# RDE-Based Transistor-Level Gate Simulation for Statistical Static Timing Analysis

Qin Tang, Amir Zjajo, Michel Berkelaar, Nick van der Meijs
Circuits and Systems Group, Delft University of Technology
Q.tang@tudelft.nl

## ABSTRACT

Existing industry-practice statistical static timing analysis (SSTA) engines use black-box gate-level models for standard cells, which have accuracy problems as well as require massive amounts of CPU time in Monte-Carlo (MC) simulation. In this paper we present a new transistor-level non-Monte Carlo statistical analysis method based on solving random differential equations (RDE) computed from modified nodal analysis (MNA). In order to maintain both high accuracy and efficiency, we introduce a simplified statistical transistor model for 45nm technology and below. The model is combined with our new simulation-like engine which can do both implicit non-MC statistical simulation and deterministic simulation fast and accurately. The statistics of delay and slew are calculated by means of the proposed analysis method. Experiments show the proposed method is both run time efficient and very accurate.

## Categories and Subject Descriptors

J.6 [**Computer-Aided Engineering**]: Computer-Aided Design; I.6.5 [**Simulation and Modeling**]: Model Development; G.1.7 [**Ordinary Differential Equations**]: Differential-Algebraic Equations;

## General Terms

Algorithms, Performance, Design, Theory

## Keywords

transistor-level modeling, non-Monte Carlo, statistical static timing analysis.

## 1. INTRODUCTION

In recent years, the increasing impact of process variations (PVs) in the successive technology nodes has made timing analysis more and more challenging. One way to analyze the impact of the PVs is to use statistical representations of timing in SSTA. This paper tries to pave the road for highly accurate yet run time efficient SSTA. We see several weaknesses in the current industry practice SSTA solutions, which are getting worse for every technology generation. Firstly, SSTA primarily uses gate-level models (GLMs), such as current source models, for standard cells. Therefore, the accuracy of GLMs is crucial for timing analysis.


*DAC'10*, June 13-18, 2010, Anaheim, California, USA.


However, GLMs fail to work with a multi-port coupled interconnect load since the complicated load is modeled as an effective capacitance ($C_{eff}$) only in order to simplify the model. They also fail to deal with noisy or atypical input signals which are modeled as a saturated ramp only, and fail to capture (near-) simultaneous multi-input switching (MIS) and internal charging effect for high-stack and some complicated cells [1-2]. Secondly, separate calculations for cell and wire delay adds to the inaccuracy, as does just representing signal waveforms by the delay and slew. Thirdly, generating SSTA models for all standard cells of a library takes a huge amount of CPU time due to the necessary MC-based simulations, up to several months [3]. The accuracy issues result mainly from the black-box-based property of GLMs, ignoring intrinsic circuit behaviors.

Intensive efforts have been made to obtain more accurate GLMs. For the input waveform, a Weibull-based waveform model is proposed in [4]. Some other papers model gate current and capacitance (or node charge) as a function of port voltages to build waveform-independent models [1-2, 5]. In [5] a non-linear $C_{eff}$ model is described although its accuracy still needs to be evaluated further. Some more advanced GLMs for MIS issues model internal nodes to achieve better accuracy [2]. They attempt to optimize GLMs to maintain acceptable accuracy for all types of gates but unfortunately the black-box-based property is the essential root of those issues. A good model should be independent of input waveform and output load, should get full access to waveform data, should be able to capture the important circuit behaviors, and should have high accuracy and efficiency as well. The extreme method is running Spice or Spectre to do transistor-level simulation using sophisticated transistor models like BSIM4 model. However, the complicated device models dramatically slow down the simulation. Therefore, a fast transistor model is necessary for transistor-level timing analysis. Simplified transistor models can be categorized to three types: *i)* look-up table (LUT) methods combined with advanced interpolation [6], *ii)* polynomial function of terminal voltages [7], and *iii)* simplified transistor models based on device physical properties [8-9]. The first two methods ignore the device physical behaviors which do not allow statistical extension easily. For instance, by using LUTs, the sensitivities of current to process variations are functions of time-varying terminal voltages, requiring a lot of CPU time to characterize the sensitivity LUTs. In contrast, by using closed forms, the sensitivities to variations can be obtained from the partial differential of the current expressions directly.

Relatively few papers mention the stochastic waveform propagation problem for SSTA, because the dominant method is MC simulation. A stochastic waveform propagation method based on current-source-based GLMs in [10] is efficient for simple GLMs without input and Miller capacitors. Nevertheless, the Miller capacitor is important for GLMs [11] and the elements of

the gates (current source and capacitances) could be statistical as well. Once these issues are considered, it is not clear whether the method in [10] is still practical.

