Size Optimization Technique for Logic Circuits that Considers BTI and Process Variations

MICHITAROU YABUUCHI¹,a) KAZUTOSHI KOBAYASHI¹,b)

Received: December 5, 2015, Revised: March 8, 2016, Accepted: May 9, 2016

Abstract: In this paper we outline a transistor size optimization technique for logic circuits that takes into account BTI (Bias Temperature Instability) and process variations. We demonstrate the accuracy of our results with statistical analysis. Since variations have a large impact on the scaling process, dependable circuit designs should include a quantitative analysis if they are to become more reliable in the future. In this study we used an algorithm to prove that with our technique we efficiently lowered the timing margin of the logic path by 4.4% below the margin achieved by conventional techniques. We also observed that the lifetime of the optimized circuits extended without any area overhead.

Keywords: BTI, process variation, dependability

1. Introduction

Bias Temperature Instability (BTI) is one of the most significant aging degradations of MOSFETs fabricated in a 65 nm process or later process nodes [1]. The threshold voltage, $V_{th}$, increases with the operating time of the circuits. This degradation leads to timing violations in logic circuits and read/write failures in memories. Compensations such as adding timing margins are generally applied to circuit designs to cope with this problem.

There are two types of BTI. The first is negative BTI (NBTI) that appears in PMOSFETs, so named because $V_{th}$ of PMOSFETs increases over time when their gates are stressed by negative bias [2]. The second, positive BTI (PBTI) occurs on NMOSFETs when their gates are stressed by positive bias [3]. In CMOS inverters, either the NMOSFET or the PMOSFET is degraded during its operation. When input is high, the NMOS is degraded by PBTI and when low, PMOS is degraded by NBTI.

BTI originates from defects which trap and detraps carriers in the gate oxide [4]. Therefore, the amount of BTI-induced degradations of the MOSFETs are distributed statistically [5]. The timing margins become larger in the scaled process because the distribution of BTI variations changes drastically. BTI variation is also called “BTI variability” [6], [7].

The main sources of variability in scaled MOSFETs are 1) BTI resulting from random discrete-charges and 2) process variation. Process variation can originate from many sources including: random dopant fluctuation (RDF), line edge roughness (LER), and variations in oxide thickness [8], [9], [10], [11]. If more reliable chips are to be produced, both BTI and process variations should be considered by chip designers [12].

In this paper, we propose a transistor size optimization technique for logic circuits that takes into account both BTI and process variations by considering stress conditions and circuit lifetime. Our aim in doing this is to introduce a design technique that overcome these variations with minimum overhead. Three key features of our technique are 1) we used variation distributions to provide statistically proven results of the optimized lifetime delay, 2) we extended circuit lifetime and reduced the timing margin, 3) our technique does not require any area overhead. Conventional studies of transistor size optimization fail to consider the statistical distributions of variations [13], [14] even though this analysis is critical if these variations are to be predicted accurately. We believe this to be sub-optimal since variations have a larger impact in smaller devices. Incorporating this statistical component in our technique allowed us to produce accurate delay distributions in this study. Variations have a great impact on the scaling process and timing margins can become too large to handle in real designs. We believe it is necessary to reduce the variation effect within the constraints of a small area.

This paper is organized as follows: Section 2 provides an overview and discussion of the analytical circuit simulation methods for analyzing BTI and process variations. Section 3 outlines the proposed transistor size optimization technique considering the variations and the experimental results of the estimated circuit lifetime. Section 4 concludes this paper.


In this section, we describe two analytical models for BTI and process variations, the circuit simulation and the delay analyses method.
2.1 Analytical Model of BTI Variation

BTI-induced shifts of the threshold voltage $\Delta V_{th}$ are determined by both the characteristics and states of the defects in the gate dielectric [15]. Table 1 shows the parameters used in calculating BTI-induced $\Delta V_{th}$. When the MOSFET has $n$ defects, $\Delta V_{th}$ at time $t$ can be calculated by Eq. (1):

$$\Delta V_{th}(t) = \sum_{j=1}^{n} k_j(t) \cdot v_j$$

(1)

where, $j$ is the index of defects (1 - $n$), and $k_j$ is the state of the $j$th defect. The variable $k_j$ becomes 0 or 1 when $j$th defect captures or emits carriers, respectively.

