A Gate Delay Model Focusing on Current Fluctuation over Wide-Range of Process and Environmental Variability

Ken'ichi Shinkai†, Masanori Hashimoto†, Atsushi Kurokawa‡, and Takao Onoye†

† Dept. Information Systems Engineering, Osaka University, Osaka, Japan
{shinkai.kenichi, hasimoto, onoye}@ist.osaka-u.ac.jp

‡ Semiconductor Technology Academic Research Center (STARC), Kanagawa, Japan

ABSTRACT

This paper proposes a gate delay model that is suitable for timing analysis considering wide-range process and environmental variability. The proposed model focuses on current variation and its impact on delay is considered by replacing output load. The proposed model is applicable for large variability with current model constructed by DC analysis whose cost is small. The proposed model can also be used both in statistical static timing analysis and in conventional corner-based static timing analysis. Experimental results in a 90nm technology show that the gate delays of inverter, NAND and NOR are accurately estimated under gate length, threshold voltage, supply voltage and temperature fluctuation. We also verify that the proposed model can cope with slow input transition and RC output load. We demonstrate applicability to multiple-stage path delay and flip-flop delay, and show an application of sensitivity calculation for statistical timing analysis.

Categories and Subject Descriptors

B.7.2 [**Integrated Circuits**]: Design Aids—*Simulation*

General Terms

Design, Experimentation, Performance

Keywords

gate delay model, variability, static timing analysis, statistical timing analysis

1. INTRODUCTION

Process variation is predicted to get severer [1], which prompt us to develop a timing analysis method that can statistically estimate circuit delay before fabrication. Currently, statistical static timing analysis (SSTA) is intensively studied [2–4]. Because of tightly allotted design time and expensive mask cost, pre-fabrication timing verification considering both manufacturing and environmental variability will become a key technology for successful chip design. On the other hand, recent chip design with lowered supply voltage yet increasing power dissipation enforces us to endure delay variation due to supply voltage fluctuation [5, 6]. As the number of low power application increases, dynamic techniques for reducing switching and leakage power dissipation. For example, DVS (dynamic voltage scaling) and variable V_{th} design are getting incorporated into a SoC design flow. In these designs, intentional V_{dd} and V_{th} variations are generally larger than usual environmental variability, and hence an obstacle is timing verification in numerous operation modes and corners. It means that conventional gate delay modeling tailored for each mode and corner goes bankrupted in characterization cost.

Thus a gate delay model that can handle manufacturing and environmental variability is eagerly demanded. A naive approach is to increase the number of parameters in table look-up model, but it is difficult to use, because its characterization cost increases exponentially. Sensitivity-based model using Taylor expansion (Equation (1)) is a reasonable approach for variational analysis,

$$d = d_0 + \sum_{\forall p_i} \left[\frac{\partial f}{\partial p_i}\right]_0 \triangle p_i, \tag{1}$$

where d_0 is the nominal value of d, p_i is the i-th variational parameter, $\left[\frac{\partial f}{\partial p_i}\right]_0$ is the delay derivative with parameter p_i at the nominal, and $\triangle p_i$ is the difference of p_i from the nominal value. The sensitivity-based model is widely used in SSTA [2, 3, 7], because it has desirable statistical characteristics [2] and it can cope with various variational parameters in a similar way. However, delay is inversely proportional to current [8], and in nature it has a strong non-linear dependency on variation sources, which is a reason why sensitivity model is not tolerant of large variability. Another problem is that the sensitivity depends on output load and input signal waveform, and hence a number of additional transient analyses are necessary for characterization, which is computationally very expensive. Moreover, the sensitivity computed at the nominal condition in Equation (1) might be much different from the sensitivity in the actual operating condition, such as the case that the average supply voltage drops at the center of a chip.

This paper presents a gate delay modeling technique that can cope with large amount of manufacturing and environmental variability. A key point is that we focus on current variation and not on delay directly. Gate delay is the time required to charge/discharge output load. The proposed method translates and maps the current variation caused by manufacturing and environmental variability into variation in output load (Figure 1).

The advantage of the proposed model is summarized as follows: (1) Large applicable variation range of the proposed method is suit-

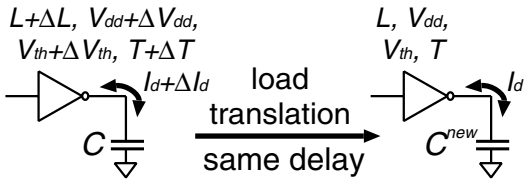

Figure 1: Proposed concept of translating variability into output load.

ICCAD '06 November 5-9, 2006, San Jose, CA

able not only for variation-aware SSTA but also for DVS and variable V_{th} design. In addition, the proposed model can be used to compute the sensitivity in the operating condition shifted from the nominal condition. (2) Additional time-consuming transient analysis for characterization is not required at all. The proposed current variation modeling needs only DC analysis, which is independent of input slope and output load, and hence the additional characterization cost is quite small. (3) We can compute gate delay in any conventional manners by using the replaced output load, and hence every timing analyzer can perform variation-aware timing analysis with the pre-process of output load translation. Delay is calculated by table look-up model, for example, as follows:

Prior preparation: (1) look-up table using input transition time and output load at nominal condition,
(2) proposed current estimation model.

Input: (1) input transition time, (2) output load,
(3) variational parameter values.