In this paper, we propose three new building blocks:

1) A statistical simplified transistor model (SSTM) for transistor-level SSTA: no more black-box models;
2) A non-Monte Carlo statistical simulation method based on solving RDEs [12] which describe the statistical behavior of the circuits. Given the mean ($\mu$) and standard deviation ($\sigma$) of process variations and/or input waveforms at every time point, the proposed simulation method calculate the $\mu(t)$, $\sigma(t)$ and covariance of output voltages and probability density functions (*pdf*) of $\eta$% crossing time, which are used to derive the *pdf* of delay and output slew;
3) Full access to waveforms rather than abstracting them to delay and slew values. A stochastic waveform propagation method is provided in the RDE-based statistical method.

Our simulation engine supports statistical models intrinsically, no more need for lengthy MC runs. The proposed design flow is shown in Fig 1. The simulation results show high accuracy for both deterministic and statistical timing analysis.

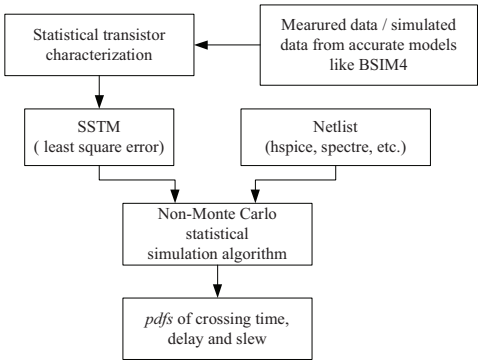


**Fig 1 The flowchart of our proposed method**

## 2. STATISTICAL SIMPLIFIED TRANSISTOR MODEL

The critical foundation of transistor-level timing analysis is the transistor model, which must be carefully built, considering an accuracy-efficiency trade-off. Our target is to develop a transistor model which captures sufficient second-order effects to allow accurate waveform and delay calculation in digital design.

The BSIM4 model is regarded as one of the most accurate transistor models but it is also the most complicated model with several hundred parameters. The currents and capacitances of the BSIM4 model are determined by solving complex equations, which are functions of many process parameters. Consequently, when they are applied to simulators which calculate node voltages and branch currents by solving ordinary differential equations, simulation becomes unacceptably slow for large digital circuits. The widely-used gate-level models are the so-called current source models which typically model every gate by several capacitors and a current source [11], as shown in Fig 2a. Every gate (or, more accurately, arc) is basically modeled as a single transistor. Although the current-source GLM is less accurate in its representation of a gate for nanometer technology, the simple model is, however, appropriate for transistors. The proposed SSTM shown in Fig 2b represents transistors by a nominal current source $I_{ds}$, a statistical current source $\delta_{ids}$ caused by process variations and five parasitic capacitances which have statistical parts with respect to random parameters of interest as well.

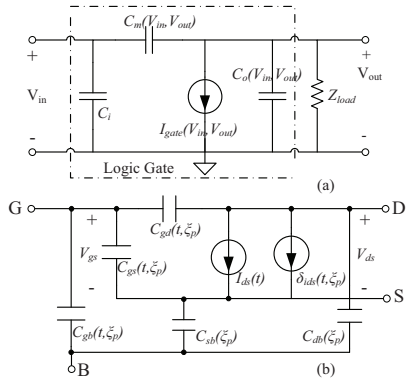


**Fig 2 a) Current-source model; b) Proposed SSTM**

### 2.1 I-V model

Conventionally, the Shockley-Sah MOSFET model was widely used for fast circuit simulation due to its simplicity. Since it is unable to reproduce *I-V* characteristics of short channel transistors, the Shichman-Hodges model was proposed to take channel length modulation into account. As these models do not consider the second-order effects of deep sub-micron MOSFETs, they were gradually replaced by deep submicron MOSFET models (DSMM) [13]. Although, in general, this model improves accuracy substantially for sub-micron MOSFET behavior, our experiments in a 45nm CMOS technology still show significant errors: *i)* due to channel length modulation, drain induced barrier lowering (DIBL) and the substrate current induced body effect, the channel length modulation parameter $\lambda$ is a complicated function of $V_{gs}$ and $V_{ds}$. As a consequence, the method to model the saturation current as a linear function of $V_{ds}$ with constant slope starting from $I_{ds}(V_{dsat})$ is not accurate enough; *ii)* in the linear region, $I_{ds}$ is no longer proportional to ($V_{gs}$-$V_{th}$-½$V_{ds}$). In fact, the ½ should be replaced by a factor which depends on $V_{gs}$-$V_{th}$; *iii)* in 45nm and below, the cut-off current can not be ignored any more. Simulation results show that if $V_{gs}$ is smaller than $V_{th}$ by a small amount, the current still has similar shape as the currents when $V_{gs}$>$V_{th}$, which cannot be modeled as zero if the input slew and load capacitance are small.

