This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

Statistical Analysis and Modeling of Random Telegraph Noise Based on Gate Delay Measurement

A.K.M. Mahfuzul Islam, *Member, IEEE,* Tatsuya Nakai, *Student Member, IEEE,* and Hidetoshi Onodera, *Senior Member, IEEE.*.

***Abstract*—We propose a characterization methodology for Random Telegraph Noise (RTN) based on gate delay measurement. To convert delay change to MOSFET threshold voltage fluctuation, ΔV_{T}, a topology-reconfigurable ring oscillator (RO) is utilized. We discuss the issue of detecting RTN-induced discrete fluctuations in the delay and develop a kernel density based method to detect the fluctuations. Characterization results of several RTN properties from a test chip fabricated in a 65 nm bulk process are presented. Particular focus is given on the suitable distribution to present RTN-induced overall ΔV_{T} distribution and its gate area dependency. The results show that Lognormal distribution is better at representing the total ΔV_{T} distribution. RTN-induced delay fluctuation of 40 % has been observed for a single gate under weak inversion operation. Local process variation and RTN amplitude are found to be uncorrelated. The proposed methodology is thus suitable for characterizing RTN of devices operating under switching condition.**

I. Introduction

As the scaling of transistor dimensions continues, transistor variability poses a challenge on designing reliable computing systems [1]. With the emergence of IoT (Internet of Things), the need for energy-efficient computing is more demanding now. Supply voltage lowering is the most effective way to increase energy efficiency [2], but the effect of variability also becomes severe as the circuit performance is now more sensitive to variation. With transistor scaling, variability such as process variation and random telegraph noise (RTN) are reported to increase [3, 4]. For design time optimization and reliability analysis of circuits, RTN needs to be modeled accurately for several gate widths, gate lengths, bias conditions and so on.

RTN is a dynamic phenomenon which is strongly influenced by both the process technology and the design parameters such as the gate width [5]. From a digital circuit's perspective, RTN can be treated as the total V_{T} variation that a transistor may encounter during its operating lifetime. Thus, the time-dependent property can be omitted for simplicity and the ΔV_{T} amplitude can be modeled for use in analysis. A common technique to characterize RTN is to measure device ON-current variation over time under DC bias [4, 6]. AC bias can also be applied to investigate RTN under high-frequency operation of the device [7]. As a transistor in a digital circuit operates under switching condition, it is useful to characterize devices embedded in a digital environment. The cell-based design approach can realize the digital environment of the devices under test. We present an in-situ RTN characterization methodology based on gate delay variation measurement. The circuit is designed with a standard cell-based approach. A topology-reconfigurable ring oscillator (RO) is utilized to measure the gate delay variations such that the delay change can be converted to V_{T} fluctuation directly. To associate the measured delay variations to either pMOSFET or nMOSFET, the gate delay is made sensitive to either pMOSFET or nMOSFET. Thus, we have a detailed view of the transistor level RTN characteristics from a delay perspective. We show that individual RTN states, trap numbers, and the overall ΔV_{T} can be evaluated. Implementing ROs with several gate widths allow us to characterize RTN for different gate widths.

This paper is an extension of Ref. [8]. In this paper, we have included a detailed explanation of the characterization methodology and its verification. We have also updated the RTN detection method to characterize RTN amplitudes and trap numbers. In Sec. II, we describe the statistical modeling approach of V_{T} fluctuation due to RTN. In Sec. III, we describe our RTN characterization methodology from delay variation measurement. Sec. IV demonstrates some results of various RTN properties. In Sec. V, we put our concluding remarks.

II. Characterization and Modeling of RTN

A. RTN Model

RTN is caused by the trapping and de-trapping of carriers into the oxide interface [5]. The number of traps is reported to follow Poisson distribution [6]. However, we can only evaluate the detectable traps, as background noise and measurement resolution limit the trap detectability. V_{T} fluctuation due to a single trap is often expressed by (1).

$$\Delta V_{\mathrm{T}} = \frac{q}{C_{\mathrm{OX}} W L} \tag{1}$$

Here, W is the gate width, L is the gate length, q is the electron charge, and C_{OX} is the gate capacitance per unit area.

RTN-induced ΔV_{T} distribution is reported to have a long tail, and therefore we cannot use a Normal distribution to model them [6, 9–11]. Literature suggests two distributions to model RTN-induced ΔV_{T} variation. One is the Exponential distribution, $\mathrm{Exp}(\beta)$, and the other is the Lognormal distribution, $\ln\mathcal{N}(\mu, \sigma^2)$. 3-D CAD simulations and device measurements are performed to investigate the scaling effect of the statistical model parameters [6, 9, 10]. Distribution of

A.K.M. Mahfuzul Islam is with the Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, JAPAN (e-mail: mahfuzul@iis.u-tokyo.ac.jp). Tatsuya Nakai and Hidetoshi Onodera are with the Graduate School of Informatics, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, JAPAN (e-mail: {tnakai,onodera}@vlsi.kuee.kyoto-u.ac.jp). Hidetoshi Onodera is also with the JST, CREST.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

single-trap induced ΔV_{T} is reported to follow Exponential distribution [6, 9, 11]. Complex RTN results from multiple traps with different capture and emission times as well as different effects on the channel potential. Overall V_{T} fluctuation due to complex RTN is often modeled as a superposition of multiple single-trap-induced RTNs. Under this assumption, overall ΔV_{T} is obtained by combining the statistics of the number of traps with the statistics of single-trap amplitudes, ΔV_{T} [10]. 95% percentile of overall ΔV_{t} distribution is reported to have $W^{-1}L^{-1.5}$ dependency [10]. However, the assumption, in this case, is that the trap effects are independent of each other which is not the fact as inter-trap interactions are reported to occur [7, 12]. Ref. [7] reports that such an approach largely overestimates the ΔV_{T} distribution. Thus, from the limitations of detecting independent single trap induced V_{T} fluctuations, and the complexity of the inter-trap interactions, it is more reliable to characterize and model the overall ΔV_{T} distribution for circuit performance analysis.

