Worst-case Performance Analysis Under Random Telegraph Noise Induced Threshold Voltage Variability

A.K.M. Mahfuzul Islam
Institute of Industrial Science
The University of Tokyo
Meguro-ku, Tokyo, Japan
Email: mahfuzul@iis.u-tokyo.ac.jp

Hidetoshi Onodera
Graduate School of Informatics
Kyoto University
Sakyo-ku, Kyoto, Japan
Email: onodera@vlsi.kuee.kyoto-u.ac.jp

***Abstract*—RTN induced threshold voltage distribution has a long tail that can degrade the worst-case distribution severely. In this paper, we analyze the effect of RTN on worst-case performance based on variability models extracted from a 65 nm silicon-on-thin-body low threshold voltage process. Monte Carlo based simulation results reveal that with the lowering of supply voltage, RTN can degrade the worst-case delay by more than 10 % when the number of critical paths is 10. The worst-case delay degradation can go as high as 100 % if the critical path number increases to 100. Because of the RTN induced threshold voltage fluctuation, several outliers appear at near/sub-threshold operation. Considering RTN amplitude can increase at weak-inversion operation, low-voltage operation needs careful consideration of RTN.**

I. Introduction

With technology scaling, random telegraph noise (RTN) has become an increasing concern [1, 2]. There are reports of RTN effects on the performance of SRAMs [3], ring oscillators (RO) [4], and Flash memories [5]. At low voltage operation, frequency fluctuation of 10.4 % has been reported for a 40 nm process [4]. It is known that worst-case distribution of random variables is sensitive to long tail in the parent distribution. RTN induced ΔV_{T} distribution is reported to follow Lognormal or Exponential distribution meaning the distribution has a long tail. On the other hand, random dopant fluctuation (RDF) induced WID random variation is reported to follow a Normal distribution [6, 7]. Normally, RDF induced WID random variation is much larger than that induced by RTN. A transistor encounters both of the within-die (WID) random variation and RTN induced random variation simultaneously. As a result, the total variation is a statistical sum of WID random variation and RTN induced random variation. It is not clear whether RTN can make a large impact on the worst-case distribution when both of the WID variation and RTN are considered. Therefore, a theoretical analysis of worst-case performance is useful to asses the impact of RTN on performance degradation.

In analyzing the effect of RTN on performance, following difficulties arise.

1) Total ΔV_{T} distribution of transistors is the sum of Normal (WID) and Lognormal (RTN) distributions.
2) At low voltage, because of the non-linearity in delay against ΔV_{T}, suitable distribution to represent the delay variation is difficult to define.

RTN induced ΔV_{T} variation is much smaller than WID ΔV_{T} variation. Even though, RTN is reported to affect the performance of SRAM and Flash memories at low-voltage operation where large number of devices operate in parallel. Besides, anomalous current fluctuations have been reported because of RTN at sub-threshold operation [8]. Our hypothesis is that the long tail in the RTN amplitude distribution causes anomaly in the performance distribution. According to our knowledge, there is no report on the theoretical analysis of worst-case performance distribution considering RTN. We, therefore, attempt to provide a theoretical framework to asses the impact of RTN on worst-case distributions. In this paper, we focus firstly on the worst-case ΔV_{T} distributions. Then, we perform Monte Carlo based analysis to obtain worst-case delay distribution using analytical delay models. We investigate the effect of supply voltage on the worst-case distributions. The WID and RTN induced ΔV_{T} distribution models we used in this analysis are derived from a 65 nm silicon-on-thin-body low threshold voltage process. We use skewed ring oscillators (RO) to capture the delay fluctuations and extract variability models for ΔV_{T} distributions. Our delay analysis based on the extracted variability models show that worst-case performance degradation due to RTN can be more than 10 % at low-voltage operation. Because of the long tail induced by the RTN, the worst-case performance is sensitive to parallel number of paths. For a parallel number of 100, the worst-case delay degradation due to RTN can be as high as 100 % at near/sub-threshold operation. Therefore, it is important to accurately characterize RTN variability and its impact on circuit performance especially for low-voltage operation.

The remainder of the paper is organized as follows. We describe a variability modeling methodology considering both of the WID variation and RTN induced variation in Sec. II. In Sec. III, we present maximum value distribution theory and its properties that are the base of our analysis. We then present

978-1-5386-6365-3/PATMOS/©2018 IEEE

our analysis results in Sec. IV. Finally, we put our concluding remarks in Sec. VI.