Step 1: Translate variability into current fluctuation by the proposed current estimation model,

Step 2: Replace output load based on current fluctuation,

Step 3: Calculate delay by look-up table with the replaced load.

The rest of the paper is organized as follows. Section 2 describes the key idea, that is translation of current variation into output load. Section 3 gives an implementation of the proposed idea for RC output load and slow input transitions. Section 4 shows experimental results to confirm the accuracy. We also demonstrate a comparison with the conventional sensitivity-based method. Finally, we conclude the paper in Section 5.

2. KEY IDEA OF PROPOSED MODEL

A feature of the proposed gate delay calculation is variational current modeling instead of direct delay modeling. Delay of a CMOS logic gate is the time required to charge/discharge its output load. Variability sources, such as gate length, change charging/discharging current, which results in gate delay variation. Therefore, variation-aware current estimation can guide gate delay computation for variability. We observe that charging/discharging current has an almost linear relation with respect to variation sources, and it is easier to model.

2.1 Relationship among delay, current and variability

We first review the relationship among gate delay, charging/discharging current and variability sources. Saturation current of MOSFET I_{dsat} is expressed as follows in alpha-power law MOSFET model [9].

$$I_{dsat} = k\frac{\mu\epsilon_{ox}W}{t_{ox}L}(V_{gs} - V_{th})^{\alpha}, \tag{2}$$

where μ denotes an effective mobility, ϵ_{ox} a dielectric constant of a gate oxide, t_{ox} a gate oxide thickness, W a channel width, and L a channel length. α is a coefficient to denote carrier velocity saturation effect, and it becomes close to 1 in advanced technologies. V_{th} is a threshold voltage, and k is a coefficient. Equation (2) shows that the saturation current is roughly proportional to supply voltage, threshold voltage, channel width and mobility, and inversely proportional to gate length and gate thickness. An example of current variation in a 90nm CMOS technology is shown in Figure 2, where dV_{th} is the variation of V_{th} from the nominal value. I_{step} (I_d at $V_{gs} = V_{ds} = V_{dd}$) has a linear relation with V_{dd}, V_{th} and temperature T and it is roughly inversely proportional to L. We therefore expect that current fluctuation can be modeled by a simple expression of

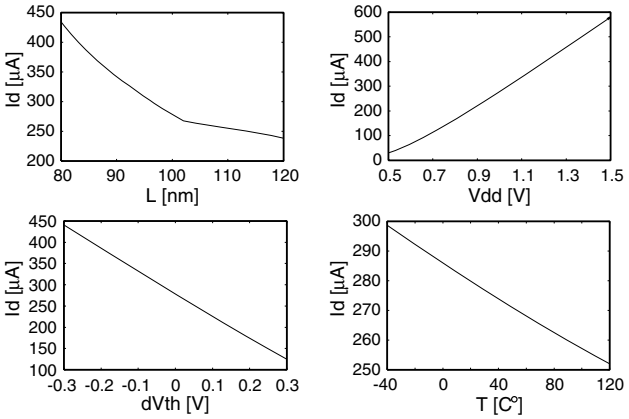

Figure 2: Example of I_{step} (I_d at $V_{gs} = V_{ds} = V_{dd}$) variation.

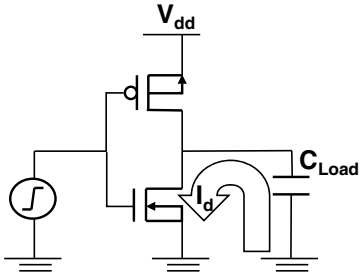

Figure 3: Drain current of an inverter.

L^{-1}, V_{dd}, dV_{th} and T. Gate delay is basically inversely proportional to current [8]. Intuitively speaking from the above discussion, the gate delay is inversely proportional to $(V_{dd} - V_{th})$. This means that the sensitivity-based delay model for variability, which is widely adopted in SSTA, is not suitable for large variability of V_{dd} and V_{th}, whereas it really works well for small amount of variation. In order to develop a gate delay model applicable to wide-range fluctuation of V_{dd} and V_{th}, which is eagerly demanded in DVS and variable V_{th} design, focusing and modeling current fluctuation is a reasonable approach.

2.2 Translation of current fluctuation into output load

We here show the key idea of the proposed model. We use a fall transition, as an example, of a 90nm single-stage inverter with 100fF output load when a step rise input transition is given (Figure 3). The fall gate delay depends on the drain current of NMOS I_d. The amount of charge to discharge, Q, is expressed as

$$Q = CV_{dd}, \tag{3}$$

where V_{dd} is supply voltage and C the total capacitance to discharge. C includes C_{Load} and parasitic capacitance inside a cell. Generally speaking, when the drain current increases twice by process variation, the delay becomes half. When the output load becomes half, the delay also decreases to the half. The above discussion means that the variation of current can be translated into output load while keeping the gate delays equal. Considering that Q varies with V_{dd} fluctuation, we thus should translate I_d variability into C^{new} as follows,

$$(I_d + \Delta I_d) : I_d = Q : Q^{new}, \tag{4}$$

$$= C(V_{dd} + \Delta V_{dd}) : C^{new}V_{dd}, \tag{5}$$

$$\Longrightarrow Delay(I_d + \Delta I_d, Q) = Delay(I_d, Q^{new}), \tag{6}$$

$$C^{new} = \frac{I_d}{I_d + \Delta I_d}\frac{V_{dd} + \Delta V_{dd}}{V_{dd}}\,C. \tag{7}$$

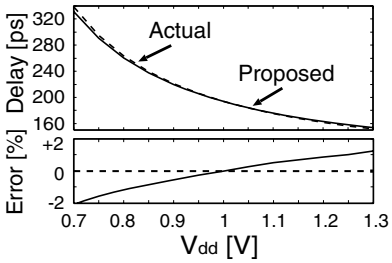

Figure 4: Estimation of delay variation under V_{dd} fluctuation (inverter, step input, fall transition, output load 100fF).