The probability mass function of $n$ is shown in Eq. (2).

$$f_n = \frac{N^n}{n!} \cdot e^{-N}$$

(2)

Where, $N_i$ is the expected value of $n$ and explained as Eq. (3).

$$N_i = L \cdot W \cdot D$$

(3)

Where, $L$ and $W$ are the length and width of the channel of the transistor, respectively, and $D$ is the density of the defect in gate oxide. We assume $D = 4 \times 10^{-3}$/nm$^2$ [16], [17].

The probability density function of $V_{th}$ step of single defect $\nu$ is shown in Eq. (4).

$$f_{\nu} = \frac{1}{\eta} \exp(-\nu/\eta)$$

(4)

Where, $\eta$ is the expected value of $\nu$ explained as Eq. (5) [18].

$$\eta = \frac{s}{LW}$$

(5)

Where, $s$ is the coefficient of $\eta$. We assume $s = 9 \times 10^3$ mV-nm$^2$ [18].

Variables $\tau_{el}$ and $\tau_{eh}$ represent the time constants of emission and capture when the gate is biased, respectively. Variables $\tau_{el}$ and $\tau_{eh}$ represent the time constants of emission and capture when the gate is zero-biased, respectively. Time constant $\tau_{el}$ is assumed to distribute log-normally from $10^{-9}$ s to $10^8$ s [19], [20]. The relationship between $\tau_{el}$ and the other time constants are as follows:

$$\tau_{th} \approx 0.01 \tau_{el}, \ \tau_{eh} \approx 100 \tau_{el}, \ \tau_{cl} \approx 100 \tau_{el}$$

The duty factor is shown in Eq. (6).

$$DF = f \times t$$

(6)

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Parameters to calculate BTI-induced degradation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>$L$</td>
<td>Device parameter</td>
</tr>
<tr>
<td>$W$</td>
<td>Device parameter</td>
</tr>
<tr>
<td>$DF$</td>
<td>$f \times t$</td>
</tr>
<tr>
<td>$t$</td>
<td>Time</td>
</tr>
<tr>
<td>$N_i$</td>
<td>Expected value of $n$</td>
</tr>
<tr>
<td>$D$</td>
<td>$4 \times 10^{-3}$</td>
</tr>
<tr>
<td>$n$</td>
<td>Poisson dist. of $N_i$</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Exponential dist. of $\nu$</td>
</tr>
<tr>
<td>$s$</td>
<td>$9 \times 10^3$</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Expected value of $\nu$</td>
</tr>
<tr>
<td>$\tau_{el}$</td>
<td>Capture time const. of high gate voltage [s]</td>
</tr>
<tr>
<td>$\tau_{eh}$</td>
<td>Emission time const. of high gate voltage [s]</td>
</tr>
<tr>
<td>$\tau_{cl}$</td>
<td>Capture time const. of high gate voltage [s]</td>
</tr>
</tbody>
</table>

Fig. 1 Stress time distribution of BTI-induced threshold voltage degradation, $t = 10^5, 10^6, 10^7$ and $10^8$ s, $DF = 0.5, L = 45 \text{ nm}, W = 200 \text{ nm}$. Where, $f$ and $t$ are the frequency of the gate input signal and time during input is high, respectively. Long-term $P_C$ is shown in Eq. (7) [22].

$$P_C(t) = \frac{1}{\tau_e} \left[ 1 - \exp \left( -\left( \frac{1}{\tau_{eh}} + \frac{1}{\tau_{cl}} \right) t \right) \right]$$

(7)

$$\frac{1}{\tau_e} = \frac{DF}{\tau_{eh} + \tau_{cl}}$$

(8)

$$\frac{1}{\tau_e} = \frac{DF}{\tau_{eh} + \tau_{cl}}$$

(9)

Where, $\tau_{eh}$ and $\tau_{cl}$ are the effective capture time constants, and the effective emission time constants, respectively. Note that Eq. (7) is valid if the characteristics of the time constants are significantly larger than the period of the stress signal.