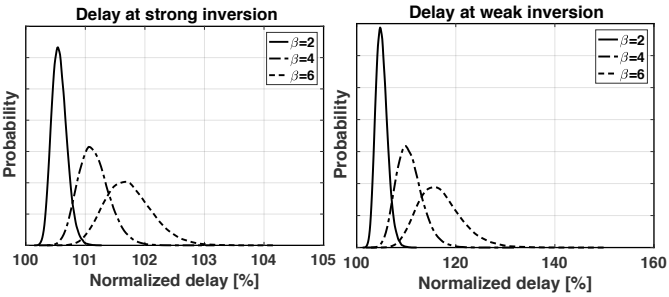
Recently, increase of MOSFET dynamic V_{T} variability with the reduction of rise time of gate voltage is reported [13]. The dynamic variability here is due to low-frequency noise and RTN. The standard deviation of MOSFET dynamic ΔV_{T} is found to follow $1/\sqrt{WL}$. The time-dependent variability reported here suggests that circuits operating at higher switching speed will face higher dynamic variability. Thus, we need to evaluate RTN properties of devices for different switching conditions for accurate modeling. An RO based approach can easily realize different switching speeds by varying the load conditions.

B. Process Variation and RTN

Here, we make a simple comparison between the static process variation and RTN. Static process variation is a result of random dopant fluctuation, line edge roughness, metal granularity, and other process dependent phenomena [14]. Process variation can be global which accounts for total variations across different dies, or local which accounts for mismatches on the same die [15]. Local variation or mismatch in a die can be modeled by a Normal distribution [3, 16]. The Pelgrom model expresses the gate area dependency of the standard deviation of local V_{T} variation as in (2) [17].

$$\sigma_{V_{\mathrm{T,WID}}} = \frac{A_{V_{\mathrm{T}}}}{\sqrt{WL}}, \tag{2}$$

Here, W is gate width, and L is gate length. $A_{V_{\mathrm{T}}}$ is a process dependent parameter. Once $A_{V_{\mathrm{T}}}$ is characterized for a process, the circuit designers use (2) to optimize gate widths and analyze their circuit performance. A statistical analysis can be performed to evaluate the performance spread of a circuit. Process variation is fixed after the manufacture of the transistors, whereas RTN is a time-dependent phenomenon that is caused by the defects in the oxide interface. RTN amplitude is dependent on the defect location, defect energy, channel dopant profile and so on. [5, 18]. As a result, RTN needs to be characterized and modeled for different processes as well as different bias conditions, channel profiles and switching speeds. Characterization of RTN requires time and area as not only a large number of devices has to be measured, but also the devices need to be measured for a long period of time. A digital measurement system would facilitate characterization of RTN. In this paper, we propose an area- and cost-efficient characterization methodology of RTN where the target transistors operate under switching conditions.



(a) Strong inversion. (b) Weak inversion.

Fig. 1: Simulated RTN-induced delay distribution of a delay chain consisting of 20 stages of inverter gates. Three different Exponential distributions with β values of 2, 4 and 6, are assumed for the V_{T} fluctuations.

C. Effect of RTN on Digital Circuit

Compared with an analog circuit, a digital circuit consists of a large number of transistors with different gate lengths and widths. For statistical analysis of circuit delay, distribution of each logic stage needs to be calculated. For a typical digital circuit, the effect of random process variation is averaged out. However, RTN only has a positive fluctuation of V_{T} meaning V_{T} of a device can become higher only. As a result, these effects do not average out. Especially at low voltage operation, where the delay changes exponentially to V_{T} changes, can become a problem. Fig. 1(a) shows simulated delay distributions due to RTN for a 20-stage inverter delay chain. An Exponential distribution is assumed for the RTN-induced ΔV_{T} distribution. Three values of the distribution parameter β are used. β of 2 is observed in our measurement which represents a 65 nm bulk process. β values of 4 and 6 are assumed for finer processes where the channel area shrinks by multiple times. Fig. 1(a) shows the delay distribution when the supply voltage is much higher than the transistor V_{T}. Fig. 1(b) shows the delay distribution when the supply voltage is below transistor V_{T}. At weak inversion operation, RTN-induced delay fluctuation becomes more than 40 % when β is 6. Maximum of 10.4 % of frequency fluctuation is observed in a 7-stage RO for a 40 nm process at a supply voltage of 0.65 V, which supports our observation [19]. There are also reports indicating the increase of RTN with the decrease of gate overdrive voltage [20]. Thus, for low voltage operation, RTN needs to be considered for reliable circuit operation.

In this paper, we investigate the most suitable distributions that represent RTN-induced overall ΔV_{T} distributions when the devices operate in a digital environment. Lognormal and Exponential distributions are compared for the purpose. We also look at the number of traps distributions for different gate widths and fit the distributions to Poisson distributions.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

TABLE I: Key characteristics of and symbols used for different distributions.

	Lognormal	Exponential	Poisson
Parameters	μ, σ	β	λ
PDF of PMF	$\frac{1}{x\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}$	$\frac{1}{\beta}\mathrm{e}^{-\frac{x}{\beta}}$	$\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}$
CDF	$\frac{1}{2}\left[1+\mathrm{erf}\left(\frac{\ln x-\mu}{\sigma\sqrt{2}}\right)\right]$	$1-\mathrm{e}^{-\frac{x}{\beta}}$	$\frac{\Gamma(k+1,\lambda)}{k!}$
Mean	$\mathrm{e}^{\mu+\sigma^2/2}$	β	λ
Variance	$(\mathrm{e}^{\sigma^2}-1)\mathrm{e}^{2\mu+\sigma^2}$	β^2	λ
Sum of same distributions	Approx. Lognormal	Gamma	Poisson
Sum of different distributions	Approx. Lognormal	Not gamma	Poisson

We also look at the correlation between the within-die random variation and RTN. For clarity, The parameters and symbols used in the paper to represent the Lognormal, Exponential and Poisson distributions are summarized in Table I.

D. Contribution of this paper

We show that gate delay variation due to RTN can be related to transistor ΔV_{T} variation using a topology-reconfigurable RO circuit. Using a carefully designed circuit topology, we propose a calibration method to extract ΔV_{T} from delay deviation. We then show an extraction methodology to characterize single transistor ΔV_{T} distribution. In this study, we target weak inversion operation of the target transistors. Cell-based design approach is taken to layout the RO circuit so that the design mimics a conventional digital layout. We perform a comparative analysis of RTN-induced ΔV_{T} variation using both of the Exponential and Lognormal distribution. Then, we evaluate gate area dependency of the model parameters which can be useful in generating distributions for transistors of arbitrary gate widths. We also investigate on the correlations between process variation and RTN.