II. Variability Modeling

A. Modeling

We need to consider both of the RDF induced within-die static ΔV_{T} variation and RTN induced ΔV_{T} variation. Thus, the V_{T} of a transistor can be expressed as follows.

$$V_{\mathrm{T}} = V_{\mathrm{T0}} + \Delta V_{\mathrm{T,WID}} + \Delta V_{\mathrm{T,RTN}}, \quad (1)$$

$$\Delta V_{\mathrm{T}} = \Delta V_{\mathrm{T,WID}} + \Delta V_{\mathrm{T,RTN}}. \quad (2)$$

Note that we have omitted the global die-to-die variation component and assume that V_{T0} includes the global component. Furthermore, we assume that most of the current fluctuations are caused by the change in transistor threshold voltage for simplicity. Here, we focus on the worst-case distribution based on the statistical max operation. In order to obtain the total ΔV_{T} distribution, we need to confirm the following properties.

1) Suitable distributions for WID variation and RTN induced ΔV_{T} variation, and
2) Correlation between the WID and RTN variations.

There is one significance difference between the two types of variations of WID and RTN when estimating performances of a delay path for example. WID variation is static and thus statistical sum operation can be performed. However, RTN is a dynamic phenomenon where current fluctuations occur at different probabilities at different transistors. Therefore, a straight forward statistical sum operation cannot be performed. In this paper, we make the following approximation so that we can treat RTN induced ΔV_{T} variation as static.

> *Assumption 1:* Given enough time of operation, there exist at least one time frame where all fluctuations occur simultaneously in a delay path.

The above approximation will result in pessimistic evaluation of performance distributions. None the less, analysis based on the above approximation will give us useful insights. Now that we treat both of the WID and RTN ΔV_{T} distributions as static, we can perform convolution to obtain the overall ΔV_{T} distribution given that $\Delta V_{\mathrm{T,WID}}$ and $\Delta V_{\mathrm{T,RTN}}$ are independent to each other. The threshold voltage V_{T} of a transistor is related to the average surface potential, whereas RTN amplitude depends on the exact surface potential value under the trap. Therefore, it is expected that no significant correlation exists between transistor V_{T} and RTN induced ΔV_{T}. Our measurement results also validate this fact which will be presented later in the section. Literature reports also support the assumption [9]. Thus, WID random variation and RTN induced ΔV_{T} can be treated as independent to each other.

B. Measurement Circuit

In this paper, two different types of skewed ROs are used to evaluate WID random variation and RTN.

1) RO type #1 (pMOSFET dominant): small pMOSFET and large nMOSFET.
2) RO type #2 (nMOSFET dominant): large pMOSFET and small nMOSFET.

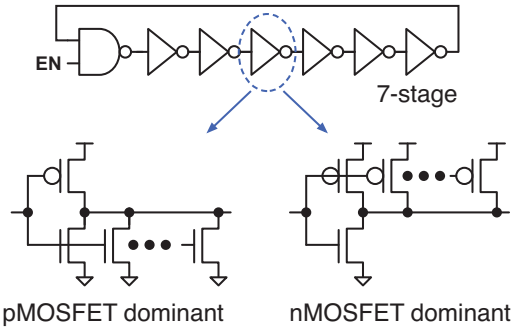

Fig. 1: Ring oscillator structure to characterize pMOSFET and nMOSET RTN. Large skewed inverter cells are used. Different P/N ratio is utilized to decouple nMOSFET and pMOSFET RTN.

The chip contains 208 instances of identical ROs to characterize statistical properties. Each RO instance is measured for 10 s with an integration time of 1 ms to capture RTN characteristics. Supply voltage is varied from 0.4 V to 1.0 V. The details of the test chip is explained in [10].

In order to estimate ΔV_{T} distributions from the measured $\Delta d/d$ distributions, A linear model as Eq. (3) based on simulation based sensitivity coefficients is utilized.

$$\frac{\Delta d}{d} = \sum_{i=1}^{N_{\mathrm{Tr}}} \left(k_i \cdot \Delta V_{\mathrm{T},i}\right). \quad (3)$$

Here, N_{Tr} is the number of transistors in the RO. k_i is sensitivity coefficient which can be calculated by SPICE simulation using MOSFET model provided by the foundry.

In RO #1, the delay fluctuations will be dominated by the smaller pMOSFETs. Similarly, in RO #2, the delay fluctuations will be dominated by the smaller nMOSFETs. Thus, utilizing the differences in the variability sources, statistical fitting on the measured distribution can be performed to estimate the parameters of ΔV_{T} distributions [10].

