is used. Take NAND2 as a case and denote the effective cdf range of two inputs as $[t_{start1}, t_{end1}]$ and $[t_{start2}, t_{end2}]$, separately. If $t_{start2} - t_{end1} > \varepsilon$ or $t_{start1} - t_{end2} > \varepsilon$, the correlation between two inputs will be ignored and the latest/earliest input or inputs will be propagated while the other is assumed static. However, if the effective cdf ranges are overlapping, all stochastic correlated inputs must be considered.

V. Experimental Results

The effectiveness and accuracy of the proposed approach was evaluated on some most commonly used standard cells and ISCAS85 benchmark circuits using the GVT library in the latest Nangate 45nm package [17]. The SSTM is characterized based on the simulated data using a full BSIM4 model, and every gate model is constructed by replacing every transistor in the gate by its corresponding SSTM. The whole algorithm is implemented in Matlab in a computer with single processor.

SSTM-based deterministic delay calculation for STA: Since the statistical simulation depends on the nominal value computation (x_s in (2)), the accuracy of the proposed SSTM-based gate models for deterministic timing analysis (no process variations) is important. It was tested by the minimum-sized standard cells listed in Table I. Every switching input signal is a ramp with input slew varying from $7.5ps$ to $600ps$ and the load capacitance changes from $0.40fF$ to $25.6fF$. Both rising and falling inputs are simulated. The μ and σ of gate delay relative errors and output slew relative errors after thousands of simulations are listed in the Table I. The scenarios that all input signals switch at the same time are also included in the experiments. The results indicate a high accuracy of deterministic delay and slew calculation by using our SSTM-based gate modeling.

TABLE I
THE MOMENTS OF DELAY ERROR AND OUTPUT SLEW ERROR FOR DETERMINISTIC DELAY AND SLEW CALCULATION

Standard cells	delay error		slew error	
	μ	σ	μ	σ
INV	0.2135%	0.2122%	0.1882%	0.7504%
NAND2	0.4828%	0.2346%	0.0244%	1.1621%
NOR2	0.4521%	0.2557%	0.2659%	0.7047%
AND2	0.7842%	0.2659%	0.0448%	0.8974 %
XOR2	0.0782%	0.4304%	0.2748%	0.5666%
BUF	0.3633%	0.2109%	0.2765%	0.9400%
MUX2	0.2996%	0.2463%	0.0304%	0.5285%
AOI21	0.3540%	0.2412%	0.0947%	0.8084%
AOI211	0.7570%	0.3503%	0.0651%	1.0603%
NAND4	0.9568%	0.3468%	0.7210%	1.6800%

Statistical timing analysis considering MISS: In order to evaluate the capability of our statistical simulation method for multiple variational inputs, we applied our approach in circuits with up to four inputs. All inputs of every gate are variational with signal correlations and have high possibilities to switch near-simultaneously (MISS). The multi-input cells are $NAND2$, $NOR2$, $NOR3$, $NAND3$, $AOI21$, $AOI211$, $AOI22$ and $NAND4$. Every variational input signal is modeled as a ramp signal of $40ps$ input transition time with voltage variations. The σ of voltages and the arrival time differences among input signals are varied to obtain results at diverse scenarios. The correlation among every two voltage variations varies from 0 to 0.8. All the statistical simulation results are compared to $10K$ Spectre Monte Carlo (MC) simulations. Fig. 4 illustrates the relative errors of all the experiments. Most of μ relative errors are within 1% while AOI211 has over 1% relative errors when correlation coefficient is 0.8 and variance is large. All the σ relative errors are within 6% except two biggest σ cases (6.42% and 6.71%) coming from NAND4 and AOI21 respectively. All of the skewness errors are within 8%. The average μ, σ and γ relative errors are 0.38%, 2.30% and 2.87% respectively. Fig. 5 shows the discrete pdf with 50 samples and the histogram of MC simulation in Spectre of AOI21. All inputs have the exact same mean value of arrival times (MISS). The discrete pdf was scaled to provide a straightforward shape comparison.