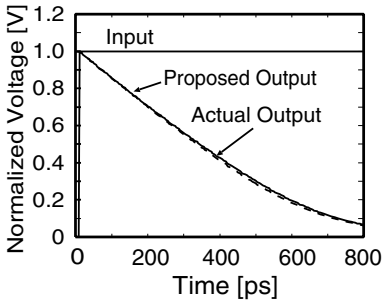

Figure 5: Waveform evaluation (inverter, fall transition, V_{dd} 0.7V(varied) 1.0V(nominal), output load 100fF).

where Δ denotes variability and superscript ^{new} corresponds to values in the nominal condition. The rise delay is calculated similarly.

In order to realize the above translation, we need two techniques: (1) current estimation and (2) estimation of parasitic output capacitance. The second technique is also required in transistor-level timing analyzer, such as [10], and it is not novel. We therefore do not discuss capacitance estimation further in this paper. With regard to (1), the saturation current I_{step} is reasonable for Equation (7) when the input signal is a step signal, because the NMOS works like a constant current source in the saturation region. In practical designs, however, input signals have a certain amount of transition time. Current modeling for slow input transition will be discussed later in Section 3.

A paper proposes a gate delay model that translates difference of power/ground voltage level between driver and receiver into output load toward power/ground-noise-aware timing analysis [6]. Though the idea that delay variation is expressed by output load replacement is the same, the variability of this paper includes both manufacturing and environmental fluctuation, and the target is different.

2.3 An analysis example

We demonstrate an example to validate the idea of the proposed method in a 90nm CMOS technology. This example supposes that power supply voltage V_{dd} fluctuates. The upper figure in Figure 4 shows delay variation estimated by circuit simulation with the actual varied V_{dd} and with the proposed idea. The lower figure shows the delay estimation error of the proposed model. In this analysis, the estimation error is below 2% even though the supply voltage is varied from 0.7 to 1.3V by 0.6V, where the nominal supply voltage is 1.0V. The proposed idea works well for wide-range variation. Figure 5 shows a waveform example, where the difference of supply voltage is removed by normalization to 1.0V for waveform comparison. Gate delay is measured after the voltage normalization as well. We can see that the output transition waveform is well estimated. We also verify the effectiveness under variations of L, V_{th} and T in a similar way.

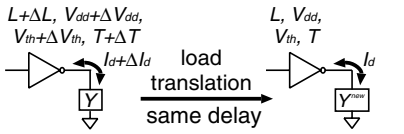

Figure 6: Concept for Generic RC load.

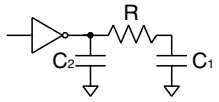

Figure 7: CRC π model.

3. COPING WITH RC OUTPUT LOAD AND SLOW INPUT TRANSITION

To implement the proposed idea in a timing analyzer, there are two obstacles in current modeling: (1) RC output load and (2) slow input transition. This section shows the solutions for each obstacle.

3.1 Translation for RC output load

Coping with RC output load is indispensable in nano-meter technology era, because wire resistance is large especially in inter-block interconnections. We here show that the proposed idea can be applied consistently to generic output load.

Figure 6 shows the concept for a generic output load, where Y is the driving point admittance of the actual circuit, and Y^{new} is that after translating variability into output load. Similar to capacitive output load, Y^{new} is calculated by

$$Y^{new} = \frac{I_d}{I_d + \Delta I_d} \frac{V_{dd} + \Delta V_{dd}}{V_{dd}} Y. \tag{8}$$

We can see that Equation (7) is one of the special cases of Equation (8). In the case of CRC π model in Figure 7, C_1^{new}, C_2^{new} and R^{new} can be simply expressed as follows,

$$C_1^{new} = \frac{I_d}{I_d + \Delta I_d} \frac{V_{dd} + \Delta V_{dd}}{V_{dd}} C_1, \tag{9}$$

$$C_2^{new} = \frac{I_d}{I_d + \Delta I_d} \frac{V_{dd} + \Delta V_{dd}}{V_{dd}} C_2, \tag{10}$$

$$R^{new} = \frac{I_d + \Delta I_d}{I_d} \frac{V_{dd}}{V_{dd} + \Delta V_{dd}} R. \tag{11}$$

3.2 Current estimation for slow input transition

We here discuss current estimation used to translate variability into output load in Equations (7) and (8). When a fast input transition is given, i.e. the output transition is much slower than the input transition, I_{step} is used as I_d, as explained in Section 2.2. This section supposes a situation that the input transition is slower than the output transition in Figure 3. The operating condition of the NMOS depends on the input waveform and the output load, and hence it is difficult to identify the discharging current that should be used for the proposed translation.To solve this problem, we have observed discharging current in various conditions of input waveform and output load, and found that the discharging current becomes almost maximum when the output goes across 50% of the supply voltage. Also obviously, the discharging current starts to flow when the input voltage reaches to the threshold voltage of the NMOS.