III. RTN Characterization based on Gate Delay

A. Characterization Methodology

There are previous reports of RTN characterization methodology [6, 9, 10]. A basic approach is to measure transistor I–V characteristics. This method uses g_{m} to convert ΔI_{D} to ΔV_{T}. However, I–V based approach has the following limitations. Firstly, constant bias is applied which is not representative of the switching condition in a logic circuit. The dynamic nature of RTN, however, requires the verification from in-situ characterization under switching conditions. In-situ characterization of RTN effect on SRAM write-instability is reported [21, 22]. Ref. [13] reports a 25% of performance degradation of the static noise margin of high frequency operating CMOS SRAM cell. Ref. [23] reports effects of RTN on sense amplifier performance and finds that some combinations of RTN amplitudes lead to error. In-situ characterization of RTN using a representative digital circuit such as a ring oscillator (RO) thus is very helpful.

There are RO based methodologies for characterization and modeling of process variation [24–26]. ROs are also used for characterizing NBTI [27]. Although RTN measurement results based ring oscillators are reported [19, 28], the challenge lies in converting delay fluctuation to ΔV_{T}. Because RO oscillation period is influenced by RTNs from multiple transistors. Therefore, the RO structure needs to be designed such that the oscillation period is a strong function of a particular transistor. One way to achieve this is to up-size all other transistors except the transistor of concern. However, up-sizing transistors have also its downside that the layout may not be representative of an actual digital circuit. Furthermore, the load capacitance will be large that is also not representative of typical circuit operation. Another way is to reduce the gate–source voltage of the target transistor. Reduced gate voltage can be achieved by inserting a pass-gate transistor before the gate [26]. Pass-gate based design results in compact layout so that transistor characteristics are as close as to those in an actual digital circuit.

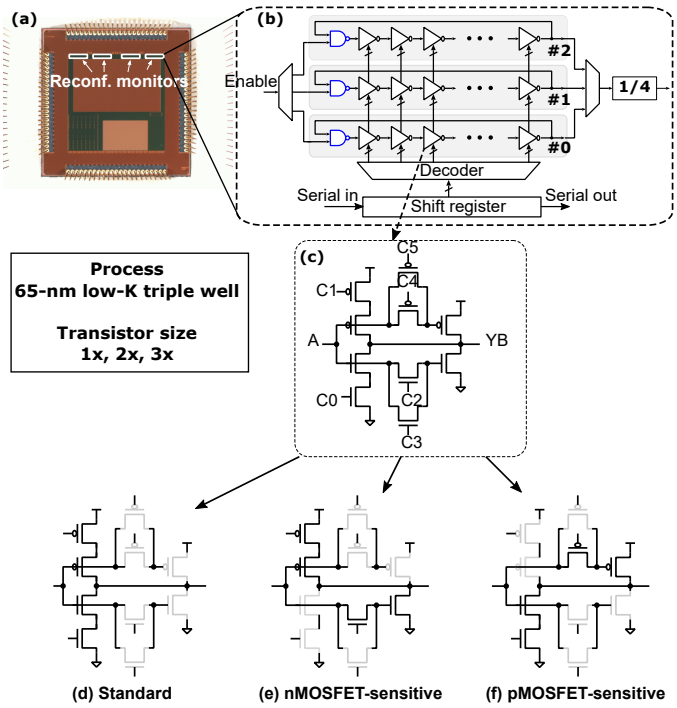


Fig. 2: Topology-reconfigurable ROs for gate delay variation measurement implemented in a 65-nm test chip. ROs with three different gate widths of 120 nm, 240 nm and 360 nm are implemented.

As RTN is a statistical phenomenon, a large number of samples are required. A conventional RO array based approach will consume large area thus reduce sample numbers from a limited area. We, therefore, use a topology-reconfigurable RO based approach as shown in [29]. Using a single RO to measure multiple samples has additional benefits. First, precise calibration can be performed. Second, systematic error due to on-chip location correlated variation is reduced.

B. Circuit Topology

Fig. 2 shows our test chip which comprises four modules with each consisting of three topology-reconfigurable ROs.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

TABLE II: Gate widths used for three ROs.

RO Index	nMOS gate width [nm]	pMOS gate width [nm]
#0	120	120
#1	240	360
#2	360	240

The topology-reconfigurable RO enables gate-level delay evaluation. In the topology, each inverter delay is configured to have multiple times higher sensitivity to V_{T} variation than the other stages. The higher sensitivity is utilized to extract RTN of each inverter stage. Furthermore, making the inverter delay sensitive to only pMOSFET or nMOSFET provides independent pMOSFET and nMOSFET characterizations, respectively.

ROs of three different gate widths are implemented and presented in this paper for demonstration. For detailed understanding and modeling, gate widths with a wider range are preferable. The gate widths for each RO are shown in Table II. Minimum gate length that the process technology allows is used for all the transistors. Each RO is 127 staged. The first stage is a NAND gate to turn ON and OFF the oscillation, and the last stage is used as a buffer. From a single RO, 125 samples for pMOSFET dominated delay, and 125 samples for nMOSFET dominated delay are measured by reconfiguring the topology of each stage. To increase the sample number from a chip, we have implemented four instances of each RO. Thus, we obtain 500 samples for each of the gate widths. Six control signals are used to reconfigure the topology to achieve the following delay characteristics.

1) Equal rise/fall delay as shown in Fig. 2(d),
2) larger fall delay sensitive to nMOSFET as shown in Fig. 2(e), and
3) larger rise delay sensitive to pMOSFET as shown in Fig. 2(f).