Similarly, the $\alpha$-power law MOSFET model [9] is also widely used in digital circuit simulation. This model assumes that near- and sub-threshold region modeling is not important in delay calculation for digital circuits, so the linear region current is only approximated by linear lines and the saturation region current is constant. However, if the load capacitance and input slew are both quite small, the inaccuracy of the linear-region current significantly impacts the output waveform at the end of the transition, introducing a large error for output slew. Taking these issues into consideration, the proposed BSIM4-based nominal *I-V* model of SSTM in equation form is given below:

$$I_{ds} = \begin{cases} He^{(V_{gst}/nV_t)}(1-e^{(-V_{ds}/V_t)}) & V_{gs} < V_{th} \\ \dfrac{W}{L}\cdot\left\{JV_{gst}V_{dseff}\left(1-\dfrac{V_{ds}}{2V_b}\right)\Big/\!\Big/\left(1+\dfrac{V_{ds}}{V_c}\right)\right\}\cdot\left[1+\lambda(V_{ds}-V_{dseff})\right] & V_{gs} \geq V_{th} \end{cases} \quad (1)$$

where $V_{gst}$ is $V_{gs}$-$V_{th}$ and $V_t$ is the thermal voltage. Other terms are described below:

$$V_b = V_{gst} + 2V_t \qquad V_c = E_{sat} \cdot L \tag{2}$$

$$V_{dseff} = V_{dsat} - \frac{1}{2}\left(V_{dsat} - V_{ds} - \gamma + \sqrt{(V_{dsat} - V_{ds} - \gamma)^2 + 4\gamma V_{dsat}}\right) \tag{3}$$

$$V_{dsat} = V_c \cdot \left(V_{gst} + 2V_t\right) \Big/ \left(V_c + V_{gst} + 2V_t\right) \tag{4}$$

In order to link the continuous linear current with the saturation current, a smooth function (3) based on BSIM4 is used. $V_{dseff}$ enables a unified expression for both linear and saturation currents. Since the $V_{dsat}$ expression in BSIM4 is simple, we use it directly here. $\gamma$ in (3) partially determines the shape of the *I-V* curve in the transition region between the linear and saturation regions. For the 45nm PTMLP technology [14], $\gamma$=0.06 is sufficiently accurate for NMOS and PMOS transistors. Instead of using complicated expressions, the parameter *J* considers several effects, including mobility degradation. It should be noticed that the cut-off current could simply be modeled as zero if sharp input ramps and extremely small load capacitances are rare in digital design. Then the proposed model is simplified further to the 2nd equation in (1) where only *J* and $\lambda$ are obtained in the characterization stage. Threshold voltage $V_{th}$ in (1) divides the *I-V* plane to two parts, thus accurate $V_{th}$ modeling is important. According to the BSIM4 model, a linear dependence of $V_{th}$ on $V_{ds}$ is a sufficient approximation. We simplify the $V_{th}$ model as:

$$V_{th} = V_{th0} - \alpha \cdot V_{ds} + K_1\left(\sqrt{\Phi_s - V_{bs}} - \sqrt{\Phi_s}\right) - K_2 \cdot V_{bs} \tag{5}$$

where $V_{th0}$ is the zero-biased long-channel device $V_{th}$, and $\alpha$ is a coefficient for drain/source charge sharing and DIBL effects on $V_{th}$. The coefficients $K_1$, $K_2$ and surface potential $\Phi_s$ are obtained and derived from the technology library file.

To include a statistical description of *I-V* model we evaluate the drain current $I_{ds}$ variation as:

$$\delta_{i_{ds}}(t,\xi) = \sum_{k=1}^{m} \frac{\partial I_{ds}}{\partial p_k}\Big|_{p_k = p_{k0}}(t) \cdot \xi_k = \sum_{k=1}^{m} \chi_k(t) \cdot \xi_k \tag{6}$$

$$p_k = p_{k0} + \xi_k \quad (k = 1 \sim m) \tag{7}$$

where $p_k$ is the $k_{th}$ random process parameter which is the sum of nominal value $p_{k0}$ and random variable $\xi_k$ with zero mean value and same standard deviation as $p_k$. $\chi_k(t)$ is the differential function of $I_{ds}(t)$ by elements of the corresponding process variation.