Figure 8 shows a typical current waveform when a slow input transition is injected. The current shape can be approximated as a trapezoid as shown in Figure 8 and the integrated area of the current is equal to the area of a rectangle whose height is $I_{avg} = \frac{1}{2}(I_{V_{in}=V_{th}} + I_{V_{out}=0.5V_{dd}})$, where $I_{V_{in}=V_{th}}$ represents the drain current when the input voltage equals the threshold voltage, and $I_{V_{out}=0.5V_{dd}}$ is the drain current when the output voltage equals half of the supply voltage. Please recall that gate delay is the time required to charge/discharge the output load. The output voltage depends on the amount of charge that has been charged/discharged through the MOS. In other words, the integral average of the charging/discharging current determines the delay characteristics. This is why I_{avg} is reasonable for slow input transition. We therefore decide to use I_{avg} as I_d. Here, you may

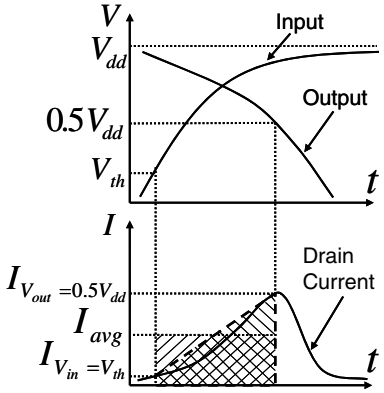

Figure 8: Current computation for slow input transition.

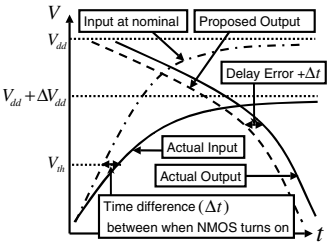

Figure 9: Variability shifts a timing at which NMOS turns on and output starts to change.

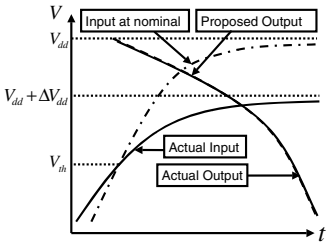

Figure 10: Compensation of timing offset.

think that $I_{V_{in}=V_{th}}$ is not necessary because of its small value, but we experimentally confirm that $I_{V_{in}=V_{th}}$ contributes to accurate delay estimation, especially in the case that V_{dd} is close to V_{th}.

A requirement for current estimation is flexibility and robustness that can accept new variation parameters and multiple input cells. We, in this paper, use a response surface method as a candidate. However, other methods also can be used as long as their accuracy is sufficient. In the response surface method, we can choose the order of polynomials according to the required accuracy. We derive polynomial expressions for $I_{V_{in}=V_{th}}(L^{-1}, V_{dd}, dV_{th}, T, V_{th})$ and $I_{V_{out}=0.5V_{dd}}(L^{-1}, V_{dd}, dV_{th}, T, V_{in})$. V_{th} is dependent on L, V_{dd} and T as well as dV_{th}, and hence V_{th} in $I_{V_{in}=V_{th}}$ is separately estimated beforehand by the response surface method.

When deriving $I_{V_{in}=V_{th}}$, the output voltage is reasonably assumed to be V_{dd}, because at that timing the output is about to start a transition. A remaining problem is how to compute V_{in} when the output is $0.5V_{dd}$. Note that the output waveform shape and/or the crossing timing of $0.5V_{dd}$ can be computed by the delay model in the nominal condition with the replaced output load. The input waveform is also given, and hence we can estimate V_{in} when the output is $0.5V_{dd}$ without transient simulation. A chicken-and-egg problem of current and V_{in} estimation exists, that is current estimation needs V_{in}, and V_{in} estimation requires the output load computed based on current estimation. We experimentally confirm that V_{in} can be estimated by iteratively updating the output load with the proposed translation, which will be shown in Section 4.2.1.

3.3 Problem of starting time of switching

The proposed method is expected to be applicable for large supply voltage and threshold voltage variation. This situation leads to a problem that the time at which the output starts to change varies (Figure 9). "Proposed Output" is normalized to the actual V_{dd}. In this example, we stretch the waveform with respect to voltage from "Actual Input" to "Input at nominal", since the supply voltage drops.

Table 1: Condition of the variability for Monte Carlo analysis.

Variable parameter	Mean	3σ
L	100nm	20nm
V_{dd}	$1.0V$	$0.5V$
dV_{th}	$0V$	$0.3V$
T	$37.5°C$	$82.5°C$

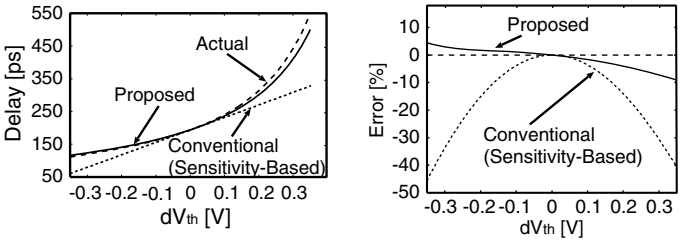

Figure 11: Comparison in applicable variation range.