C. Gate Delay Evaluation

The following way obtains delay variation between inverter stages. The RO oscillation period when the i-th inverter has a higher fall or rise delay sensitivity is expressed as follows.

$$D^i = D_0^i + d^i. \tag{3}$$

Here, d^i is the inverter fall or rise delay for the i-th stage. D_0^i is the delay contribution from all the inverter delays except the i-th stage fall delay. Thus, the delay of the target i-th stage can be obtained as follows if D_0^i can be estimated.

$$d^i = D^i - D_0^i. \tag{4}$$

Although we can only measure D^i with our circuit, for a sufficiently large number of stages, D_0^i can be approximated with D_0 which is the delay when all the inverter stages are configured as in Fig. 2(d). If D_0 does not show any discrete fluctuation, then the discrete delay fluctuations observed in ΔD^i can be attributed to the i-th inhomogeneous stage. As a result, Δd^i is evaluated by measuring ΔD^i.

$$\Delta d^i \approx \Delta D^i. \tag{5}$$

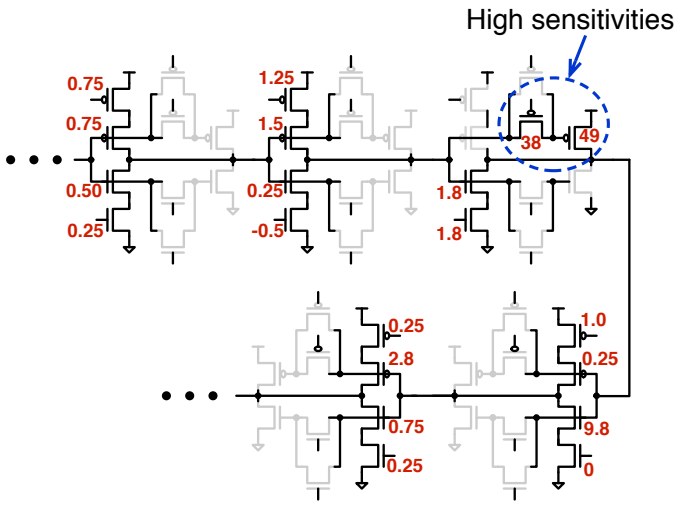


Fig. 3: Sensitivity coefficient values for each transistor.

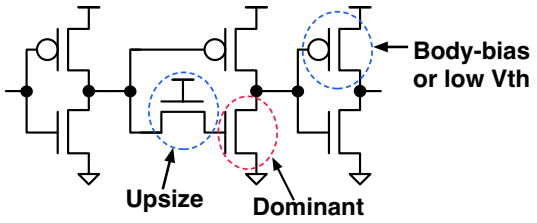


Fig. 4: Tuning of circuit parameters to improve single transistor dominance to the delay variation.

In summary, $(D^i - D_0)$ gives the static process variation for the i-th stage, and ΔD^i over time provides the RTN-induced fluctuation for the same i-th stage.

D. Sensitivity Analysis

Next, we show that by using our proposed gate delay evaluation method RTN of each gate can be characterized accurately. We show sensitivity coefficients of each transistor for a pMOSFET-sensitive inhomogeneous configuration in Fig. 3. Each sensitivity coefficient value represents $\Delta D/\Delta V_{\mathrm{T}}$. The unit of the values is ps per mV. The two pMOS transistors of the inhomogeneous stage have high sensitivity values of 38 and 49, whereas other transistors have sensitivity coefficients less than 2 except for the nMOSFET of the stage next to the inhomogeneous one. The nMOSFET of the next stage have a higher sensitivity of 9.8 which is due to the larger input slew of this gate. In the circuit, there are two pMOSFETs with very high sensitivities and one nMOSFET with relatively higher sensitivity but multiple times smaller compared to the pMOSFETs. In our test circuit, ROs with pMOSFET-sensitive configurations have sensitivity ratios of more than 4 times for gate widths of 120 nm, 240 nm, and 360 nm. However, for the nMOSFET-sensitive configuration, nMOSFET with the gate width of 120 nm, have sensitivity ratio less than 2 times. So, only the gate widths of 240 nm and 360 nm can be characterized in this test circuit.

Various circuit parameters can be employed to make only a single transistor dominant to the delay variation in the RO

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

topology of Fig. 2. Fig. 4 shows a simplified circuit topology with an nMOSFET-sensitive stage as the inhomogeneous stage. The pass-gate transistor is up-sized. pMOSFET of the next stage can be forward body biased or low V_{T} devices can be used. pMOSFET can also be up-sized to reduce its noise. In our demonstration, we choose standard V_{T} devices and similar gate sizing is used for all the transistors. We have used the same cell footprint for all the cells, and the standard cell height is used to mimic the digital circuit environment. Although any arbitrary layout can be used and the design parameters can be tuned further.

E. RTN Detection

RTN is considered to be discrete fluctuations of drain currents due to trapping and de-trapping of charges at the gate oxide interface. The power spectrum of RTN is a Lorentzian or $1/f^2$ spectrum, whereas flicker noise or $1/f$ noise is considered to be a superposition of multiple RTN events [5]. Both the RTN and flicker noise is reported to have a strong gate area dependency. Therefore, characterizing single trap induced fluctuations is difficult in the presence of flicker and other noises. A widely used practice is to detect the discrete fluctuations and characterize the discrete changes as RTN.

Detecting discrete fluctuations from measured samples over time is difficult in the presence of flicker and other noises. Although several detection mechanisms are used in the literature, it is possible that some small discrete fluctuations will be undetected by any method. We illustrate this phenomena by two examples which are shown in Figs. 5 and 6. Background noise is assumed to follow Normal distribution with a standard deviation of 13 ps. In Fig. 5(b), large delay fluctuation due to one trap is considered. As the measured delay distribution is a convolution between the background noise PDF and RTN PDF, the resulted distribution becomes a bimodal PDF as shown in Fig. 5(c). Here, the dotted line shows the convoluted distribution. Open circles refer to the measured distribution from an RO of our test chip. The two distributions overlap each other. A simple peak detection mechanism using a kernel density estimation will suffice for RTN detection in this case. However, when multiple RTN states are very close to each other, as shown in Fig. 6(b), the resulted PDF follows a Normal distribution where no peak exists.

Gaussian Mixture Model (GMM) based on Expectation-Maximization (EM) algorithm is used to find the best PDF [30]. A large number of iterations are performed in this case. Markov chain based technique to filter out the noises is also proposed [31]. Weighted techniques to distinguish RTN states are proposed in Ref. [32]. Here we show that kernel density estimation can be used as an alternative method to detect the peaks in the distribution. In the delay measurement, depending on the measurement time and the time constants of the traps, intermediate delay values between two discrete states are also measured. Figure 7(a) shows one such example. Here, many intermediate delay values are observed which can happen when the time constant is smaller than the measurement time. Using a simple 1-dimensional kernel density based peak detection thus can give many pseudo-RTN states as shown in Fig. 7(c). If we utilize a time lag plot as shown in Fig. 7(b), and concentrate on the $y = x$ line of the 2-dimensional kernel density, the pseudo-RTN states are mostly eliminated which is shown in Fig. 7(d). We, therefore, use a 2-dimensional kernel density estimation, and then use the density on $y = x$ to find the peaks. Ref. [33] also uses a similar approach.