## 2.2 C-V model

The most accurate way to model non-linear capacitances is to represent them as voltage dependent terminal charge sources [15]. Characterization of such a model would involve generating charge tables for a range of terminal voltages. All capacitances are derived from the charge to ensure charge conservation. Each capacitance is computed by $C_{ij}=\partial Q_i/\partial V_j$ at every time step, where *i* and *j* denote the transistor terminals. Although this approach might be the most accurate, performance would be a problem for SSTA. Furthermore, characterization also becomes a runtime intensive task. In the 45nm node and beyond, parasitic capacitances become more non-linear. As an example, $C_{gd}$ is shown in Fig 3. In order to improve accuracy while maintaining good computational efficiency, SSTM treats the five capacitances differently. In each of the operating regions, gate channel capacitances (GCC) $C_{gs}$, $C_{gd}$ and $C_{gb}$ are modeled as constant values. For junction depletion capacitances $C_{sb}$ and $C_{db}$, SSTM uses constant values through all regions of operations since they are one to two orders of magnitude smaller than GCCs. In the statistical *C-V* model (8), $C_{j0}(t)$ is the nominal value of the $j^{th}$ capacitance in Fig 2b and the sensitivities $\zeta(t)$ are characterized by perturbing method:

$$C_j(t,\xi) = C_{j0}(t) + \sum_{k=1}^{m} \frac{\partial C_j}{\partial p_k}\Big|_{p_k = p_{k0}}(t) \cdot \xi_k = C_{j0}(t) + \sum_{k=1}^{m} \zeta_k(t) \cdot \xi_k \tag{8}$$

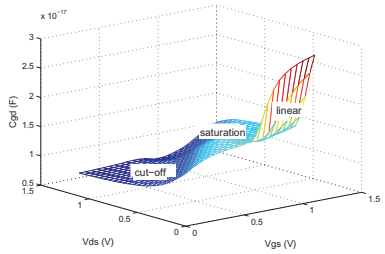


**Fig 3 The $C_{gd}$ of a minimum-size NMOS transistor**

# 3. RDE-BASED STATISTICAL SIMULATION

## 3.1 Random MNA for process variations and stochastic waveform propagation

In general, for time-domain analysis, modified nodal analysis (MNA) leads to non-linear ordinary differential equations (ODE) or differential algebraic equations (DAE) system that, in most case, is transformed to a nonlinear algebraic system by means of numerical integration methods [16]. At every integration step, a Newton-like method is then used to solve this nonlinear algebraic system. Defining $x'$ as the time derivative of the variable $x$, the MNA equations for any circuit can be expressed in compact form:

$$F(x', x, t, p_0) = 0 \quad x(t_0) = x_0 \tag{9}$$

where $x$ is the vector of the circuit state variables consisting of $n$ nodal voltages and $e$ branch currents, and $p_0$ is the nominal process parameter vector with elements $p_{k0}$ introduced in (7). Let $x_s$ be the solution to (9). Transient analysis in a conventional circuit simulator solves for $x_s$ using numerical integration methods for ODEs and DC analysis solves for the initial vector before transient analysis is started. By introducing the process variations in (9), the deterministic ODE can be expanded as:

$$F(x', x, t, p) = 0 \quad x(t_0) = x_0 + \delta_{x_0} \tag{10}$$

where $p$ is the random process parameter vector with elements $p_k$ from (7), and $\delta_{x0}$ is the initial variation caused by PVs. It is worth noticing that the main difficulty in solving (10) is related to the nonlinearity and a large set of correlated random variables. Therefore, in order to make the problem manageable, we linearize (10) with a truncated Taylor expansion and employ principal component analysis (PCA) to model a large set $m$ in (7) of correlated random variables $p$ to a $m'$-dimensional vector of uncorrelated random variables. To avoid notational cluttering, the notation $p$ representing the uncorrelated process variables after PCA is further used in this paper. The linear Taylor expansion of

$F$ around $x_s$ (Eq. 11) is used considering the possibility and complexity of solving the resulting random differential equations.

$$F(x',x,t,p) \cong F(x'_s,x_s,t,p_0) + \frac{\partial}{\partial x'}F(x',x,t,p)\Big|_{\substack{x'=x_s' \\ p=p_0 \\ x=x_s}}(x'-x'_s) + \frac{\partial}{\partial x}F(x',x,t,p)\Big|_{\substack{x'=x_s' \\ p=p_0 \\ x=x_s}}(x-x_s) + \frac{\partial}{\partial p}F(x',x,t,p)\Big|_{\substack{x'=x_s' \\ p=p_0 \\ x=x_s}}\cdot\xi \tag{11}$$