Table 2: Applicable range within 10% delay error.

Variable parameter	Proposed	Sensitivity-Based
V_{dd}	±0.5V	±0.2V
V_{th}	±0.35V	±0.16V

In this case, the timing when the input reaches to the threshold voltage gets earlier by Δt due to voltage drop, which causes delay increase that is irrelevant to current variation. A similar problem also occurs when V_{th} varies. Thus, the crossing timing offset ($=\Delta t$) of threshold voltage is separately computed and compensated in delay estimation (Figure 10). The input transition waveform is given in STA, and hence Δt can be calculated without any special information nor simulation.

4. EXPERIMENTAL RESULTS

This section demonstrates the accuracy of the proposed gate delay model and its efficiency in applicable variation range. We assume that L, V_{dd}, dV_{th} and T vary as Table 1, where dV_{th} denotes a V_{th} offset excluding the influence of the other parameters.

V_{dd} and V_{th} are widely varied, because the proposed model also aims to cover DVS and variable V_{th} design. Thus, there are combinations in which MOSFETs works only in sub-threshold region. The sub-threshold current changes in an exponential manner with respect to V_{dd} and V_{th}, and hence the regression to polynomials is not appropriate. However, generally speaking, in those combinations, gate delay is unacceptably large. We therefore eliminate those combination of V_{dd} and V_{th} in experiments.

In the 90nm technology we use for experiments, the value of 2fF corresponds to an input capacitance of a standard inverter. The results of fall delay in an inverter are shown, and L, V_{dd}, dV_{th} and T vary simultaneously and independently as long as there are no special notations in the following experiments. In the following figures of waveform example, solid lines denote actual waveforms, and dotted lines correspond to the proposed model.

4.1 Comparison with conventional methods

In this section, we demonstrate advantages of the proposed model compared with conventional methods.

4.1.1 Applicable range

We demonstrate the applicable variation range of the proposed method comparing with the sensitivity-based model. Figure 11 shows a comparison in delay estimation error when dV_{th} is varied. The proposed method provides accurate estimation even though dV_{th} varies by ±0.35V. When delay estimation within 10% error is demanded, V_{dd} variation of ±0.5V and dV_{th} variation of ±0.35V are

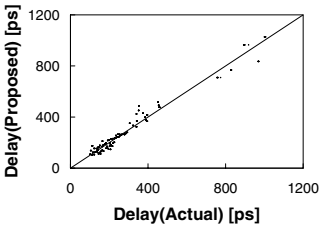

Figure 12: Delay estimation accuracy w/ direct delay model.

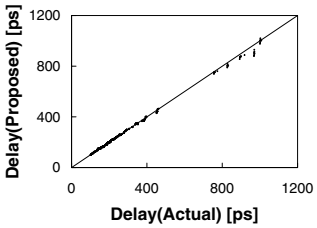

Figure 13: Delay estimation accuracy w/ proposed model.

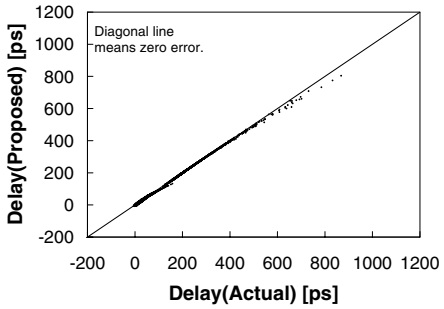

Figure 14: Delay estimation accuracy (input transition time 100ps, output load 2fF, 10fF, 100fF).

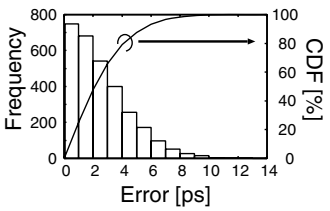

Figure 15: Error histogram (input transition time 100ps, output load 2fF).

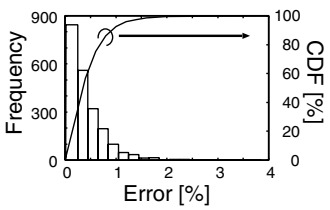

Figure 16: Error histogram (input transition time 100ps, output load 100fF).

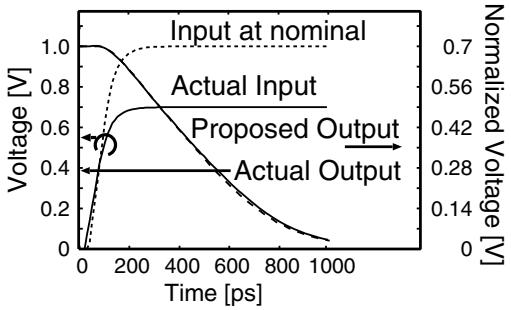

Figure 17: Estimated transition waveform (V_{dd} 0.7V, input transition time 100ps, output load 100fF).

acceptable in the proposed method (Table 2), which reveals that the proposed model is suitable for DVS and variable V_{th} design.

4.1.2 Comparison with direct delay modeling

We compare the proposed model with a model which deals with delay directly. The direct delay model is derived by the response surface method. The delay is expressed by a second-order polynomial of L, V_{dd}^{-1}, $(V_{dd} - dV_{th})^{-1}$ and T. Input transition time and output load are fixed to 100ps and 100fF, respectively. Similarly, for the proposed method, a second-order polynomial of L^{-1}, V_{dd}, dV_{th} and T is derived as a current model. The evaluation configuration in variability is the same for both models. Figures 12 and 13 show delay estimation accuracy. RMS error of the direct delay model is 39.4ps (14.0%), whereas that of the proposed model is 11.3ps (4.0%). As discussed in Section 2.1, modeling current variability is more reasonable than direct delay modeling.