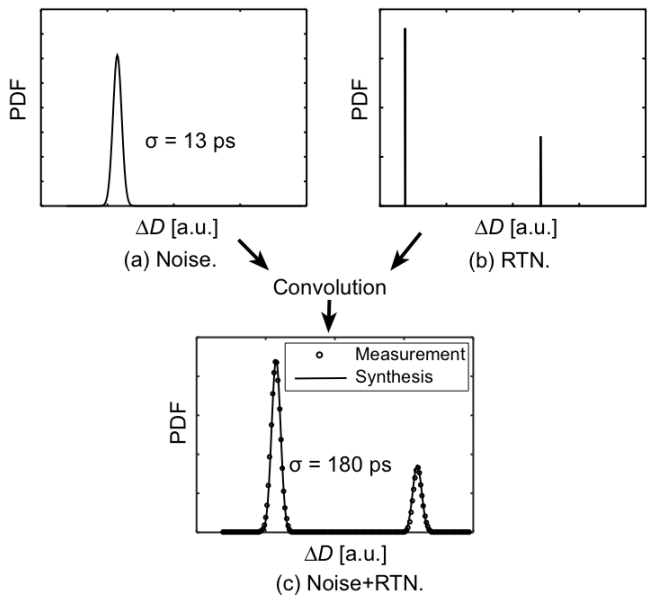


Fig. 5: Delay fluctuation distribution with distinguishable RTN levels. Convolution of noise and RTN are compared with the measurement.

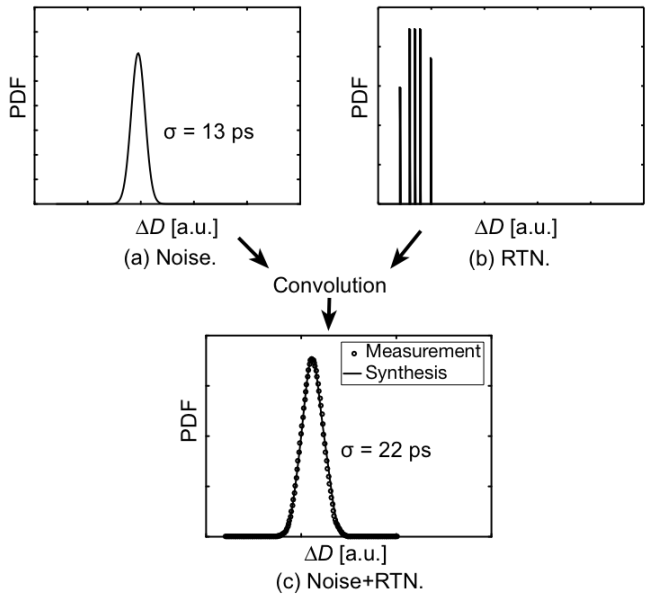


Fig. 6: Delay fluctuation distribution with non-distinguishable RTN levels. Convolution of noise and RTN are compared with the measurement.

F. Delay to Threshold Voltage Conversion

After the discrete fluctuations in the gate delays are detected, the delay fluctuations then need to be converted to V_{T}

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

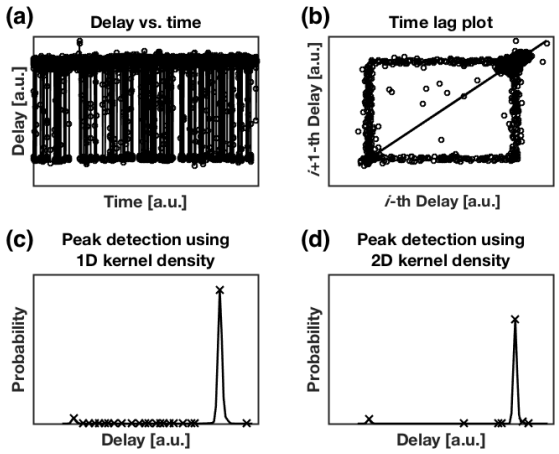


Fig. 7: Detecting RTN with 1-dimensional and 2-dimensional kernel density estimation. (a) Delay versus time plot, (b) time lag plot, (c) 1-dimensional kernel density, (d) kernel density on $y = x$ line of time lag plot.

fluctuations. Using the circuit topology of Fig. 2, we show that $\Delta V_{\rm T}$ of pMOSFET and nMOSFET can be obtained from the gate delay fluctuations. Two transistors mainly contribute the gate delay of pMOSFET- or nMOSFET-sensitive inverter topology. One is the pull-up or the pull-down transistor, and the other is the pass-gate transistor. Delay d for each stage can be modeled by (6) when the pull-up or pull-down transistors operate at weak inversion region [8].

$$d = K \cdot \mathrm{e}^{\alpha \mathrm{V_{T1}} + \beta \mathrm{V_{T2}}}. \tag{6}$$

Here, $V_{\rm T1}$ is the threshold voltage of the pull-up or pull-down transistor, and $V_{\rm T2}$ is the threshold voltage of the corresponding pass-gate transistor. α and β are defined by the sub-threshold, body-effect and DIBL coefficients [29]. Although these coefficients vary randomly in scaled devices, here we consider them as constants.

Next, we show how to extract the WID random process variation of $V_{\rm T}$ between the transistors directly from the gate delay variations. First, we select a particular gate delay as the reference, and divide all other gate delays with the reference delay as shown in (8).

$$\frac{d}{d^0} = \mathrm{e}^{\alpha \Delta \mathrm{V_{T1}} + \beta \Delta \mathrm{V_{T2}}}, \tag{7}$$

$$\Delta V_{\rm T1} + c \cdot \Delta V_{\rm T2} = \frac{1}{\alpha} \cdot \ln\left(\frac{d}{d^0}\right). \tag{8}$$

Here, $c = \beta/\alpha$ and can be considered as a constant as both α and β are constant for a particular process. d is the gate delay of the i-th stage, and d_0 is the delay of any arbitrary gate stage. As discussed in Sec. III, the pass-transistor noise can be made small by up-sizing. Then, $\Delta V_{\rm T2}$ can be ignored, and we obtain $\Delta V_{\rm T1}$ directly from the delay variation.

Next, we show the conversion of Δd due to RTN to $\Delta V_{\rm T}$ for an inhomogeneous stage. First, we normalize RTN-induced delay fluctuation Δd by the gate delay d as shown in (10).