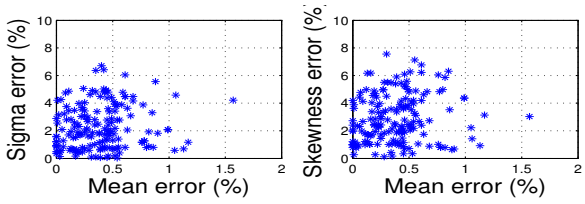

Fig. 4. All moment percentage relative errors comparison

Statistical timing analysis with L_{eff} and V_{th} variations: Effective length L_{eff} and threshold voltage V_{th} are chosen as the representative process variables, which both have 3σ equal to 20% of the mean value with correlation coefficients of 0, 0.2, 0.5 and 0.8. We firstly applied the proposed method

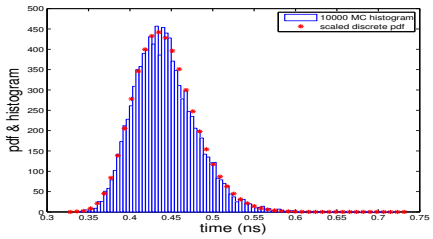

Fig. 5. The *pdf* and histogram comparison of AOI21

to nine common standard cells with different input transitions. Fig. 6 illustrates the average relative errors (absolute values) of μ and σ for nine common standard cells. The worst σ relative errors are -4.03% and 3.04% from AOI211 and XOR2 with falling output respectively. Fig. 7 shows what the variational waveforms look like. The discrete *pdf* of the 50% crossing time is shown on the upper right corner of Fig. 7.

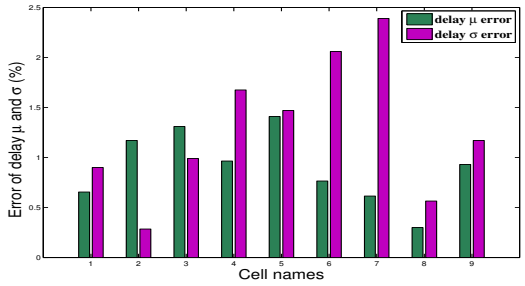

Fig. 6. RDE-based simulator for standard cells. The numbers of 1 to 9 stand for INV, NAND2, NOR2, BUF, AND2, XOR2, AOI211, NAND4 and MUX2 respectively.

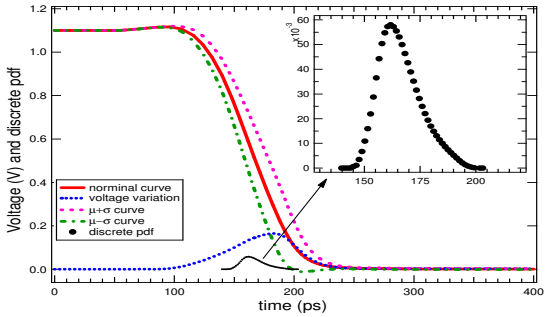

Fig. 7. Variational waveforms of AOI211 and its discrete *pdf*. The discrete *pdf* of 50% crossing time is amplified on the upper right corner.

Secondly, we used the proposed transistor-level statistical timing analysis method for some circuits listed in Table II, where the absolute μ and σ relative errors of delay distribution calculation are included. The gates with more than 3 inputs in C432 and C499 are replaced with several logic gates with no more than 3 inputs provided by gate library. 10K Monte Carlo-based Spectre simulations are used for accuracy and efficiency comparison. The correlation coefficient (ρ) among process variations are considered to be constant 0 (independent), 0.2, 0.5 and 0.8. The results show high accuracy of our method.

Compared to Spectre MC runs, our method achieves $712\times$ speed-up on average. The ratio T_{Ψ}/T_{STA} of C17, Adder, C432 and C499 are $1/2.27$, $1/2.25$, $1/3.68$ and $1/2.25$ respectively. It indicates that $T_{SSTA} \approx 1.383 \cdot T_{STA}$ on average for two process variations. It is expected that using a more efficient solving method would contribute to even higher efficiency.