Define $y(t)=x(t)-x_s(t)$ as the variation vector of $x(t)$ due to process variation vector $\xi$ with zero $\mu$ and finite $\sigma$ mentioned in (7). Denoting the partial derivative of $F$ to $x'$, $x$ and $p$ as $C$, $A(t)$ and $B(t)$ respectively, the compact expression of (11) is obtained as

$$Cy'(t) + A(t)y(t) + B(t)\xi = 0 \quad y(t_0) = \delta_{x_0} \tag{12}$$

The nonlinear equation (10) is converted to a linear random differential equation (RDE) in $y$ with time-varying coefficient matrices $A$ and $B$. $C$ is constant since a nominal value is selected for every capacitor at every operating state. We rewrite (12) as:

$$y'(t) = E(t)y(t) + F(t)\xi \quad y(t_0) = \delta_{x_0} \tag{13}$$

Let $\Phi(t,t_0)$ be the homogeneous solution of (13) satisfying $\Phi'(t,t_0)=E(t)\Phi(t,t_0)$. According to the mean square (*m.s.*) integral theorem [13], there exists a unique *m.s.* solution which can be represented by:

$$y(t) = \Phi(t,t_0)\,y_0 + \int_{t_0}^{t}\Phi(t,u)F(u)\xi du \tag{14}$$

Even if the random variable $\xi$ is not strictly Gaussian, a second-order probabilistic characterization yields sufficient information for most practical problems:

$$E\{y(t)\} = 0 \tag{15}$$

$$Var\{y(t)\} = \Phi(t,t_0)Var\{y_0\}\Phi^T(t,t_0) + \left(\int_{t_0}^{t}\int_{t_0}^{t}\Phi(t,u)F(u)F^T(v)\Phi^T(v,t_0)dudv\right)Var\{\xi\} + \Phi(t,t_0)Cov\{y_0,\xi^T\}\left(\int_{t_0}^{t}\Phi(t,u)F(u)du\right)^T + \left(\int_{t_0}^{t}\Phi(t,u)F(u)du\right)Cov\{\xi,y_0^T\}\Phi(t,t_0)^T \tag{16}$$

Assuming that the initial condition $x_0$ is set to a fixed value, $y(t)$ and variance of $y(t)$ can be written as:

$$y(t) = \left(\int_{t_0}^{t}\Phi(t,u)F(u)du\right)\cdot\xi = \alpha(t)\cdot\xi \tag{17}$$

$$Var\{y_j(t)\} = E\{y_j{}^2(t)\} = \sum_{k=1}^{m'}\alpha_{jk}{}^2(t)Var\{\xi_k\} \tag{18}$$

where $y_j(t)$ is the $j^{th}$ element of the $y(t)$ vector. Since the variation vector $y(t)$ is proportional to $\xi$, the coefficient $\alpha(t)$ can be calculated by substituting $y(t)$ with $\alpha(t)\xi$ in (13). As a result, the solution of $y(t)$ becomes the solution of a linear ODE:

$$\alpha'(t) = E(t)\alpha(t) + F(t) \qquad \alpha(t_0) = 0 \tag{19}$$

which can be solved by fast numerical methods. The mean and variance of $x(t)=x_s(t)+y(t)$ are expressed as:

$$E\{x(t)\} = E\{x_s(t) + y(t)\} = x_s(t) \tag{20}$$

$$Var\{x_j(t)\} = E\{y_j(t)^2\} = \sum_{k=1}^{m'}\alpha_{jk}{}^2(t)Var\{\xi_k\} \tag{21}$$

where $x_j(t)$ is the $j^{th}$ element of vector $x(t)$. As long as $\alpha(t)$ is calculated, the expression for $y(t)$ is known, thus the covariance matrix of the solution $x(t)$ at two different time points $t_a$ and $t_b$ can be written as:

$$Cov(x_a,x_b) = E\left\{\left(x_a - E\{x_a\}\right)\cdot\left(x_b - E\{x_b\}\right)^T\right\} = E\{y_a\cdot y_b{}^T\} = E\{\alpha(t_a)\xi\xi^T\alpha^T(t_b)\} = \alpha(t_a)\cdot diag\left(Var\{\xi_1\},\cdots,Var\{\xi_{m'}\}\right)\cdot\alpha^T(t_b) \tag{22}$$

The $j^{th}$ diagonal entries of $Cov(x_a,x_b)$ is the covariance of the $j^{th}$ element of $x$ at the time $t_a$ and $t_b$, which is crucial for statistical slew calculation. If the covariance of different node voltages is required (e.g. statistical delay calculation), the corresponding non-diagonal entries are kept. Otherwise, the non-diagonal entries can be set to zero without calculation.

The results of the proposed analysis are elements of $x(t)$ and their variations $y(t)$ which are linear functions of $\xi$. If the input signal $x_j(t)$ is deterministic, the elements $y_j'(t)$ and $y_j(t)$ corresponding to it, and the $j^{th}$ column of $C$, $A$ and $B$ matrices in (12) are all zeros. However, if the input signals are stochastic processes, those entries will have non-zero values, and in essence provide the means to propagate stochastic signals. For STA, the output is a time-indexed voltage array $x(t)$. For SSTA, the statistical output waveform is modeled by its mean, variance and covariance values.