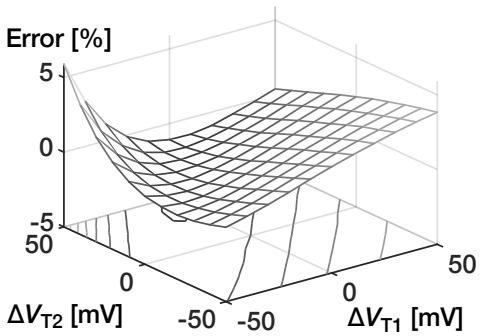


Fig. 8: Error between delay model and simulation for pMOSFET with 240 nm gate width.

$$\frac{\Delta d}{d} = \mathrm{e}^{\alpha \Delta \mathrm{V_{T1}} + \beta \Delta \mathrm{V_{T2}}} - 1, \tag{9}$$

$$\Delta V_{\rm T1} + c \cdot \Delta V_{\rm T2} = \frac{1}{\alpha} \cdot \ln\left(\frac{\Delta d}{d} + 1\right). \tag{10}$$

Thus, we obtain $\Delta V_{\rm T1} + c \cdot \Delta V_{\rm T2}$ for each stage. If the pass-gate width is up-sized by several times, $\Delta V_{\rm T2}$ can be ignored and we obtain $\Delta V_{\rm T1}$ directly for the gate delay fluctuation. As (10) does not depend on the parasitic capacitances and carrier mobility, the method yields accurate conversion of $\Delta V_{\rm T}$. In summary, Δd due to RTN and process variation can be characterized for each inverter stage. The Δd can then be attributed to $\Delta V_{\rm T}$ of a particular transistor.

Next, when pass-gate $\Delta V_{\rm T}$ cannot be ignored, we can estimate $\Delta V_{\rm T}$ distribution from the $\Delta V_{\rm T1} + c \cdot \Delta V_{\rm T2}$ distribution assuming that pass-gate and pull-up(down) transistors have similar $\Delta V_{\rm T}$ distribution. For example, when a Lognormal distribution is assumed for $\Delta V_{\rm T}$, $\Delta V_{\rm T1} + c \cdot \Delta V_{\rm T2}$ can be approximated by another Lognormal distribution. Using Fenton's method [34], $\mu_{\Delta V_{\rm T}}$ and $\sigma^2_{\Delta V_{\rm T}}$ are computed from Eq. (10) with the following equations.

$$\sigma^2_{\Delta V_{\rm T}} = \ln\left[\frac{(1+c)^2 \exp(\sigma^2_{\Delta V_{\rm T1} + c \cdot \Delta V_{\rm T2}}) - 2c}{1 + c^2}\right], \tag{11}$$

$$\mu_{\Delta V_{\rm T}} = \mu_{\Delta V_{\rm T1} + c \cdot \Delta V_{\rm T2}} - \frac{\sigma^2_{\Delta V_{\rm T}}}{2} - \ln(1+c) + \frac{\sigma^2_{\Delta V_{\rm T1} + c \cdot \Delta V_{\rm T2}}}{2}. \tag{12}$$

When an Exponential distribution is assumed for $\Delta V_{\rm T}$ distribution, $\Delta V_{\rm T1} + c \cdot \Delta V_{\rm T2}$ will follow a distribution whose CDF function becomes as follows.

$$F_X(x) = 1 - \frac{c}{c-1}\mathrm{e}^{-\frac{\mathrm{x}}{c\beta}} + \frac{1}{\mathrm{c}-1}\mathrm{e}^{-\frac{\mathrm{x}}{\beta}}. \tag{13}$$

As c is known, fitting (13) with the measured distribution would yield β of the Exponential distribution.

G. Verification

We verify our proposed characterization methodology to extract $\Delta V_{\rm T}$ distributions from ΔD distributions using (10). First, we verify the model equation of (6). In Fig. 8, we show the error between the model (6) and simulation results performed with the foundry provided SPICE transistor model.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

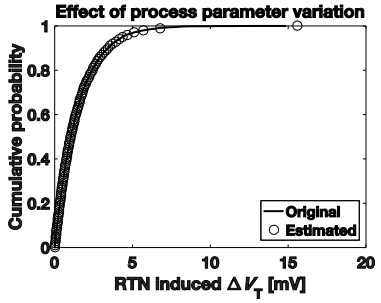


Fig. 9: Effect of variations in sub-threshold swing coefficient values in ΔV_{T} distribution estimation.

In the figure, X-axis indicates the change of V_{T} of the pull-up transistor and Y-axis shows the change of V_{T} for the pass-gate transistor. Pass-gate and pull-up threshold voltages are varied from -50 mV to 50 mV. The Maximum error lies within ± 7 % for all the gate widths.

In (6), we assumed that several process-related parameters are constant. However, in scaled devices, these parameters vary [35]. We show that those variations do not affect the estimation of ΔV_{T} distributions. For simplicity, we consider only ΔV_{T1} in (6) to show the effects of process related parameter variations. Here, mean values of α and β are in the order of 10^1, whereas the mean value of ΔV_{T} is in the order of 10^{-3}. For a product random variable, $Z = XY$, we have the following relationships for the mean and variance using the laws of total expectation and variance, respectively [36].

$$\mathrm{E}(XY) = \mathrm{E}(X) \cdot \mathrm{E}(Y), \tag{14}$$

$$\mathrm{Var}(XY) = \mathrm{Var}(X) \cdot \mathrm{Var}(Y) + \mathrm{Var}(X) \cdot \mathrm{E}(Y)^2 + \mathrm{Var}(Y) \cdot \mathrm{E}(X)^2. \tag{15}$$

Considering $X = \alpha$ and $Y = \Delta V_{\mathrm{T}}$, we find in (15) that the first and the second term on the right side can be ignored even we consider a 10 % variability for the α parameter. Thus, the mean and the variance of ΔV_{T} can be used to approximate the product distribution with small error. This is verified by performing Monte Carlo simulation. The result is shown in Fig. 9. In the figure, the solid line shows the original ΔV_{T} distribution used in the Monte Carlo simulation, and the open circles show the estimated ΔV_{T} distribution under the presence of 10 % variability in the α parameter. The distributions are almost identical. Thus, we conclude that the extraction of ΔV_{T} distributions based on gate delay measurement with a topology-reconfigurable RO is valid.