We show the characteristics of $\Delta V_{th}$ of the stress time, duty factor and gate width. Figure 1 shows the stress time characteristics. The x and y axes denote $\Delta V_{th}$ and the probability (CDF), respectively. Input assignments for calculations are as follows: $t = 10^5, 10^6, 10^7$ and $10^8$ s, $DF = 0.5, L = 45 \text{ nm}, W = 200 \text{ nm}$. There are 10,000 samples for each $t$.
Table 2 Average μ and standard deviation σ of BTI-induced threshold voltage distribution, \( t = 10^5 \), \( 10^6 \), \( 10^7 \) and \( 10^8 \) s, \( DFS = 0.5 \), \( L = 45 \) nm, \( W = 200 \) nm.

<table>
<thead>
<tr>
<th>( t ) [s]</th>
<th>( \mu ) [mV]</th>
<th>( \sigma ) [mV]</th>
<th>( \sigma/\mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^5</td>
<td>26.4</td>
<td>7.3</td>
<td>0.277</td>
</tr>
<tr>
<td>10^6</td>
<td>28.2</td>
<td>7.4</td>
<td>0.262</td>
</tr>
<tr>
<td>10^7</td>
<td>29.9</td>
<td>7.8</td>
<td>0.261</td>
</tr>
<tr>
<td>10^8</td>
<td>31.3</td>
<td>7.9</td>
<td>0.253</td>
</tr>
</tbody>
</table>

Fig. 2 Duty factor distribution of BTI-induced threshold voltage degradation, \( t = 10^8 \) s, \( DFS = 0.1, 0.5, 0.9 \) and \( 1.0 \), \( L = 45 \) nm, \( W = 200 \) nm.

Table 3 Average μ and standard deviation σ of BTI-induced threshold voltage distribution, \( t = 10^8 \) s, \( DFS = 0.1, 0.5, 0.9 \) and \( 1.0 \), \( L = 45 \) nm, \( W = 200 \) nm.

<table>
<thead>
<tr>
<th>( DFS )</th>
<th>( \mu ) [mV]</th>
<th>( \sigma ) [mV]</th>
<th>( \sigma/\mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>19.3</td>
<td>6.2</td>
<td>0.324</td>
</tr>
<tr>
<td>0.5</td>
<td>31.3</td>
<td>7.9</td>
<td>0.253</td>
</tr>
<tr>
<td>0.9</td>
<td>34.3</td>
<td>8.3</td>
<td>0.243</td>
</tr>
<tr>
<td>1.0</td>
<td>35.0</td>
<td>8.4</td>
<td>0.241</td>
</tr>
</tbody>
</table>

Fig. 3 Gate width distribution of BTI-induced threshold voltage degradation, \( t = 10^8 \) s, \( DFS = 0.5 \), \( L = 45 \) nm, \( W = 100, 200, 300 \) and \( 400 \) nm, respectively. Input values are assigned as follows: \( t = 10^8 \) s, \( DFS = 0.1, 0.5, 0.9 \) and \( 1.0 \), \( L = 45 \) nm, \( W = 200 \) nm. There are 10,000 samples for each \( DFS \).

Table 4 Average μ and standard deviation σ of BTI-induced threshold voltage distribution, \( t = 10^8 \) s, \( DFS = 0.5 \), \( L = 45 \) nm, \( W = 100, 200, 300 \) and \( 400 \) nm.

