Voltage Sensitivity Calculation for ViVo-based Gate Models Considering Process Variations

Qin Tang, Amir Zjajo, Michel Berkelaar, Nick van de Meijs
Circuits and Systems, Delft University of Technology
Mekelweg 4, Delft, The Netherlands
+31(0)152783865; Q.tang@tudelft.nl

***Abstract*—The variational waveform model is employed to represent and propagate voltages in statistical delay calculation in the presence of process variations. In the existing variational waveform models, the voltage sensitivity with respect to process variations is required. In this paper, we present a first-order model based method which is used to generate the sensitivity equation from random nodal analysis (NA) or modified NA (MNA) equation of every gate. The sensitivity calculation method is general for all ViVo-based gate models. The proposed algorithm is verified with standard cells, simple digital circuits and ISCAS benchmark circuits for statistical delay calculation. The results demonstrate the high accuracy and speedup of our algorithm compared to SPICE Monte-Carlo-based simulations.**

I. Introduction

Static Timing Analysis (STA) tools are widely used for performance verification due to their ability to perform efficient timing checks on large chips. In early times, the nonlinear delay model (NLDM) was widely used for STA which models the gate delay as a nonlinear function of input slew (S_{in}) and output effective capacitance (C_{eff}). As technology downscaled into ultra-deep sub-micron region, noise and coupling considerations require advanced gate modeling for STA. Criox and Wong proposed a current source drive models (CSDM) which model every gate by a current source and one capacitor [1]. These elements are dependent on input and output voltages, thus the CSDM can handle arbitrary input waveform and output load. However, since CSDMs model gate with only one input and one output and have the assumption that one input is switching while others are static, some effects, such as internal charge sharing and multiple input simultaneous switching (MISS) [2], are not considered. These issues are addressed by transistor-level gate models which achieves higher accuracy for STA [3]–[5].

The downscaling of technology brings a significant increase in the device and interconnect manufacturing process variations. Therefore, there is a need for advanced analysis tools which can handle variability caused by imperfect manufacturing processes. In order to capture the impact of process variations on gate behavior, the corner based method can be used which performs STA at multiple corners. Clearly, the results are still deterministic not statistical. Although STA is accurate at every corner thanks to the higher-quality gate models, the corner-based method is too pessimistic since it is close to impossible for all process parameters to have extreme values. Additionally, if the number of process variations is N_p, there are 2^{N_p} process corners —far too many to analyze. Consequently, the increasing number of process variations poses a major obstacle for deterministic corner-based STA, and statistical timing analysis becomes more and more attractive.

In statistical timing analysis, the signal propagation is critical for both block-based and path-based timing verification. The accuracy of the signals propagated through gates highly depends on the accuracy of variational waveform models (VWMs). Although there are some waveform models for accurate STA, such as Weibull waveform model [6], they have not been extended to consider process variations. To realize the potential accuracy of CSDM and transistor-level gate models, a variational waveform models which can accurately represent real waveforms is required. There are two major published variational waveform models: i) VWM1. The voltage at every time point is a linear function of process variations [7], [8]; ii) VWM2. The voltage at every time point is modeled by four parameters to represent voltage scaling, voltage shifting, time scaling and time shifting [9]. Singular value decomposition [10] and principal component analysis [11] can also be utilized for variational waveform propagation. However, the efficient application for statistical timing analysis of large circuits is unclear.

VWM1 model was used for statistical delay calculation from variational waveform in [7], [8], [12]. Based on time domain integration of statistical variables, the variational output voltages are calculated considering only the variational input signals in [8]. All elements in simple CSM are modeled as a linear function w.r.t. process variations in [7]. The variational voltages are computed from a stochastic first-order expression in terms of process variations. Then the output voltage is treated as a Markovian process for delay distribution calculation [7]. In [9], the variational waveform model considers voltage scaling and shifting and time scaling and shifting. In both VWM1 and VWM2, the voltage sensitivity with respect to (w.r.t.) process variations of interest is required. For simple CSDMs, due to the one-input one-output simplified model, the equation for gate simulation is easy to solve. However, for the optimized CSDMs and transistor-level gate models, since more nodes are considered for MISS and internal charge effects, the gate equation becomes more complex. How to efficiently calculate the voltage sensitivity for complex gate models is challenging.