## 3.2 Probability density function of delay

The time-domain analysis of a non-linear dynamic circuit consists of the successive solutions of many linear resistive circuits approximating the original circuit at special operating points. Therefore, from a numerical point of view, the output signal is discrete rather than continuous thus the *pdf* of crossing time as well as delay have discrete *pdfs*. Unlike [17], where firstly the cumulative distribution function (*cdf*) of crossing time is calculated to find a discrete *pdf* at any time point, we compute the *pdf* of the crossing time for both rising $t_{r\eta}$ and falling $t_{f\eta}$ output voltage $V_o$, directly:

$$\begin{aligned}\Pr(t_{r\eta} = t) &= \Pr(V_o(t-\Delta t) \le V_\eta \cap V_o(t) \ge V_\eta) \\ &= \Pr(V_o(t-\Delta t) \le V_\eta) - \Pr(V_o(t-\Delta t) \le V_\eta \cap V_o(t) \le V_\eta)\end{aligned} \tag{23}$$

$$\begin{aligned}\Pr(t_{f\eta} = t) &= \Pr(V_o(t-\Delta t) \ge V_\eta \cap V_o(t) \le V_\eta) \\ &= \Pr(V_o(t) \le V_\eta) - \Pr(V_o(t-\Delta t) \le V_\eta \cap V_o(t) \le V_\eta)\end{aligned} \tag{24}$$

where the crossing time $t_\eta$ is the time for node voltage to cross any corresponding threshold voltage $V_\eta=\eta_{\%}\cdot V_{dd}$ and $P_r(V_o(t-\Delta t) \le V_\eta \cap V_o(t) \le V_\eta)$ is the joint *cdf* of $V_o$ at two successive time steps. Since the node voltages are assumed Gaussian distributed, the joint *cdf* is straightforwardly calculated based on (20)-(22). Similarly, given the *pdfs* of $t_{r\eta}$ and $t_{f\eta}$ in (23) and (24) respectively, the mean and variance of crossing time can be found. Note that in contrast to [10], computing a *pdf* in such a way includes the correlation of $V_o$ at two time steps.

Given mean and variance of crossing time, the mean and variance of delay and slew can be calculated. It is worth noting that the *pdf* of crossing time is calculated only during the corresponding

transition time period rather than throughout the whole transient analysis period.

## 4. EXPERIMENTAL RESULTS

The characterization time of gate-level models for SSTA is quite long (often several months) since standard-cell libraries consist of hundreds of cells with different sizes and process corners. On the other hand, by using transistor-based models, such as SSTM, the characterization time is reduced significantly as only the transistors required in standard-cells need to be characterized. For our model, data for characterization is obtained from Spectre using BSIM4 model for the 45nm PTMLP technology [14], and imported to Matlab to acquire the required parameters.

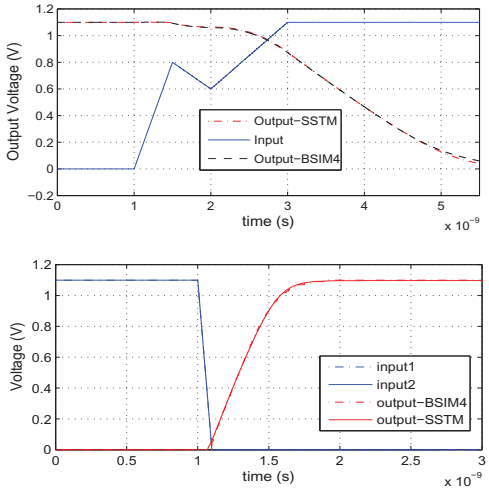


**Fig 4 Top: irregular input; Bottom: simultaneous multi-input switching for a 2-input NAND gate**

Firstly, we evaluated the nominal SSTM, when no process variation data is included in the model, in a minimum-sized inverter and 2-input NAND with different input slew and output capacitances. The input slew ranges from *1ps* to *500ps* and the load capacitance spans from *0.5fF* to *40fF*. In comparison with the Spectre BSIM4 model, the results show that the maximum relative error is within *5%*. Actually, *99.2%* of the output rise delay error and *93.9%* of the output fall delay error are within *1.6%*. The average relative errors of output rise delay and output fall delay are *0.44%* and *1.34%* respectively. The average relative error of output slew calculation using the nominal SSTM is *1.2%*, including maximum error *3.3%* without load capacitance and minimum error *0%*. Fig 4 illustrates the accuracy of the nominal SSTM used in a minimum-sized inverter with irregular input and a 2-input NAND in simultaneous multi-input switching scenario. All the results and waveforms mentioned above show a very good match between the nominal SSTM and the BSIM4 model.