C. WID Variability

It is reported that static process V_{T} distribution follows Normal distribution up to $\pm 5\sigma$ for planar and FinFET technologies [6, 7]. The normality of WID ΔV_{T} variability is important for yield and performance analysis of circuits with large number of devices such as the SRAM. For our target process, we extract the WID random variation of pMOSFET and nMOSFET from RO frequency variations. Fig. 2 shows QQ plots of delay variations for pMOSFET and nMOSFET dominant ROs. Supply voltage is set to 0.6 V such that the transistor operate in strong-inversion. In this operating region, the delay change against V_{T} change is linear. Thus, as the plot is a near straight line, we confirm that the WID random variation follows normal distribution. We then estimate the μ and σ of Normal distributions for pMOSFET and nMOSFET using

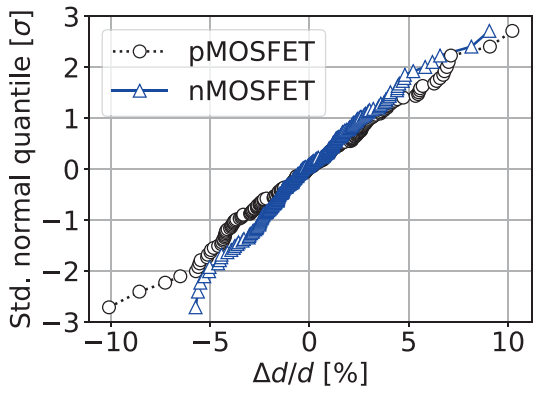

Fig. 2: QQ-plot of WID delay variation. Supply voltage is 0.6 V.

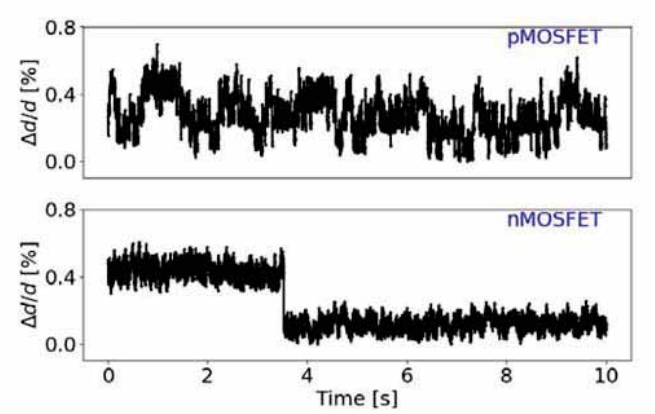

Fig. 3: Waveform examples RTN induced delay fluctuations.

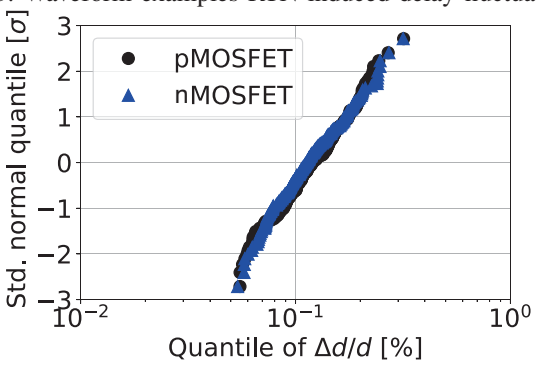

Fig. 4: QQ plot of RTN induced $\Delta d/d$ for pMOSFET and nMOSFET dominant ROs.

which the simulated distribution matches with the measured distributions. The resulted Normal distribution parameters are $\mu = 0$ mV and $\sigma = 15$ mV. The transistor V_{T} is near 0.3 V in the target process.

D. RTN Amplitude Variability

To assess whether RTN induced ΔV_{T} distribution is valid for maximum value statistics, we need a theoretical model and experimental results that support the theory. A simple model for RTN amplitude for a single trap based on surface potential fluctuation is proposed where the μ and σ of a Lognormal distribution, $\mathcal{LN}(\mu, \sigma^2)$ is expressed by Eqs. (4) and (5) [11].

$$\mu_{\ln(\Delta V_{\mathrm{T}})} = \ln\left(\frac{q}{C_{\mathrm{ox}} W L}\right), \tag{4}$$

$$\sigma_{\ln(\Delta V_{\mathrm{T}})} = \frac{q}{kT} \times \sigma_{\phi_s}. \tag{5}$$

The overall ΔV_{T} probability distribution is then computed as follows.

$$P_1(x) = \frac{1}{x\sigma\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \tag{6}$$

$$P_N(x) = \int_{-\infty}^{\infty} P_{N-1}(x-t) \cdot P_1(t) \cdot \mathrm{d}t, \tag{7}$$

$$a_N = \frac{\lambda^N \mathrm{e}^{-\mathrm{N}}}{N!}, \tag{8}$$

$$P(x) = a_0\delta(x) + \sum_{i=1}^{N} a_i P_i(x). \tag{9}$$

Here, $P_1(x)$ is the probability distribution function (PDF) for a single trap ΔV_{T} and $P_N(x)$ is the PDF of total ΔV_{T} because of N traps. $a_N(x)$ is probability mass function (PMF) for the number N of traps. $P(x)$ is the PDF of overall ΔV_{T} incorporating the trap number distribution.