In this paper, we present a first-order model based method

which is used to generate the sensitivity equation from random nodal analysis (NA) or modified NA (MNA) equation of every gate. The sensitivity calculation method is general for both CSDMs and transistor-level gate models. The accurate statistical simplified transistor model (SSTM)-based gate modeling [5] is chosen for experiments. The proposed algorithm is verified with standard cells, simple digital circuits and ISCAS benchmark circuits. The results demonstrate the high accuracy of our algorithm compared to Spectre Monte-Carlo-based simulations.

II. Voltage sensitivity in Variational waveform models

Signal propagation is the corner stone of timing analysis, the accuracy of which is highly dependent on the how signal waveforms are modeled. A good variational waveform model (VWM) should easily predict the actual waveform in the presence of variations. The construction of such VWM has not been well-addressed in the literature.

In [7], [8], [12], VWM1 is proposed to consider the voltage at every time step as a linear function of process variations ξ. In the VWM1, the coefficients are voltage sensitivity w.r.t. ξ shown in Fig. 1. To include the waveform scaling and shifting (shown in Fig. 1 (a)-(d)), a compact modeling of variational waveform VWM2 was proposed in [9] and expressed as:

$$V(t) = (1+C)V_{nom}(t - A(t-t_0) - B) + DV_{dd} \qquad (1)$$

where A, B, C and D are the values of time and voltage shifting or scaling terms. In order to capture the impact of process variations, $A \sim D$ are represented as a linear function of ξ. A and B are calculated based on the voltage sensitivity w.r.t. ξ. Both of VWMs were introduced for voltage calculation on simple CSMs. However, the calculation complexity is significantly increased when applying them for multi-port CSDMs [2] and transistor-level gate models [3]–[5]. Therefore, the voltage sensitivity is the the core of these VWMs which must be efficiently calculated for more accurate and complex gate models [2]–[5].

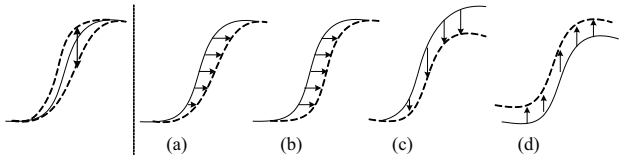

Fig. 1. The variational waveform models VWM1 and VWM2

III. Voltage sensitivity calculation

Nodal analysis (NA) or modified NA (MNA) is used for gate simulation of CSDMs and transistor-level gate models [3]–[5], [7], [8], [12], [13]. If process variations are included, the NA/MNA equation is typically a random differential equation (RDE). In this section, we propose a voltage sensitivity calculation method based on truncated Taylor expansion which needs simulation only once.

A. nominal voltage calculation

By using first-order Taylor expansion, the variational voltage $\hat{v(t)}$ can be expressed as follows:

$$\begin{aligned} \hat{v(t)} &= v_s(t) + \left(\frac{\partial \hat{v(t)}}{\partial p}|_{\hat{v(t)}=v_s(t);\xi=0} \right) \xi \\ &= v_s(t) + \mathbf{\Psi}(t)\xi \end{aligned} \qquad (2)$$

where $v_s(t)$ is the nominal voltage when there is no process variations, and p is the process parameter value with nominal value P_0 and process variation ξ. ξ has mean zero and standard deviation σ_ξ. The voltage sensitivity $\mathbf{\Psi}(t)$ is what we need to calculate for variational waveform models. It should be noted that (2) is the expression of VWM1. Since $\alpha(t)$ is the derivative around the nominal voltage $v_s(t)$, $v_s(t)$ should be calculated.

When there is no process variations, the $v_s(t)$ is calculated during typical STA procedure based on CSDMs or transistor-level gate models [3]–[5]. By using efficient threading algorithm and multiple processors, [4] 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.

B. voltage sensitivity calculation

During STA, the NA equation of every gate can be written in the compact format:

$$F(\dot{v}, v, t, p_0) = 0 \qquad v(t_0) = v_0 \qquad (3)$$

where $v(t)$ denotes the node voltage vector, $\dot{v}$ is its time derivative and p_0 represents the nominal process parameter value vector. In our simulator, the NA equation is further simplified by moving input voltages to the right hand side since they are known. Consequently, v consists of internal and output node voltages only. Since $v_s(t)$ is the nominal voltage, it is the solution of (3) which satisfies:

$$F_s = F(\dot{v_s}, v_s, t, p_0) = 0 \qquad x(t_0) = x_0 \qquad (4)$$

Since all process parameters have their nominal values p_0, $v_s(t)$ is deterministic. $v_s(t)$ is obtained by the method introduced in Section III-A. However, if process variations are considered the solution becomes statistical.