In order to evaluate the worst cases of 45nm technology, we examined minimum-sized (since they are most sensitive) 2-input NAND, NOR and Buffer cells loaded with *20fF* capacitance (Fig 5-7). Standard deviations of *9nm* and *12nm* are set for the transistors length *L* and width *W* respectively. The *pdf* of $t_{50\%}$ is magnified in Fig 5-6. The statistical waveform propagation is illustrated in Fig 7 where *vout1* designates the first inverter output which has limited $\sigma$ (*s.d.vout1*) due to the two order smaller parasitic capacitances in comparison with the buffer load capacitance. The stochastic waveform *vout1* is applied to the input of the second inverter. Note that in this case, the method covers both the statistical input waveform and the random process variables. The achieved $\mu$ error within *1.0%* and $\sigma$ error within *8.0%* of $t_{50\%}$ crossing time confirm the accuracy of the proposed model and RDE-based method for these minimum-sized gates.

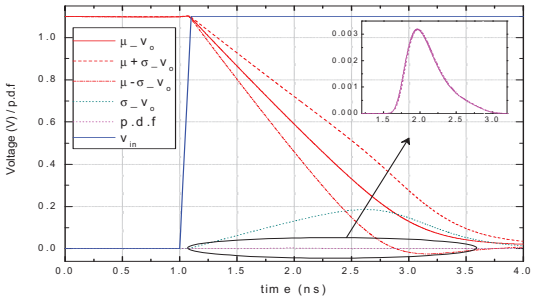


**Fig 5 Statistical analysis of a 2-input NAND**

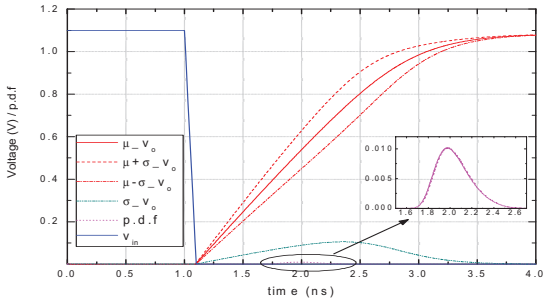


**Fig 6 Statistical analysis of a 2-input NOR**

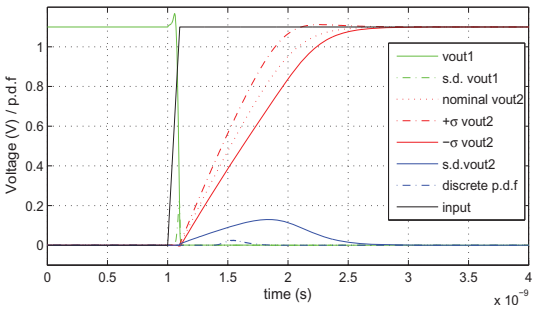


**Fig 7 Statistical analysis of a Buffer**

We also applied the SSTM and RDE-based statistical analysis method to a larger number of standard cells. The uncorrelated process variations are length and width variations with zero $\mu$. The $3\sigma$ of length and width are *20%* and *15%* of the nominal length and the largest width of every cell, respectively. The $\mu$, $\sigma$ and CPU time of rise delay are listed in Table 1. The results of fall delay are similar to the rise delay. In comparison with *1000* Monte Carlo trials in Spectre, the proposed model and analysis method achieved a relative error of within *1.4%* for $\mu$ and of within *6.8%* for $\sigma$ with an average *40*× speedup.

## 5. CONCLUSION

Statistical simulation is one of the most important steps in the evaluation of successful high-performance IC designs due to process variations that strongly affect devices behavior in today's nanometer technologies. In this paper, instead of performing statistical static timing analysis (SSTA) using thousands of Monte Carlo trials based on gate-level models, we present a transistor-level non-Monte Carlo approach, where the statistics of delay and slew are found by solving random differential equations (RDE) derived from modified nodal analysis (MNA). In order to maintain practical runtime, a statistical simplified transistor model (SSTM) is introduced without the loss of high-level accuracy. Note, however, that our non-Monte Carlo approach can be used for other transistor models and gate-level models. The logic gates built based on the SSTM are independent of input waveform and output load and are able to deal with multi-input switching. The effectiveness of the proposed SSTM and RDE-based statistical analysis was evaluated on 45nm PTMLP minimum-size gates and standard cells. As the results indicate, the suggested numerical methods provide accurate and CPU time efficient solutions.

## 6. ACKNOWLEDGMENT

We would like to thank Alessandro Di Bucchianico for valuable mathematical discussions and suggestions.