Fig. 3 shows two waveform samples for RO oscillation period. Fig. 4 shows QQ plots of $\Delta d/d$ for pMOSFET and nMOSFET dominant ROs. X-axis is plotted in logarithm scale. Y-axis shows the standard normal quantile. Straight lines in the QQ plots mean that the distributions follow Lognormal distributions. Assuming ΔV_{T} distribution follow a Lognormal distribution, $\mathcal{LN}(\mu, \sigma^2)$, μ and σ are extracted to be -0.15 and $\sigma = 0.90$, respectively. Similar ΔV_{T} distributions are estimated for both of the pMOSFET and nMOSFET.

As large tail in the RTN induced ΔV_{T} amplitude distribution originates from the surface potential variability [12], we assume that the logarithm of ΔV_{T} has Normality up to 5σ at least. Assuming surface potential variability of $\sigma_{\phi_s} = 20$ mV [13], the values of μ and σ from Eqs. (4) and (5) results in -0.33 and 0.77 at $T = 298$ K. $W = 140$ nm, $L = 60$ nm, gate oxide thickness, $T_{\mathrm{ox}} = 1.3$ nm and $\epsilon_r = 3.9$ is assumed in this calculation. Then considering a Poisson distribution $\mathrm{Poiss}(\beta)$ with $\beta = 1.2$ [9], we obtain a distribution according to Eq. (9). Fitting a Lognormal distribution to the obtained final distribution yields $\mu = -0.14$ and $\sigma = 0.98$. Thus, the reported Lognormal model for our process supports the theoretical assumptions.

Finally, we show a comparison between the distributions to represent WID random variation and RTN induced ΔV_{T} variation in Fig. 5. The left hand Y-axis shows the probability and the right hand Y-axis shows the cumulative probability of the ΔV_{T}. It is clear that RTN induced ΔV_{T} is much smaller than the WID random ΔV_{T} variation. However, we will show in Sec. IV that the long tail in the RTN induced ΔV_{T} distribution can significantly degrade the worst-case performance.

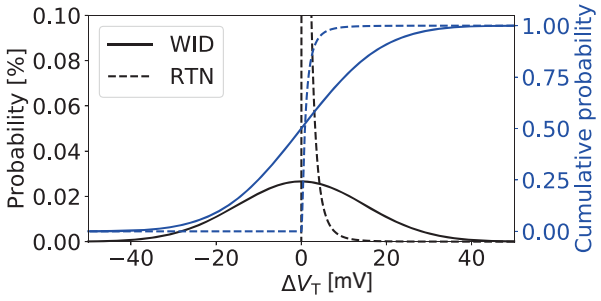

Fig. 5: ΔV_{T} distributions for WID and RTN.

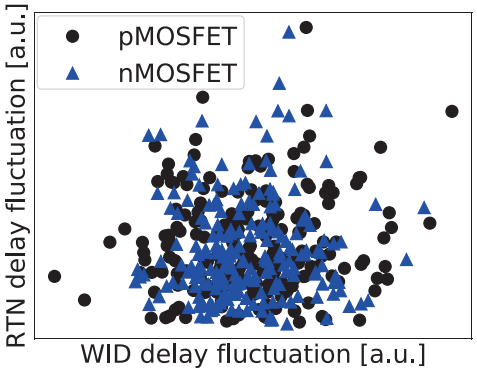

Fig. 6: Correlation between WID delay variation and RTN induced delay fluctuation. No correlation is observed.

E. Correlation between WID and RTN Variability

As explained in Sec. II-A, it is important to verify whether WID random variation and RTN induced random variation is correlated or not. In Fig. 6, we plot RTN induced delay fluctuation against WID random delay variation for a pMOSFET dominant RO and an nMOSFET dominant RO. It is clear from the figure that no correlation exists for both of the pMOSFET and nMOSFET dominated ROs.

III. Maximum Value Statistics

Here, we show how the maximum value distribution can be obtained from the parent distributions of transistor ΔV_{T}. Let $\{x_1, x_2, \cdots, x_N\}$ be a set of random variables with identical PDF and CDF of $p(x)$ and $P(x)$. The sample maximum is defined as

$$\text{Maximum } M = \max(x_1, x_2, \cdots, x_N). \tag{10}$$

If the joint probability distribution $P(x_1, x_2, \cdots, x_N)$ is known, then the cumulative distribution of maximum M can be expressed as follows.

$$Q_N(x) = \text{Prob}[M \leq x, N] \tag{11}$$

$$= \text{Prob}[x_1 \leq x, x_2 \leq x, \cdots, x_N \leq x] \tag{12}$$

$$= \int_{\infty}^{x}\int_{\infty}^{x}\cdots\int_{\infty}^{x} \mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_N P(x_1, x_2, \cdots, x_N). \tag{13}$$

For independent and identical random variables, the joint PDF factorises and we get

$$Q_N(x) = \left[\int_{\infty}^{x} \mathrm{d}x\ p(x)\right]^N = \left[1 - \int_{x}^{\infty} \mathrm{d}x\ p(x)\right]^N. \tag{14}$$