If we take into account process variations, (3) becomes a random differential equation (RDE):

$$F_x = F(\dot{\hat{v}}, \hat{v}, t, \xi) = 0 \qquad \hat{v(t_0)} = x_0 + \delta_0 \qquad (5)$$

where ξ is the process variation vector which includes both global and local variations, and δ_0 denotes the initial condition variation caused by process variations. It is worth noticing that the main difficulty to solve (5) is the high nonlinearity with respect to random variables ξ and the large number of process variations including local variations. In order to prevent exponential increase in the number of variables, the local variables can be collapsed into a single variable [4]. In

order to make (5) manageable, it is linearized by a truncated Taylor expansion around v_s and p_0.

$$\begin{aligned} F_x &\approx F_s + \frac{\partial F_s}{\partial \dot{v_s}}(t)(\dot{\hat{v}}(t) - \dot{v_s}(t)) \\ &+ \frac{\partial F_s}{\partial v_s}(t)(\hat{v(t)} - v_s(t)) + \frac{\partial F_s}{\partial p_0}(t)\xi = 0 \end{aligned} \quad (6)$$

Inserting (2) to (6) and replacing the matrices $\partial F_s/\partial \dot{v_s}$, $\partial F_s/\partial v_s$ and $\partial F_s/\partial p_0$ with $\mathbf{C(v_s)}$, $-\mathbf{E(v_s)}$ and $-\mathbf{F(v_s)}$, respectively, we obtain:

$$\mathbf{C}(v_s)\dot{\mathbf{\Psi}}(t)\xi = \mathbf{E}(v_s)\mathbf{\Psi}(t)\xi + \mathbf{F}(v_s)\xi \quad (7)$$

$$\mathbf{C}(v_s)\dot{\mathbf{\Psi}}(t) = \mathbf{E}(v_s)\mathbf{\Psi}(t) + \mathbf{F}(v_s) \quad (8)$$

The $\mathbf{C}$, $\mathbf{E}$ 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 $\mathbf{\Psi}(t)$ equation is extracted from the nonlinear random differential equation (5). The difficulty of solving (8) is from the high-dimensionality of the matrix $\mathbf{\Psi}(t)$ which is $N_v \times N_p$.

Based on moment matching, (8) is split into N_p ordinary differential equations (ODEs):

$$\mathbf{C}(v_s)\dot{\Psi}_j(t) = \mathbf{E}(v_s)\Psi_j(t) + F_j(v_s) \quad j = 1 : N_p \quad (9)$$

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 v_s-dependent coefficients $\mathbf{C}$, $\mathbf{E}$ and F_j, (9) becomes a linear algebraic equation (LAE). As a result, the voltage sensitivity equation (8) is converted into N_p typical LAEs which are solved by N_p iterations. The LAEs 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 (9) requires LU decomposition only once to solve these N_p ODEs.

C. statistical delay calculation

After computing the nominal voltage $v_s(t)$ and voltage sensitivity $\mathbf{\Psi}(t)$, the variational voltage can be obtained based on (2). According to the relationship in (2), the mean of voltage and the standard deviation and covariance of every two voltages can be calculated from the following:

$$E\{v(t)\} = v_s(t) \quad (10)$$

$$Var\{v(t)\} = \mathbf{\Psi}(t)E\{\xi\xi^T\}\mathbf{\Psi}^T(t) \quad (11)$$

$$Cov\{v(t_a), v(t_b)\} = \mathbf{\Psi}(t_a)E\{\xi\xi^T\}\mathbf{\Psi}^T(t_b) \quad (12)$$

For timing analysis, the problem of interest is to compute the moments of arrival time, gate delay or in general the crossing time. The crossing time t_η is defined as the first time for voltages to cross the threshold voltage $V_\eta = \eta\% \cdot V_{dd}$. By using a numerical integral method, e.g, backward Euler or the trapezoidal rule, the solution of v_s and $\mathbf{\Psi}$ at a specific time point are calculated from that at the previous time point, making the output $v(t)$ a Markovian process [7], [14]. During the period when the nominal voltage is in transition, the calculation of crossing time *cdf* (F_n in (13)) starts and for a rising transition this is expressed as:

$$F_n = P(t_\eta \le t_n) = 1 - P(t_\eta > t_n) = 1 - G_n \quad (13)$$

$$G_n = P(v_1 \le V_\eta \cap v_2 \le V_\eta \cap \ldots \cap v_n \le V_\eta) \quad (14)$$