TABLE II
THE ABSOLUTE VALUES OF DELAY μ AND σ RELATIVE ERRORS (UNIT: %) OF SOME CIRCUITS WITH DIFFERENT CORRELATION COEFFICIENTS ρ, COMPARED WITH 10K SPECTRE MC RESULTS

ρ	0		0.2		0.5		0.8	
name	μ	σ	μ	σ	μ	σ	μ	σ
C17	0.50	0.35	0.47	2.33	0.27	2.52	0.44	2.52
Adder	0.01	0.05	0.54	0.40	1.00	2.28	1.04	2.63
C432	0.18	2.00	0.71	1.46	0.88	1.04	1.15	0.50
C499	0.81	2.19	0.37	0.95	0.47	2.32	1.07	2.97

VI. CONCLUSIONS

In this paper, we have presented a new transistor-level gate model based statistical timing analysis method. The gate models are constructed based on statistical simplified transistor models for higher accuracy. Correlations among input signals and among process variations are preserved during simulation since the voltages and all elements in gate models have the same model format. Furthermore, the multiple input switching problem is addressed by considering all input signals together for output information. The variational waveforms of the gate output are calculated by RDE-based statistical simulations, which is used for delay distribution calculation. The experiments demonstrate the high accuracy and efficiency of the proposed method for both deterministic delay calculation and statistical timing analysis.

REFERENCES

[1] J. F. Croix and D. F. Wong, "A fast and accurate technique to optimize characterization tables for logic sythesis," in *DAC*, 1997, pp. 337–340.
[2] Synopsys. CCS timing white paper. [Online]. Available: www.synopsys.com/products/solutions/galaxy/ccs/cc_source
[3] H. Fatemi, S. Nazarian, and M. Pedram, "Statistical logic cell delay analysis using a current-based model," in *DAC*, 2006, pp. 253–256.
[4] A. Goel and S. Vrudhula, "Statistical waveform and current source based standard cell models for accurate timing analysis," in *DAC*, 2008, pp. 227–230.
[5] B. Liu and A. B. Kahng, "Statistical gate level simulation via voltage controlled current source models," in *BMAS Workshop*, Sept. 2006, p. 23.
[6] B. Liu, "Gate level statistical simulation based on parameterized models for process and signal variations," in *ISQED*, 2007, pp. 257–262.
[7] C. Amin, C. Kashyap, N. Menezes, k. Killpack, and E. Chiprout, "A multi-port current source model for multiple-input switching effects in cmos library cells," in *DAC 2006*, June 2006, pp. 247–252.
[8] P. Kulshreshtha, R. Palermo, M. Mortazavi, C. Bamji, and H. Yalcin, "Transistor-level timing analysis using embedded simulation," in *ICCAD*, 2000, pp. 344–349.
[9] S. Raja, Varadi, M. Becer, and J. Geada, "Transistor level gate modeling for accurate and fast timing, noise, and power analysis," in *DAC*, 2008, pp. 456–461.
[10] Q. Tang, A. Zjajo, M. Berkelaar, and N. P. van der Meijs, "Transistor level waveform evaluation for timing analysis," in *VARI*, 2010, pp. 1–6.
[11] ——, "RDE-based transistor-level gate simulation for statistical static timing analysis," in *DAC*, 2010, pp. 787–792.
[12] ——, "Statistical delay calculation with multiple input simultaneous switching," in *ICICDT*, 2011, pp. 1–4.
[13] X. Li, P. Li, and L. T. Pileggi, "Parameterized interconnect order reduction with explicit-and-implicit multi-parameter moment matching for inter/intra-die variations," in *ICCAD*, 2005, pp. 806–812.
[14] T. T. Soong, *Random diffrential equations in science and engineering*. New York: Academic Press, 1973.
[15] J. P. Wilson and J. C. Gallant, *TERRAIN analysis: principlea ans applications*. John Wiley & Sons, Inc., Canada, 2002.
[16] D. F. Morrison, *Multivariate statistical methods*. New York: McGraw-Hill, 1976.
[17] Nangate 45nm open cell library. [Online]. Available: http://www.nangate.com/index.php?option=com_content&task=view&id=137&Itemid=137