Finally, contributions of transistors in the inhomogeneous stage on the delay distribution is verified taking a pMOSFET-sensitive configuration as an example. Fig. 10 shows the contributions of pMOSFETs and nMOSFETs in a quantile–quantile plot. The X-axis shows the quantile of delay fluctuation due to RTN, and Y-axis shows the standard normal quantile. The bigger dotted line shows delay distribution due to RTN in the nMOSFET of the following stage after the inhomogeneous one. The smaller dotted line shows delay distribution due to RTN in the pMOSFETs in the inhomogeneous stage.

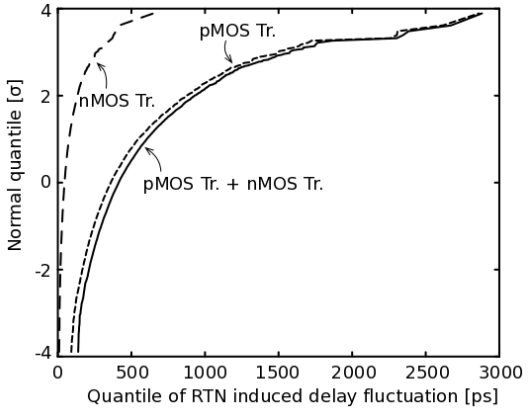


Fig. 10: RTN contribution of each transistor.

The solid line shows overall delay distribution. Above 2σ, overall delay distribution and distribution due to pMOSFETs converge. The difference can be further minimized by upsizing the nMOSFET. Thus, while characterizing pMOSFET, the error due to the higher sensitivity in the nMOSFET of the next stage is considered small.

IV. Measurement and Analyses

A. Measurement Setup

To evaluate gate delay operating at weak inversion region, a supply voltage of 0.8 V is used. Because of the V_{T} drop across the pass-gate transistor, the gate overdrive for the pull-up or pull-down transistor in the inhomogeneous stage is in weak inversion. For each inhomogeneous configuration, only the devices under test operate at weak inversion, and all the other devices operate above at strong inversion. Four ROs are measured for both of the pMOSFET and nMOSFET V_{T} characterizations. As a result, a total of $125 \times 4 = 500$ measurement samples is obtained for each gate width of pMOSFET and nMOSFET. Control signals are generated inside the chip and the output frequency is measured outside the chip using a frequency counter implemented in an FPGA. For each RO, the oscillation frequency is measured over 35 s with an integration time of 1 ms. The RO frequency is divided by 16 inside the chip so that the outside frequency at around 4 MHz. The signal is then measured by a clock signal of 24 MHz with an integration time of 1 ms. Thus, a single clock count corresponds to $(1/N_{\mathrm{count}}) \times (T_{\mathrm{clk}}/N_{\mathrm{div}}) \approx 0.65$ ps of delay change. Here, N_{count} is the number of count of the oscillation signal in the FPGA, and N_{div} is the number of divider in the test chip. T_{clk} is the clock period to measure the oscillation frequency.

B. Process Variation

First, we confirm that whether our circuit is capable of measuring V_{T} variation by estimating the process variation. Fig. 11 shows the estimated ΔV_{T} variations due to local process variation. Gate area dependency is observed to follow the Pelgrom model which suggests that characterization based on delay model (6) is valid.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

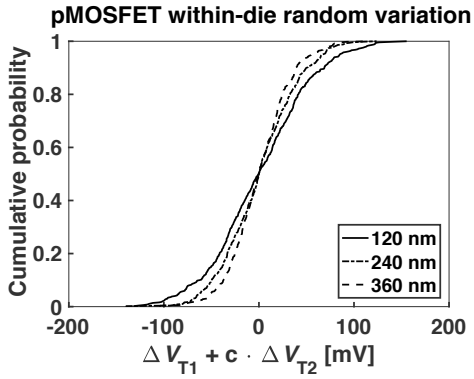


Fig. 11: Cumulative probability plot for static process variation of pMOSFET for three different gate widths.

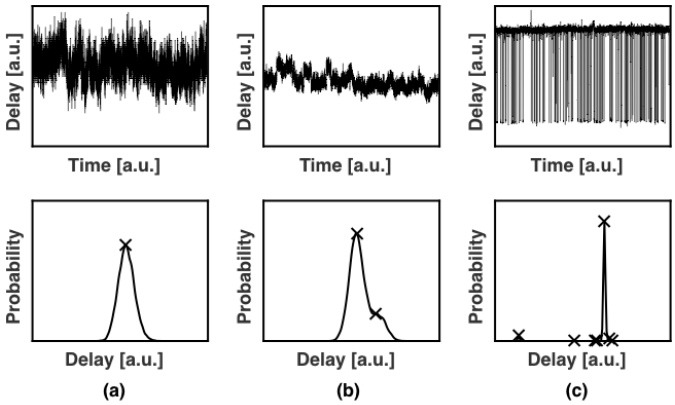


Fig. 12: Samples with different RTN profiles.

C. RTN Measurements

Discrete gate delay fluctuations are then measured for each gate width to evaluate RTN caused by pMOSFET and nMOSFET separately. Discrete states are detected using the method presented in Sec. III. Several measurement samples for pMOSFET RTN are shown in Fig. 12. Fig. 12 shows measurement samples of oscillation period observed over time for three different inhomogeneous stages. Fig. 12(a) shows a sample with probability density having a single peak. We term this kind of samples as non-detectable samples rather than treating them as zero. Samples that show only one peak in the distribution refers that their discrete fluctuations cannot be detected. In our evaluation, the numbers of valid samples showing at least two peaks are 371, 381, and 427 for pMOSFET gate widths of 120 nm, 240 nm, and 360 nm respectively. For nMOSFET, the numbers with at least two peaks are 397 and 353 for nMOSFET gate widths of 240 nm and 360 nm. Fig. 12(b) shows a sample with probability density having two distinguishable peaks. In Fig. 12(c), samples with multiple peaks are shown. With the proposed circuit and characterization system, we observe RTN of various amplitudes, time constants and a different number of traps.

D. RTN Distributions

We show the distributions of delay fluctuations for all the gate widths of pMOSFET and nMOSFET. Fig. 13 shows $\Delta d/d$ distributions due to RTN in the pMOSFETs. Distributions of three different gate widths are shown in the figure. As expected, long tails are observed in all the three distributions. We observe an RTN-induced delay fluctuation of 40 % for a gate width of 120 nm. Fig. 14 shows RTN-induced $\Delta d/d$ distributions for two different gate widths of nMOSFET. Long tails are observed for the distributions. pMOSFET and nMOSFET of the same gate width give roughly similar delay fluctuations in this process.