$$= P(v_n \le V_\eta | v_{n-1} \le V_\eta, \ldots, v_1 \le V_\eta) \cdot G_{n-1} \quad (15)$$

$$= P(v_n \le V_\eta | v_{n-1} \le V_\eta) \cdot G_{n-1}(n = 2 : N) \quad (16)$$

$$= \frac{P(v_n \le V_\eta \cap v_{n-1} \le V_\eta)}{P(v_{n-1} \le V_\eta))} \cdot G_{n-1} \quad (17)$$

where v_i is the voltage of interest at time t_i. According to the properties of a Markovian process $v(t_n)$, (15) is rewritten in (16). Based on (13) to (17) 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, the joint probability and single probability in (17) are easy to obtain. From the *cdf*, the discretized *pdf* of crossing time is calculated [14].

IV. Experimental Results

The proposed model and algorithm were implemented and verified on a set of gates and circuits in the nangate 45nm technology [15]. Transistor-level 10000 Monte Carlo(MC) simulation results are regarded as the golden reference. VWM1 is used for variational waveform model. The gate delay distribution is chosen for the metrics to verify the voltage sensitivity accuracy.

In this paper, the voltage sensitivities were calculated based on a accurate table-based statistical transistor model (STM) [5]. In STM, every transistor is modeled as a current source I_{ds} and five capacitors (C_{gs}, C_{gb}, C_{gd}, C_{sb} and C_{db}). All elements in the STM are represented as a linear function of process variations of interest.

Since the MISS is a problem in statistical delay calculation, we firstly apply the proposed method to some multiple-input cells with MISS scenarios. The multi-input cells are $NAND2$, $NOR2$, $NOR3$, $NAND3$, $AOI21$, $AOI211$, $AOI22$ and $NAND4$. Fig. 2 - Fig. 4 shows the discrete *pdf* with 50 samples and the histogram of MC simulation in Spectre of a NAND2 with falling output, AOI21 and AOI22 with rising output. All inputs of each gate have the exact same mean value of arrival times. The discrete *pdf* was scaled to provide a straightforward shape comparison. The statistical input signal at every input of a multi-input gate was modeled as ramp signals of 100ps transition time with voltage variations. The σ of voltages and arrival time differences among input signals are varied to obtain results at diverse scenarios for every gate. The μ errors are within 1%. All the σ errors are within 5% except two biggest σ cases (6.02% and 6.42%) coming from NAND4 with rising output and falling output respectively. All of the skewness errors are within 7% [14].

We secondly used the proposed method for some cells and circuits listed in Table I, where the absolute μ and σ errors of delay distribution calculation are included. Effective length L_{eff} and threshold voltage V_{th} are chosen as the

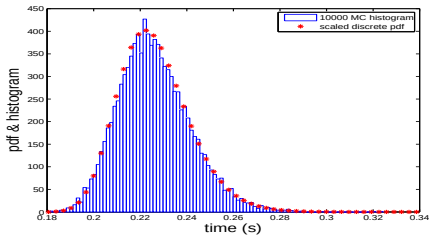

Fig. 2. The *pdf* and histogram comparison of NAND2

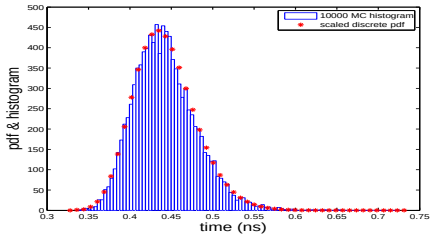

Fig. 3. The *pdf* and histogram comparison of AOI211

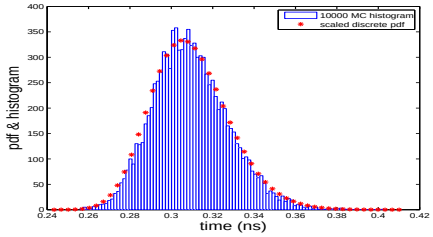

Fig. 4. The *pdf* and histogram comparison of AOI22

representative process variables, which both have 3σ equal to 20% of the mean value. We applied the proposed method to nine common standard cells with different input transitions. Fig. 5 illustrates the average relative errors (absolute values) of μ and σ for nine common standard cells. The worst σ errors are -4.03% and 3.04% from AOI211 and XOR2 with falling output respectively. 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 comparison. The results show high accuracy of our method.