<table>
<thead>
<tr>
<th>( W ) [nm]</th>
<th>( \mu ) [mV]</th>
<th>( \sigma ) [mV]</th>
<th>( \sigma/\mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>32.1</td>
<td>11.5</td>
<td>0.353</td>
</tr>
<tr>
<td>200</td>
<td>31.3</td>
<td>7.9</td>
<td>0.253</td>
</tr>
<tr>
<td>300</td>
<td>31.1</td>
<td>6.5</td>
<td>0.207</td>
</tr>
<tr>
<td>400</td>
<td>31.0</td>
<td>5.6</td>
<td>0.181</td>
</tr>
</tbody>
</table>

2.2 Analytical Model of Process Variations

Process variations tend to have a normal distribution [11]. Corner models are commonly used by chip designers to evaluate process variations. In this model, the average \( V_{th} (\mu) \), the \( V_{th} \) conditions \( \mu + 3\sigma \) and \( \mu - 3\sigma \) are defined as typical, slow and fast conditions, respectively. We analyzed the delay characteristics of these three corners in timing designs.

In this paper, we assume the distribution of threshold voltage shifts caused by process variation has a normal distribution \( N_{PV} \) as shown in Eq. (10):

\[
N_{PV} = N(\mu_{PV},\sigma_{PV}) \tag{10}
\]

where, \( \mu_{PV} = 0 \) mV and \( \sigma_{PV} = 10 \) mV [23]. Note that we assume \( \mu_{PV} \) and \( \sigma_{PV} \) are constant among various gate widths.

2.3 Circuit Simulation with Variation-Aware Netlist

Variation-aware netlists are used to analyze the characteristics of BTI and process variations: we used it in our simulation method with the HSPICE circuit simulator. Predicted values of the \( \Delta V_{th} \) at stress time \( t \) are applied to each MOSFET in the netlists. The values are calculated using the parameters \( L, W, DFS \) and \( t \) of the transistor in the netlists with the variation model (more details follows in the next section). Figure 4 shows the variation-aware netlist created from the original netlist and the variation model for the simulation analysis.

In BSIM4, the parameter of threshold voltage \( V_{TH0} \) of each MOSFET can be controlled by the parameter DELVTO [24]. \( V_{th} \) of MOSFET shifts in response to the amount of DELVTO. Consider the example of the original netlist and the variation-aware netlist shown in Fig. 5, an NMOS and a PMOS is applied to \( \Delta V_{th} \) of the values of \( dv_{th,n}(t) \) and \( dv_{th,p}(t) \), respectively. Note that we need one variation-aware netlist for each \( t \) condition.
2.4 Delay Analysis with Variation of BTI and Process

This section describes a method to analyze circuit delay caused by BTI and process variations in logic circuits. Both variations follow a normal distribution, \( N_{BTI}(\mu_{BTI}, \sigma_{BTI}) \) and \( N_{PV}(\mu_{PV}, \sigma_{PV}) \).

We consider variations that fall in the \( \pm 4\sigma \) region [25]. The parameters \( \mu_{BTI} \) and \( \sigma_{BTI} \) are calculated using the model introduced in Section 2.1.

In this experiment we simulated the rise delay time of the 45 nm CMOS inverters. The simulation conditions were as follows:

- \( V_{dd} = 1 \) V, \( DF = 0.5 \), \( L_P = L_N = 45 \) nm, \( W_P = 460 \) nm, \( W_N = 240 \) nm.
- The delay times are intervals between 0.5 V of the inputs and 0.5 V of the output.
- The output of inverters is connected to FO4 inverters.

We assume PMOSFETs and NMOSFETs have the same distributions for the purpose of simplification.

The dash line in Fig. 6 shows the result of delay analysis without the BTI variation. The x and y axes are the stress time \( t \) and the predicted delay, respectively. Note that the x axis is in log-scale. We applied the BTI-induced \( \Delta V_{th} \) for the condition of process corner \( \mu_{PV} + 3\sigma_{PV} \). The lifetime delay of the condition is the worst-case delay. However, this result is too optimistic because the BTI variation should be considered in the scaled process.

The dot-and-dash line in Fig. 6 shows the result of delay analysis with the process corner \( \mu_{PV} + 3\sigma_{PV} \), which are applied by the BTI-induced degradation of \( \mu_{BTI} + 3\sigma_{BTI} \). The worst-case delay scenario is observed when both variations are at their worst. We consider these results pessimistic because such conditions rarely occur.