Moreover, the proposed model is superior to the direct delay model in characterization cost. The current model is constructed based on DC analysis and is independent of output load and input transition time. On the other hand, delay must be evaluated at $m \times n$ sampling points, where m and n are the numbers of sampling points in output load and input transition time, respectively. The number of variability configuration required for current model derivation is $m \times n$ times as small as that for the direct delay model. If the direct model uses a higher order polynomial to make up for the larger estimation error in Figure 12, more evaluation points are necessary in general for delay model construction, and the superiority of the proposed model in characterization cost becomes significant.

4.2 Accuracy evaluation for capacitive load

4.2.1 V_{in}, I_{avg} computation

To compute $I_{V_{out}=0.5V_{dd}}$, we have to know V_{in} when $V_{out} = 0.5V_{dd}$, as explained in subsection 3.2. V_{in} depends on variability as well as output load and input transition time. We therefore compute V_{in} and I_{avg} iteratively as follows:

Step 1: Set V_{in} when $V_{out} = 0.5V_{dd}$ at nominal,
Step 2: Compute I_{avg} with V_{in} as $(V_{dd} + \Delta V_{dd}) \cdot V_{in}$,
Step 3: Replace output load by using I_{avg} in **Step 2**,
Step 4: Update V_{in} with output load in **Step 3**, and return to **Step 2**.

In **Step 1-4**, additional transient simulation is not necessary. I_{avg} computed in the above procedure is compared with that calculated by using the exact value of V_{in} obtained from transient analysis. In the results of Monte Carlo analysis, RMS error of I_{avg} estimation is 7.6% for output load of 2fF, 4.0% for 10fF, and 0.2% for 100fF. Increasing the number of the iteration from **Step 2** to **Step 4** hardly improves the accuracy of I_{avg} estimation. We conclude that V_{in} can be computed in the above process with a reasonable accuracy.

4.2.2 Accuracy of delay estimation

We next evaluate the accuracy of delay estimation. We perform Monte Carlo analysis whose evaluation count is 3000. Figure 14 shows the accuracy of delay estimation when the input transition time is 100ps. The transition time of 100ps corresponds to the transition with fan-out 8. We can see that gate delay is well estimated even for slow input transition. RMS error is 2.9ps (14.0%) for output load of 2fF, 3.3ps (7.0%) for 10fF, and 3.7ps (1.5%) for 100fF. Figures 15 and 16 show the error histograms and CDFs (cumulative distribution functions) of the proposed method. When the input transition time is 300ps, RMS error is 11.2ps for 2fF and 8.9ps (13.2%) for 10fF and 11.3ps (3.8%) for 100fF. RMS error is 11.7ps (3.4%) for a conbination of 500ps and 100fF. In the situation that the input transition is slow and the output load is small, it is difficult to estimate delay in general, however, the proposed method estimates the delay with an acceptable accuracy.

Figure 17 shows an example of transition waveform in the case of V_{dd} fluctuation. The waveform shape as well as the propagation delay is well estimated. In timing analysis, accurate estimation of output transition time is also required. RMS error is 1.7ps (5.3%) for output load of 2fF, and 6.1ps (1.4%) for 100fF when the input transition time is 100ps. The proposed model gives accurate output transition time as well as delay time.

4.2.3 Gate delay distribution

We here demonstrate that the proposed gate delay model can reproduce a statistical distribution of gate delay. We perform Monte Carlo analysis and compare the actual distribution given by circuit simulation with that of the proposed method. Figure 18 shows the

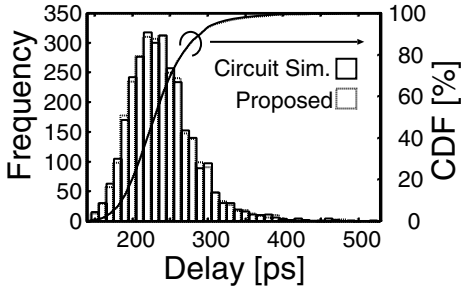

Figure 18: Delay distribution (input transition time 100ps, output load 100fF).

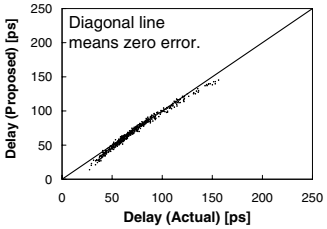

Figure 19: Delay estimation accuracy (NAND: A: LH (input transition time 100ps), B: High, output load 10fF).

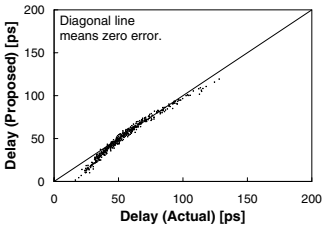

Figure 20: Delay estimation accuracy (NOR: A: LH (input transition time 100ps), B: Low, output load 10fF).

delay histograms and CDFs. The distribution shapes by the proposed method agree with those by circuit simulation. The proposed method can be used for deriving gate delay distribution.