According to Gnedenko's classical law of extremes, $Q_N(x)$ approaches the following limiting form.

$$Q_N(x) \xrightarrow[x\to\infty, N\to\infty]{z=(x-a_N)/b_N \text{ fixed}} F(\frac{x-\mu}{\sigma}) = F(z). \tag{15}$$

For a parent distribution with Gaussian like tail, $F(z)$ can be approximated by Gumbel distribution [14].

$$F(x;\mu,\sigma) = \mathrm{e}^{-\mathrm{e}^{(\mathrm{x}-\mu)/\sigma}}, \tag{16}$$

$$f(x;\mu,\sigma) = \frac{1}{\sigma} \times \mathrm{e}^{(\mathrm{x}-\mu)/\sigma} \times \mathrm{e}^{-\mathrm{e}^{(\mathrm{x}-\mu)/\sigma}}. \tag{17}$$

Thus, for large number of parallel devices, a Gumbel distribution can be used to model the worst-case distribution. If Monte Carlo simulation is performed, Gumbel distribution can be obtained by fitting the parameters to the simulated distribution.

IV. Worst-case ΔV_{T} Distribution

For circuits like SRAM for example, large number of transistors operate in parallel. We can call the logic depth to be 1 in this case. For these circuits, maximum value distribution of ΔV_{T} is of great importance. Therefore, we analyze the worst-case ΔV_{T} distributions with and without the effect of RTN. First, we demonstrate the impact of tail in the parent distribution on the resultant worst-case distribution. Fig. 7 shows the comparison between maximum value distributions when either WID random variation or RTN induced random variation is considered. Dotted lines show the distributions for WID random variations. Solid lines show the distributions for RTN random variations. Maximum distributions are shown for different transistor numbers in parallel. When the transistor number is 10, maximum distribution of ΔV_{T} induced by RTN is much smaller than that induced by WID random variation. However, with the increase of transistor number, the worst-case distribution degrades rapidly for the case of RTN compared with the case of WID random variation. Thus, long tail in the parent distribution has a severe effect on the worst-case distribution. Therefore, although the variation due to RTN of a single transistor may be negligible, its impact on circuits with large number of transistors should not be ignored.

Fig. 7 shows the worst-case ΔV_{T} distribution for two cases of WID and RTN variations separately. But a transistor in the chip encounters both the WID and RTN variations simultaneously. Therefore, the variation model of Eq. (2) is of relevant to analyze the worst-case distributions. Fig. 8 shows the worst-case ΔV_{T} distributions for two new cases. The first case corresponds to WID random variation only. The second case represents the scenario when both WID and RTN variations are involved. For a single transistor $N = 10^0$, the difference between the worst-case distributions for the two cases are insignificant. However, with transistor number crossing 10^3, clear deviation between the two cases becomes

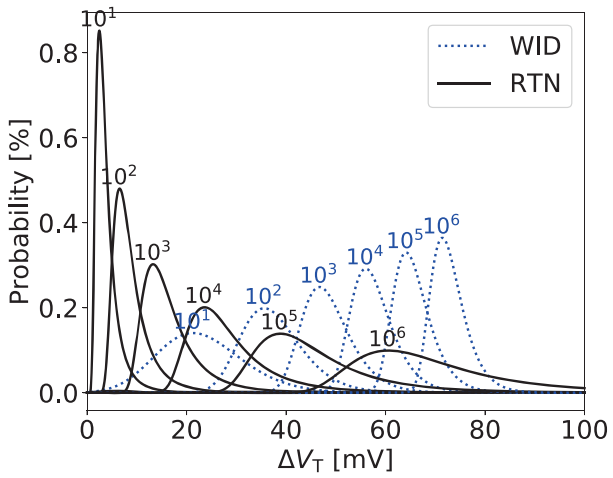

Fig. 7: Maximum distributions of ΔV_{T} with two different parent distributions of Normal and Lognormal for different transistor count, N. Normal distribution represents RDF induced process variation whereas Lognormal distribution represents RTN induced ΔV_{T} variability.

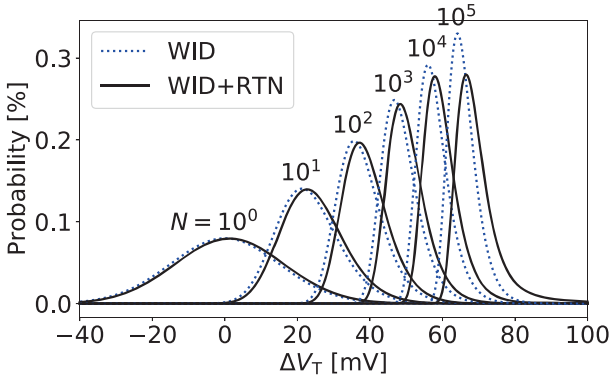

Fig. 8: Maximum distributions of ΔV_{T} for different transistor count, N when both of the static variation and RTN induced variation is considered.

visible. For transistor number $N = 10^5$, long tail appears in the worst-case distribution for the second case. This long tail may cause malfunctions in circuits. As a result, for circuit with large number of parallel devices, adequate margin is required to ensure high yield.