## 7. REFERENCES

[1] 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*, pp. 247-252, Jul. 2006.

[2] C.A.C. Menezes and N. Kashyap, "A true electrical cell model for timing, noise and power grid verification," In *DAC*, pp. 462-467, Jun. 2008.

[3] Private conversation with ST.

[4] C.S. Amin, F. Dartu and Y.I. Ismail, "Weibull-based analytical waveform model," *IEEE Trans. on CAD*, vol. 24, no.8, pp. 1156-1168, Aug. 2005.

[5] P. Li and E. Acar, "A waveform independent gate model for accurate timing analysis," In *ICCD*, pp. 363-365, Oct. 2005.

[6] S. Raja, F. Varadi, M.Becer and J. Geada, "Transistor level gate modeling for accurate and fast timing, noise and power analysis," In *DAC*, pp. 456-461, Jun. 2008.

[7] A. Dharchoudhury and S.M. Kang, "Fast and accurate timing simulation with regionwise quadratic models of MOS I-V characteristics," In *ICCAD*, pp. 190-194, Nov. 1994.

[8] M. Chen, W. Zhao, F. Liu and Y. Cao, "Fast statistical circuit analysis with finite-point based transistor model," In *DATE*, pp. 1-6, Apr. 2007.

[9] T. Sakural and A.R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE Journal of Solid-State Circuits*, vol. 25, no. 2, pp. 584-594, Apr. 1990.

[10] B. Liu and A.B. Kahng, "Statistical gate level simulation via voltage controlled current models," *IEEE Proc. of MBAS*, pp. 23-27, Sept. 2006.

[11] A. Goel and S. Vrudhula, "Current source based standard cell model for accurate signal integrity and timing analysis," In *DATE*, pp. 574-579, Mar. 2008.

[12] T.T. Soong, *Random differential equations in science and engineering*, New York: Academic Press, 1973.

[13] J.M. Rabaey, *Digital integrated circuits: A design perspective*, pp. 96-100, Upper Saddle River: Prentice Hall, 1996.

[14] Predictive Technology Model for Low-power Applications (PTMLP), Nov. 2008, [Online], Available: *http://www.eas.asu.edu/~ptm/modelcard/LP/45nm_LP.pm*

[15] BSIM4, [Online], Available: *http://www-device.eecs.berkeley.edu/~bsim3/bsim4.html*

[16] A. Brambilla, A. Premoli and G.S. Gajani, "Recasting modified nodal analysis to improve reliability in numerical circuit simulation," *IEEE Trans. on CAS*, vol. 52, no. 3, pp. 522-534, Mar. 2005.

[17] H. Fatemi, S. Nazarian and M. Pedram, "Statistical logic cell delay analysis using a current-based model," In *DAC*, pp. 253-256, Jul. 2006.

**Table 1 Comparison of the mean $\mu$, standard deviation $\sigma$ and CPU time of delay between Spectre and the RDE-based method**

| Standard cells | μ_Spectre (e-9) | μ_RDE (e-9) | relative erro_μ | σ_Spectre (e-11) | σ_RDE (e-11) | relative error_σ | cputime_ Spectre(s) | cputime_ RDE (s) |
|---|---|---|---|---|---|---|---|---|
| INX_X1 | 0.3777 | 0.3779 | 0.00% | 3.3407 | 3.4753 | 4.03% | 91 | 1.35 |
| INV_X2 | 0.1986 | 0.1998 | 0.60% | 1.6113 | 1.7140 | 6.37% | 91 | 1.35 |
| INV_X4 | 0.1148 | 0.1135 | 1.10% | 0.7103 | 0.7308 | 2.98% | 91 | 1.35 |
| BUF_X1 | 0.3914 | 0.3915 | 0.00% | 3.7228 | 3.5644 | 4.25% | 200 | 5.9 |
| BUF_X2 | 0.2215 | 0.2217 | 0.01% | 1.4150 | 1.5100 | 6.71% | 200 | 6.0 |
| BUF_X4 | 0.1554 | 0.1535 | 1.20% | 0.9995 | 1.0481 | 4.80% | 200 | 6.0 |
| NAND_X1 | 0.3830 | 0.3842 | 0.31% | 3.4579 | 3.5618 | 3.01% | 197 | 6.6 |
| NAND_X2 | 0.2063 | 0.2080 | 0.82% | 1.7487 | 1.7843 | 2.03% | 197 | 6.6 |
| NOR_X1 | 0.5040 | 0.4974 | 1.31% | 4.4148 | 4.2930 | 2.70% | 198 | 10 |
| NOR_X2 | 0.2635 | 0.2598 | 1.40% | 2.2557 | 2.1389 | 5.18% | 198 | 10 |