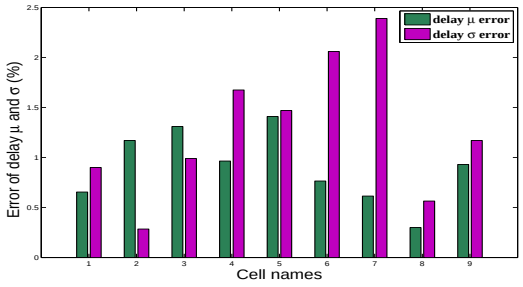

Fig. 5. First-order voltage sensitivity calculation method for delay distribution computation of standard cells. The numbers of 1 to 9 stand for INV, NAND2, NOR2, BUF, AND2, XOR2, AOI211, NAND4 and MUX2 respectively [16].

TABLE I
THE ABSOLUTE VALUES OF DELAY μ AND σ ERRORS (UNIT: %) OF SOME CIRCUITS COMPARED WITH 10K SPECTRE MC RESULTS [16].

name/value	C17	Adder	C432	C499
μ error	0.50%	0.01%	0.18%	0.81%
σ error	0.35%	0.05%	2.00%	2.19%

V. CONCLUSION

The voltage sensitivity is critical for the accuracy of many variational waveform models, which models the impact of manufacturing process variations and environment variations. In this paper, we proposed a efficient method for voltage sensitivity calculation for multiple-port CSDMs and transistor-level gate models. In the proposed method, the sensitivity equation is extracted from nodal analysis equations based on the first order Taylor expansion. The equation is converted to a linear algebraic equation after using numerical integration methods, which is solved fast since the LU decomposition needs to perform only once and no root-finding iterations are required. The experiment results show high accuracy of the proposed voltage sensitivity calculation method for statistical delay calculation.

VI. ACKNOWLEDGMENT

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

REFERENCES

[1] J. F. Croix and D. F. Wong, "Blade and razor: cell and interconnet delay analysis using current-based models," in *DAC*, 2003, pp. 386–389.
[2] 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 *In DAC 2006*, June 2006, pp. 247–252.
[3] P. Kulshreshtha, R. Palermo, M. Mortazavi, C. Bamji, and H. Yalcin, "Transistor-level timing analysis using embedded simulation," in *ICCAD*, 2000, pp. 344–349.
[4] 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.
[5] Q. Tang, A. Zjajo, M. Berkelaar, and N. van der Meijs, "Transistor level waveform evaluation for timing analysis," in *VARI*, 2010, pp. 1–6.
[6] C. Amin, F. Dartu, and Y. Ismail, "Weibull based analytical waveform model," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 24, pp. 1156–1168, Aug. 2005.
[7] H. Fatemi, S. Nazarian, and M. Pedram, "Statistical logic cell delay analysis using a current-based model," in *DAC*, 2006, pp. 253–256.
[8] B. Liu and A. B. Kahng, "Statistical gate level simulation via voltage controlled current source models," in *BMAS Workshop*, Sept. 2006, p. 23.
[9] V. Zolotov, J. Xiong, S. Abbaspour, D. J. Hathaway, and C. Visweswariah, "Compact modeling of variational waveforms," in *ICCAD*, 2007, pp. 705–712.
[10] A. Ramalingam, A. Kumar, S. R. Nassif, M. Orshansky, and D. Z. Pan, "Accurate waveform modeling using singular value decomposition with applications to timing analysis," in *DAC*, 2007, pp. 148–153.
[11] S.-A. Aftabjahani and L. Milor, "Compact variation-aware standard cell models for static timing analysis," in *ISQED*, 2008, pp. 148–151.
[12] A. Goel and S. Vrudhula, "Statistical waveform and current source based standard cell models for accurate timing analysis," in *DAC*, 2008, pp. 227–230.
[13] B. Liu, "Gate level statistical simulation based on parameterized models for process and signal variations," in *ISQED*, 2007, pp. 257–262.
[14] Q. Tang, A. Zjajo, M. Berkelaar, and N. van der Meijs, "Statistical delay calculation with multiple input simultaneous switching," in *ICICDT*, 2011, accepted.
[15] Nangate 45nm open cell library. [Online]. Available: http://www.nangate.com/index.php?option=com_content&task=view&id=137&Itemid=137
[16] Q. Tang, A. Zjajo, M. Berkelaar, and N. van de Meijs, submitted.