V. Worst-case Delay Distribution

A. Delay Model

The delay of a single stage in a delay path can be expressed by Eq. (18).

$$\tau = \frac{C_{\mathrm{L}}\, V_{\mathrm{DD}}}{I_{\mathrm{D}}}. \tag{18}$$

Here, τ is the gate delay, C_{L} is load capacitance, V_{DD} is supply voltage, and I_{D} is transistor ON current. The total delay, d of a delay path consisting of L_{d} logic gates is then expressed as follows.

$$d = \sum_{i=1}^{L_{\mathrm{d}}} \tau_i. \tag{19}$$

Here, τ_i is the delay of i-th stage. For modeling digital circuit switching delay, the α-power law current model of Eq. (20) is widely used [15].

$$I_{\mathrm{D}} = b_1 \cdot \frac{W}{L} \cdot (V_{\mathrm{DD}} - V_{\mathrm{T}})^{\alpha}. \tag{20}$$

Here, b_1 is a constant related to technology, W is gate width, and L is gate length. The value of α is typically 1.3 for scaled processes. The model equation of Eq. (20) is suitable for analyzing logic delay when the transistors operate at strong-inversion. For transistors operating at weak-inversion, the following current model is widely used.

$$I_{\mathrm{D}} = b_2 \cdot \frac{W}{L} \cdot \exp\left(\frac{V_{\mathrm{DD}} - V_{\mathrm{T}}}{n\, U_T}\right). \tag{21}$$

Here, b_2 is a technology related fitting parameter, n is the sub-threshold swing parameter and U_T is the thermal voltage. Using current models based on the operating region may cause inconsistent result across the operating regions. Therefore, to assess the variability effect for all the regions, a model that is continuous and valid from weak-inversion to strong-inversion operation is required. Such a model is the EKV model [16]. The original EKV model is intended for low-voltage analog circuit design, where velocity saturation is not considered. In order to model the velocity saturation effect, we have used a modified version of Eq. (22).

$$I_{\mathrm{D}} = b_3 \cdot \frac{W}{L} \cdot \ln\left[1 + \exp\left(\frac{V_{\mathrm{DD}} - V_{\mathrm{T}}}{\alpha\, n\, U_T}\right)\right]^{\alpha}. \tag{22}$$

Here, b_3 is a technology related parameter. The validity of the model Eq. (22) is shown in Fig. 9. Transistor ON current across V_{DD} is shown in the figure for the three models of α-power law model, the modified EKV model and the exponential model for weak-inversion. The α-power law model becomes inaccurate when V_{DD} approaches V_{T}. Consequently, the exponential model becomes inaccurate when V_{DD} crosses V_{T}. The EKV model is able to represent both the strong-inversion and weak-inversion currents as well as the region in between. There, delay analysis using EKV model is appropriate to asses the relative delay variations across the different operation regions.

B. Worst-case Delay Distribution

1) Analysis Method: Overall ΔV_{T} distribution including both of the WID random variation and RTN is a statistical sum of a Normal and a Lognormal distribution. Thus finding a valid distribution to represent the overall ΔV_{T} distribution is challenging. In this paper, we perform Monte Carlo simulation by generating random ΔV_{T} for WID and RTN induced ΔV_{T}

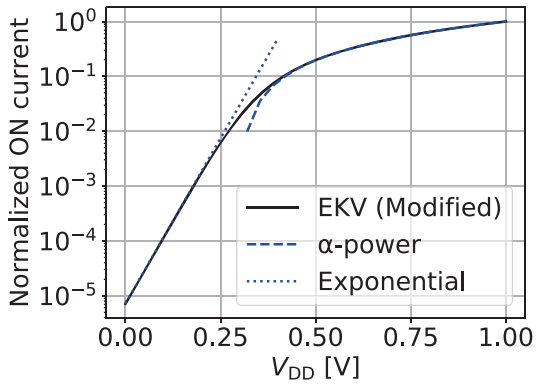

Fig. 9: Comparison between different current models.

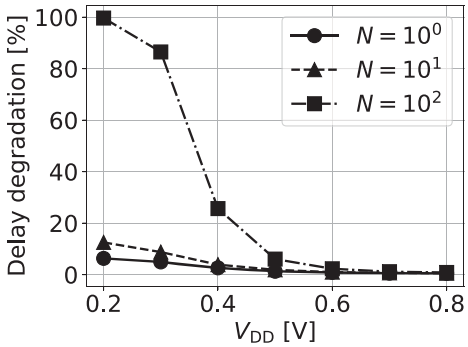

Fig. 11: Increase of worst-case delay due to RTN against that without RTN.