4.2.4 NAND, NOR

The proposed method is applicable for NAND and NOR as well as inverter. We perform a Monte Carlo analysis whose evaluation count is 500. L and dV_{th} are varied independently for every transistor. Figure 19 shows the delay estimation accuracy of a 2-input NAND gate. RMS error is 3.9ps (5.5%). In the case of 2-input NOR gate(Figure 20), RMS error is 4.9ps (9.1%). The proposed model can be applied for any single-stage CMOS gates.

4.3 Accuracy evaluation for RC output load

We next evaluate the accuracy when RC output load is given. We perform a Monte Carlo analysis whose evaluation count is 3,000. A 2.0mm-long wire is used as an RC output load. We translate the wire into a CRC π model by [11]. Figure 21 shows the accuracy of the proposed method and a waveform example is Figure 22. RMS error is 11.6ps (1.2%). We can see that the proposed method works well for RC output load as well as capacitive load.

4.4 Application to multiple-stage path and flip-flop

4.4.1 Multiple-stage path

We also evaluate the accuracy when multiple-stage path is given (Figure 23). We assume that L and dV_{th} vary independently at each transistor. The capacitance of 100fF is attached to each gate output. Figure 24 shows a waveform example. We can see the proposed method works well for multiple-stage path.

4.4.2 Flip-flop

CLK-to-Q delay of a negative-edge trigger D flip-flop in Figure 25 is examined. The signal path to consider in delay estimation is extracted and shown in Figure 26. A transmission gate is modeled such that the sum of NMOS and PMOS current is expressed in

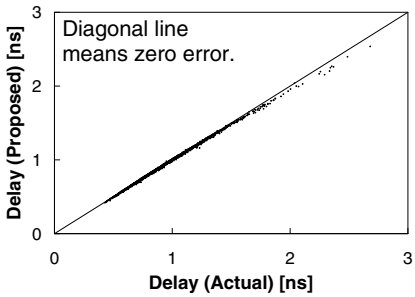

Figure 21: Delay estimation accuracy for RC output load (2.0mm-long wire).

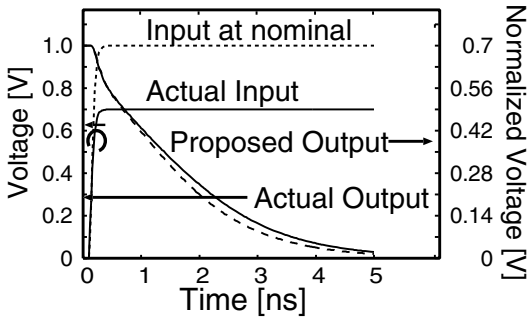

Figure 22: Waveform evaluation for RC output load (L 80nm, V_{dd} 0.7V, dV_{th} 0.1V, T 70°C, 2.0mm-long wire).

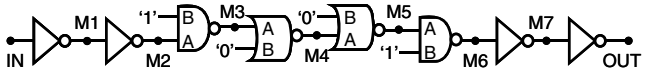

Figure 23: Example of multiple-stage path.

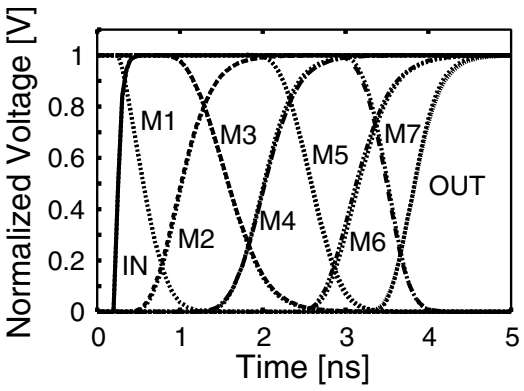

Figure 24: Waveform evaluation for multiple-stage path (solid lines denote actual waveforms, and dotted lines correspond to proposed model).

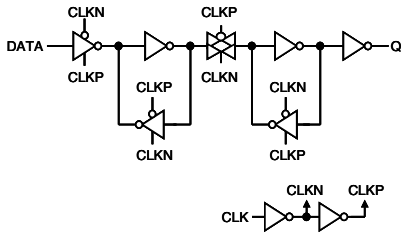

Figure 25: Negative-edge trigger DFF

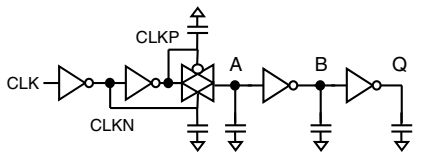

Figure 26: Signal path to consider in delay estimation.

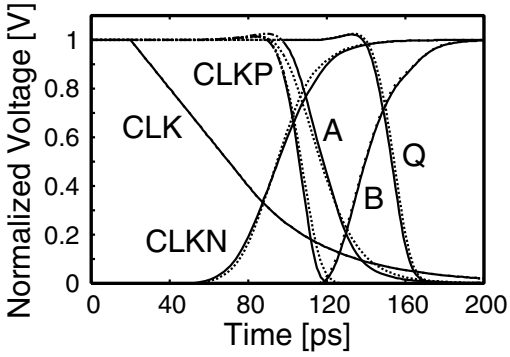

Figure 27: Waveform evaluation (DFF: clock transition time 100ps, output load 2fF, solid lines denote actual waveforms, and dotted lines correspond to proposed model).