We propose a delay analysis method that uses the sum distribution of BTI and process variations to prevent overly pessimistic or optimistic predictions. The shifts of the threshold voltage based on the assumption that BTI and process variations have a normal distribution are shown in Eq. (11).

\[
N_{\Delta V_{TH}}(\mu_{\Delta V_{TH}}, \sigma_{\Delta V_{TH}}) = N_{BTI} + N_{PV} \tag{11}
\]

\[
\mu_{\Delta V_{TH}} = \mu_{BTI} + \mu_{PV} \tag{12}
\]

\[
\sigma_{\Delta V_{TH}} = \sqrt{\sigma_{BTI}^2 + \sigma_{PV}^2} \tag{13}
\]

Note that Eq. (11) is the sum of the distributions of both variations.

The solid line in Fig. 6 shows the result of the delay analysis with the corner of the proposed model \( \mu_{\Delta V_{TH}} + 3\sigma_{\Delta V_{TH}} \). The condition is the worst case. The statistical results we have obtained from our proposed variation model will be used in the section that follows.

3. Transistor Size Optimization Considering BTI and Process Variations

In this section, we introduce a technique for transistor size optimization in logic circuits that takes into account BTI and process variations. We compared the results using our proposed technique with those obtained using conventional methods to ascertain any differences in delay degradation.

Our technique has three key features. First, the optimization is proven statistically by calculating the distributions of both BTI and process variations. Second, the extension in circuit lifetime and the reduction in timing margin were achieved by considering the stress conditions and lifetime characteristics. Third, our technique does not require any area overhead.

3.1 Size Optimization with Lifetime Delay Variation

We propose a transistor size optimization for logic gates that takes into account lifetime delay variation. The core concepts of the proposed transistor size optimization are:

- It considers lifetime delay not just initial delay
- It uses transistor size optimization to reduce lifetime delay variations

We believe that timing margins for variations (BTI and process) should be applied to circuit designs. Figure 7 is a comparison of the timing margin in our proposed transistor size optimization and a conventional design.

In a conventional design, transistor sizes of logic gates are op-
timized by considering the initial delay as a primary factor. The purpose of the optimization is defined to minimize the root mean square (RMS) of $t_{df}$ (rising delay time) and $t_{ad}$ (fall delay time) of the logic gates $d_{RMS}$ in Eq. (14).

$$d_{RMS} = \sqrt{\frac{t_{df}^2 + t_{ad}^2}{2}}$$

(14)

However, BTI variation depends primarily on both transistor size and the stress conditions. The average of the variation increases as the stress time $t$ increases and the $DF$ increases. The standard deviations of the variation increases when transistor size decreases.

The BTI-induced degradations for $t_{df}$ and $t_{ad}$ differ because degradations on PMOSFETs and NMOSFETs are not equivalent. The lifetime RMS of gate optimized by the initial delay is not the case. We limited our conditions of each input are different. As expected, the optimized size of PMOSFET is larger than indicated in the conventional method. The NMOSFET value should be large when $DF = 1.0$, and vice versa. Both PMOSFETs and NMOSFETs are degraded equivalently when $DF = 0.5$.

Figure 9 shows RMSs of the initial-based and the lifetime-based optimized inverters. The $x$ and $y$ axes denote the $W_p/W_N$ and delay RMS, respectively. In the case of initial-based inverters, RMSs of the initial delays are evaluated. The initial-based optimized size of PMOSFET is $W_p/W_N$: 460 nm/240 nm. In other words, the RMS is the smallest when the size is $W_p/W_N$: 460 nm/240 nm.