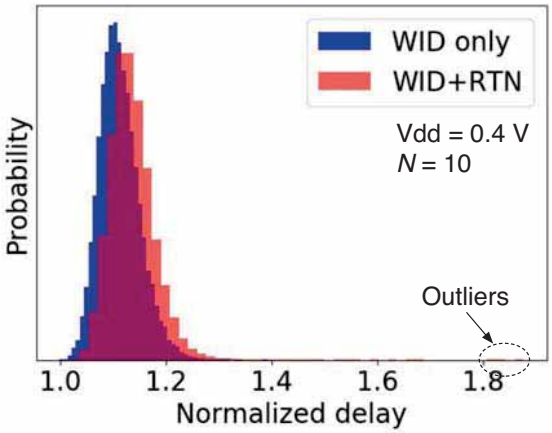

Fig. 10: Histograms of simulated delay. Outliers appear in the case of considering RTN induced ΔV_{T} variation.

and simulate the corresponding delay using Eqs. (18) and (22). We then calculate the 99.7 percentile delay value in the maximum delay distribution as the worst-case delay. $50,000$ maximum delay samples are simulated in our experiment.

2) Results: Now, we analyze the worst-case distribution for delay paths considering the RTN along with WID random variation. Distributions for two cases are simulated where in the first case only the WID random variation is considered. In the second case, RTN random variation is also considered with the WID variation. Fig. 10 shows the histogram of the maximum delay for the two cases. The parallel path number N is 10 is this case. The logic depth that is the number of logic cells in the path is set to 10 here. V_{DD} is 0.6 V. The 99.7 percentile value in the worst-case distribution increases by more than 5 % with the inclusion of RTN variation. However, as shown in the figure, multiple outliers appear in the second case where RTN is considered. Because of the long tail property in the RTN induced ΔV_{T} distribution, anomalous delay behavior occurs. When only WID is considered, no such anomaly was found.

C. V_{DD} Dependency

Fig. 11 plots the increase of 99.7 percentile delay value caused by the RTN variation against the supply voltages. When V_{DD} is much higher than V_{T}, the increase is below 1 %. Thus, we may not need to worry much on the effect of RTN on worst-case delay estimation. Worst-case delay estimation considering only the WID random variation might suffice. With the decrease of V_{DD}, the delay degradation increases. At 0.3 V, the delay degradation is more than 10 % when the parallel path number is 10. With the increase of parallel path number to 100, the delay degradation is accelerated and we observe delay degradation of near 100 % at sub-threshold operation. At low-voltage operation, there remains a possibility that the circuit may malfunction in some cases. Additional margin to account RTN induced delay degradation is a must for low-voltage operation.

D. Outliers

One key characteristic found in the worst-case delay analysis is the appearing of significant number of outliers in the distributions. To quantify the number of outliers, we define the threshold of outlier to be 1.5 times the worst-case delay. Fig. 12 shows the number of outliers against the supply voltage. When RTN is not considered, no outliers exist in the distribution. However when RTN is considered, the outliers begin to appear with the decrease of V_{DD}. Therefore, at low-voltage operation some anomalous delay can be found making low-voltage operation further challenging. Our simulation results are supported by some of the literature reports [4, 8]. We provide for the first time a thorough analytical analysis to explain the properties.

VI. Conclusion

In this paper, we have evaluated worst-case ΔV_{T} and delay distributions under the variability induced by both of the random dopant fluctuation (RDF) and random telegraph noise (RTN). Our analysis is based on measurement results of delay fluctuations from a 65 nm silicon-on-thin-body low-V_{T} process. Analysis results show that at strong inversion operation, the consideration of RTN induced ΔV_{T} variability degrades the worst-case delay by only less than 1 %. However,

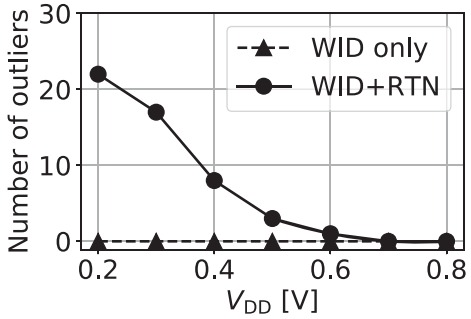

Fig. 12: Number of outliers against supply voltage.

when the supply voltage is lowered below the transistor threshold voltage, the performance degradation increases to more than 10 %. When the number of parallel paths is 100, the performance degradation can reach near 100 %. We have observed significant number of outliers appear when RTN is considered. Thus, we surmise that RTN induced ΔV_{T} variability will play a significant role at low supply voltage operation. As the underlying cause for the RDF and RTN induced variability is the surface potential fluctuation, the analysis results can be scaled to a finer process.