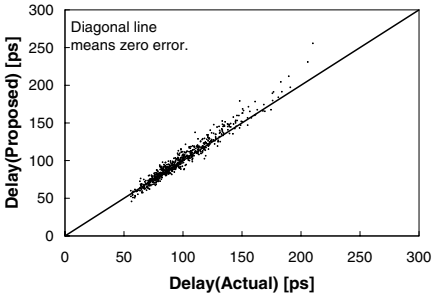

Figure 28: Delay estimation accuracy (DFF: clock transition time 100ps, output load 10fF).

a second-order polynomial assuming that step inputs of CLKN and CKLP are given at the same time for simplicity.

We vary L and dV_{th} independently at each transistor, and the same values of V_{dd} and T are given to all transistors. The transition time of given CLK signal is 100ps in this experiment. Figure 27 shows a waveform example. Labels CLK, CLKN, CLKP, A, B and Q correspond to the nodes in Figure 26. The waveforms by the proposed method are very close to the actual waveforms.

We next show the statistics of the estimation accuracy. We perform Monte Carlo analysis whose evaluation count is 1000. Figure 28 shows the accuracy of delay estimation when the output load is 10fF. RMS error is 6.7ps (6.9%). We can see that CLK-to-Q delay is well estimated.

4.5 Application to sensitivity calculation

When the average of a variational parameter shifts, the sensitivity becomes different with the sensitivity at the nominal condition. The proposed model can give the sensitivity in Equation (1) easily even when the average changes. Let us show an example. In this experiment, we vary V_{dd} such that the average and 3σ are 0.8V and 0.1V, respectively. The sensitivity to V_{dd} at 1.0V is not accurate when nominal V_{dd} is actually 0.8V. On the other hand, when we compute the sensitivity at 0.8V by using the proposed model, the estimation accuracy is much improved. Figure 29 shows the delay estimation accuracy. The proposed model can be used to compute the sensitivity at any point, which improves accuracy of SSTA that exploits the benefit of the sensitivity-based delay model (e.g. [2]).

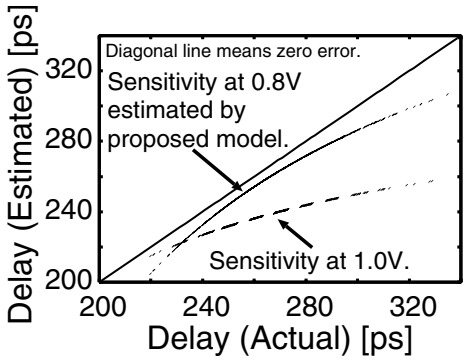

Figure 29: Comparison of delay estimation accuracy with the sensitivity at V_{dd} of 1.0V and that at 0.8V (V_{dd} variation, step input, output load 100fF).

5. CONCLUSION

We propose a method to compute variational gate delay by translating variability into output load. Every variation source is considered as a factor that fluctuates charging/discharging current. The proposed model provides accurate delay estimation in a wider range of manufacturing and environmental variability compared with sensitivity-based model, thanks to capturing variability in current modeling. Any additional characterization by time-consuming transient analysis is not necessary at all. Experimental results reveal that the RMS error of the proposed model is small even for slow input transition. The large applicable range of variability may eliminate delay characterization in numerous corners and provide a solution for timing verification in DVS and V_{th} control design.

6. ACKNOWLEDGEMENT

This work is supported in part by NEDO and STARC. We thank Dr. Masuda of STARC for a valuable discussion.

7. REFERENCES

[1] H. Masuda, S. Ohkawa, A. Kurokawa and M. Aoki, "Challenge: Variability Characterization and Modeling for 65- to 90-nm Processes," in *Proc. CICC*, pp. 593-599, 2005.

[2] H. Chang and S. Sapatnekar, "Statistical Timing Analysis under Spatial Correlations," *IEEE Trans. CAD*, Vol. 24, No. 9, pp. 1467-1482, Sep. 2005.

[3] C. Visweswariah, K. Ravindran, K. Kalafala, S. G. Walker and S. Narayan, "First-order Incremental Block-Based Statistical Timing Analysis," in *Proc. DAC*, pp. 331-336, 2004.

[4] J. Le, X. Li and L. T. Pileggi, "STAC: Statistical Timing Analysis with Correlation," in *Proc. DAC*, pp. 343-348, 2005.

[5] S. Pant and D. Blaauw, "Static Timing Analysis Considering Power Supply Variations," in *Proc. ICCAD*, pp. 365-371, 2005.

[6] M. Hashimoto, J. Yamaguchi and H. Onodera, "Timing Analysis Considering Spatial Power/Ground Level Variation," in *Proc. ICCAD*, pp. 814-820, 2004.

[7] S. Sapatnekar, "Timing," Kluwer Academic Publishers, 2004.

[8] J. M. Rabaey, A. Chandrakasan and B. Nikolic, "Digital Integrated Circuits," Pearson Education, Inc., Upper Saddle River, New Jersey, 1996.

[9] T. Sakurai 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, 1990.

[10] Synopsys Corp., "Pathmill Reference Manual," 2005.

[11] P. R. O'Brien and T. L. Savarino, "Modeling the Driving-Point Characteristic of Resistive Interconnect for Accurate Delay Estimation," in *Proc. ICCAD*, pp. 512-515, 1989.