The optimized size of the inverters when $DF = 1.0$, $DF = 0.5$ and $DF = 0.0$ are $W_p/W_N$: 440 nm/260 nm, $W_p/W_N$: 460 nm/240 nm and $W_p/W_N$: 470 nm/230 nm, respectively. As expected, the optimized size of PMOSFET when $DF = 0.0$ was larger than the conventional method, and similarly the optimized size of NMOSFET when $DF = 1.0$ was also larger than the conventional method. The optimized size when $DF = 0.5$ was the same with our technique and the conventional method.

Since the $DF$ conditions of each input are different, $c$ cells should be designed, where the number of stress conditions is $c$ and the gate has multiple $n$ inputs. We applied this principle to optimize the NAND2 gate shown in Fig. 10.

Figure 11 shows RMSs of both the initial-based optimized NAND2 gates and the lifetime-based optimized NAND2 gates. The $x$ and $y$ axes are the $W_p/W_N$ and delay RMS, respectively. These results reflect the RMSs of the input $B$ since it has delays that are more critical than those of input $A$. The initial-based optimized size is $W_p/W_N$: 410 nm/290 nm.
The optimized size of the NAND2 gates in the lifetime conditions is shown in Table 5. As expected, it shows a trend similar to the inverters. The results when $DF_A = DF_B = 0.0$ and $DF_A = 1.0, DF_B = 0.0$ are the same. This is because the effects of the degradations of PMOS2 are dominant under these conditions.

### 3.3 Estimating Path Delay with the Proposed Optimization

We estimated the lifetime delays of inverter chains which were optimized by the conventional initial-based and the proposed lifetime-based technique.

The simulated circuits are shown in Fig. 12. Circuit stress is assumed to be $DF = 0.0$. The circuit optimized by the conventional technique consists of inverters in a single size. The circuit optimized by the proposed lifetime delay method consists of inverters in two sizes optimized under the condition that $DF = 0.0$ and 1.0. The other simulation conditions were the same as in Section 3.2.

The initial delays in the conventional and proposed methods are 64.8 ps and 66.5 ps, respectively. The conventional circuit resulted in a shorter initial delay because it uses a technique that minimizes such initial delay. However, we expect the lifetime delay of our proposed circuit to be shorter.

Figure 13 shows the estimated lifetime delay in the simulation circuits. The x and y axes represent the stress time and the predicted delay, respectively. Note that the x axis is log-scale. The estimated lifetime delay of our proposed method is 73.2 ps which is a 4.4% improvement over the 76.6 ps in the conventional method.

### 4. Conclusion

Conventional lifetime-based designs for logic circuits tend not to consider BTI and process variations: this leads to overstating or understating the predicted delay degradation in circuits. In this report we propose a statistical size optimization technique for logic circuits that take these variations into consideration. This statistical method produces more accurate results because it is based on the sum distributions of the variations. Although experimental, our results clearly show that circuit lifetime is extended and the timing margins are reduced by 4.4% compared with the conventional technique. We believe it is highly likely that our technique can support more dependable and efficient chip designs than what is currently available.

**Acknowledgments** This work is supported by VLSI De-
References


Kazutoshi Kobayashi received his M.E. and Ph.D. in Electronic Engineering from Kyoto Institute of Technology, Kyoto, Japan in 1993 and 1999, respectively. Starting as an Assistant Professor in 1993, he was promoted to associate professor in the Graduate School of Informatics, Kyoto University, and stayed in that position until 2009. For two years during this time, he acted as associate professor of VLSI Design and Education Center (VDEC) at the University of Tokyo. Since 2009, he has been a professor at Kyoto Institute of Technology and has led his research group to give presentations in ASP-DAC, ISQED, IRPS, NSREC and SSDM in 2011. While in the past he focused on reconfigurable architectures utilizing device variations, his current research interest is in improving the reliability (soft errors and Bias temperature Instability) of current and future VLSIs. He was the recipient of the IEICE best paper award in 2009.

Michitarou Yabuuchi received his M.E. and Ph.D. in Electronics and System Engineering from Kyoto Institute of Technology, Kyoto, Japan in 2012 and 2015, respectively. He is presently a post doctor at Kyoto Institute of Technology. His interests are in reliability issues and dependable circuits.