Acknowledgment

The authors would like to thank Masashi Oka of Kyoto University for helping the measurements. The VLSI chip in this study was fabricated in the chip fabrication program of the VLSI Design and Education Center (VDEC), the University of Tokyo, in collaboration with Renesas Electronics. This work has been supported in part by JSPS KAKENHI 25280014 and 16H01713.

References

[1] M. Kirton and U. M. in Physics, "Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency (1/f) noise," *Advances in Physics*, vol. 38, no. 4, pp. 367–468, 1989.

[2] N. Tega, H. Miki, F. Pagette, D. Frank, M. Rooks, W. Haensch, and K. Torii, "Increasing threshold voltage variation due to random telegraph noise in FETs as gate lengths scale to 20 nm," *Symposium on VLSI Technology Digest of Technical Papers*, pp. T50–T51, 2009.

[3] H. Qiu, T. Mizutani, Y. Yamamoto, H. Makiyama, T. Yamashita, H. Oda, S. Kamohara, N. Sugii, T. Saraya, M. Kobayashi, and T. Hiramoto, "Impact of random telegraph noise on write stability in Silicon-on-Thin-BOX (SOTB) SRAM cells at low supply voltage in Sub-0.4V regime," *Symposium on VLSI Technology Digest of Technical Papers*, pp. T38–T39, 2015.

[4] T. Matsumoto, K. Kobayashi, and H. Onodera, "Impact of random telegraph noise on CMOS logic delay uncertainty under low voltage operation," in *IEEE Electron Devices Meeting*, 2012, pp. 25.6.1–25.6.4.

[5] A. Ghetti, C. Compagnoni, A. S. Spinelli, and A. Visconti, "Comprehensive analysis of random telegraph noise instability and its scaling in DecaNanometer flash memories," *IEEE Transactions on Electron Devices*, vol. 56, no. 8, pp. 1746–1752, 2009.

[6] T. Tsunomura, A. Nishida, F. Yano, A. T. Putra, K. Takeuchi, S. Inaba, S. Kamohara, K. Terada, T. Hiramoto, and T. Mogami, "Analyses of 5σ Vth Fluctuation in 65nm-MOSFETs using Takeuchi Plot," in *Symposium on VLSI Technology Digest of Technical Papers*, 2008, pp. 156–157.

[7] M. Giles, A. N. Radhakrishna, D. Becher, A. Kornfeld, K. Maurice, S. Mudanai, S. Natarajan, P. Newman, P. Packan, and T. Rakshit, "High sigma measurement of random threshold voltage variation in 14nm logic FinFET technology," in *Symposium on VLSI Technology Digest of Technical Papers*, 2015, pp. T150–T151.

[8] J. Campbell, L. yu, K. Cheung, J. Qin, J. Suehle, A. Oates, and K. Sheng, "Large random telegraph noise in Sub-Threshold operation of Nano-Scale nMOSFETs," in *IEEE International Conference on IC Design and Technology*, 2009, pp. 17–20.

[9] A. K. M. M. Islam, T. Nakai, and H. Onodera, "Statistical analysis and modeling of random telegraph noise based on gate delay measurement," *IEEE Transactions on Semiconductor Manufacturing*, vol. 30, no. 3, pp. 216–226, 2017.

[10] A. K. M. M. Islam and H. Onodera, "Effect of supply voltage on random telegraph noise of transistors under switching condition," in *IEEE/ACM International Symposium on Power and Timing Modeling, Optimization and Simulation (PATMOS)*, 2017, pp. 1–8.

[11] K. Sonoda, K. Ishikawa, T. Eimori, and O. Tsuchiya, "Discrete dopant effects on statistical variation of random telegraph signal magnitude," *IEEE Transactions on Electron Devices*, vol. 54, no. 8, pp. 1918–1925, 2007.

[12] A. Asenov, R. Balasubramaniam, A. Brown, and J. Davies, "RTS amplitudes in decananometer MOSFETs: 3-D simulation study," *IEEE Transactions on Electron Devices*, vol. 50, no. 3, pp. 839–845, 2003.

[13] G. Slavcheva, J. Davies, A. Brown, and A. Asenov, "Potential fluctuations in metal-oxide-semiconductor field-effect transistors generated by random impurities in the depletion layer," *Journal of Applied Physics*, vol. 91, no. 7, pp. 4326–4334, 2002.

[14] S. Majumdar and A. Pal, "Extreme value statistics of correlated random variables," *arXiv preprint arXiv:1406.6768*, 2014.

[15] T. Sakurai and A. 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.

[16] C. C. Enz, F. Krummenacher, and E. A. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications," *Analog integrated circuits and Signal Processing*, vol. 8, no. 1, pp. 83–114, 1995.