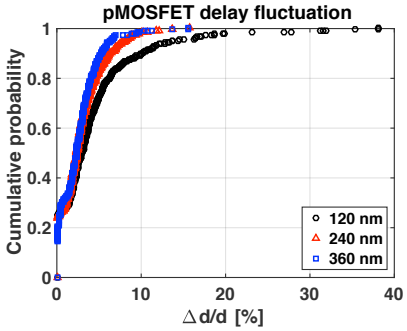


Fig. 13: Cumulative probability of detected RTN-induced $\Delta d/d$ of pMOSFET dominant gates.

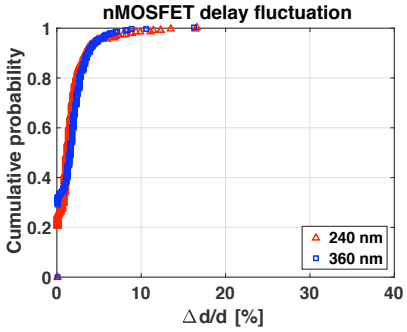
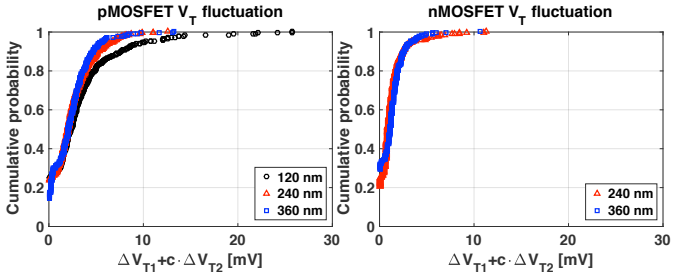


Fig. 14: Cumulative probability of detected RTN-induced $\Delta d/d$ of nMOSFET dominant gates.



(a) pMOSFET. (b) nMOSFET.

Fig. 15: Cumulative probability of $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ for pMOSFET and nMOSFET.

Next, the RTN-induced $\Delta d/d$ distributions are converted to $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ distributions using (10). $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ distribution here represents the sum of two ΔV_{T} distributions. Fig. 15(a) shows the $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ distributions for pMOSFET with gate widths of 120 nm, 240 nm, and 360 nm. 120 nm pMOSFET show 99 % quantile value of 22 mV Increasing the gate width by twice reduces the value to 9.6 mV.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

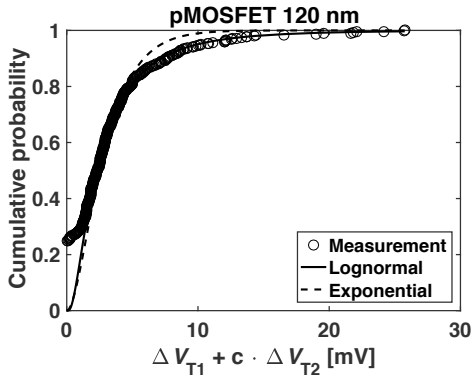
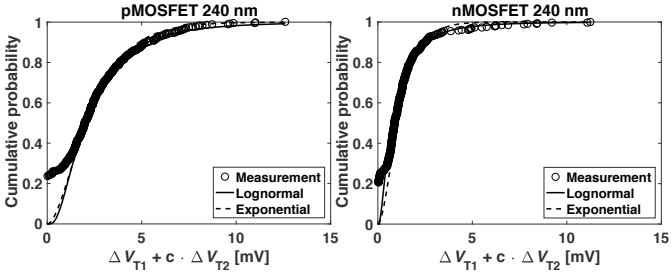


Fig. 16: Cumulative probability of $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ for pMOSFET of 120 nm gate width.



(a) pMOSFET. (b) nMOSFET.

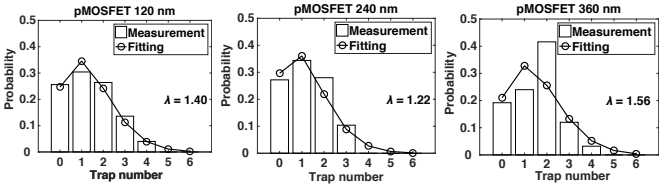
Fig. 17: Cumulative probability of $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ for pMOSFET and nMOSFET of gate width of 240 nm.

However, further increasing the gate width by 1.5 times gives a value of 9.0 mV. Fig. 15(b) shows the $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ distributions for nMOSFET with gate widths of 240 nm, and 360 nm. 99 % quantile values are 7.9 mV and 5.6 mV for gate widths of 240 nm and 360 nm. Thus, for nMOSFET increasing the gate width by 1.5 times decreases V_{T} fluctuation by 1.4 times.

Fig. 16 shows comparisons between Lognormal and Exponential fitting for $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ distribution of pMOSFET a with gate width of 120 nm. Solid and dotted lines show the fitting with a Lognormal and an Exponential distribution. We find that Lognormal model is better at representing RTN-induced overall V_{T} fluctuations. $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ distributions for pMOSFET and nMOSFET with the same gate width of 240 nm are shown in Fig. 17 along with the fitted curves. Both the Lognormal and the Exponential distributions fit the data well with Lognormal being slightly better at 240 nm. So, overall we observe that Lognormal distribution is better at representing RTN-induced overall V_{T} fluctuations for smaller channel areas.

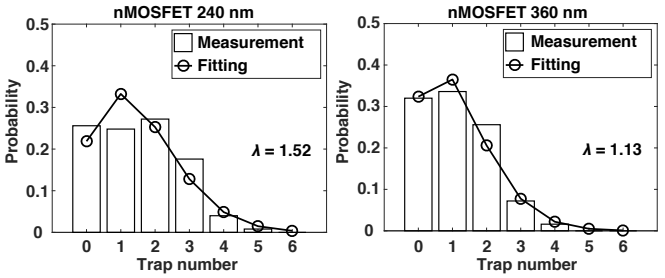
E. Number of traps

We show the distributions of number of detected traps for all the gate widths of pMOSFET and nMOSFET. In theory, number of traps is believed to follow a Poisson distribution. Sum of two Poisson distributions is also a Poisson distribution and the parameter λ of the resulted Poisson distribution is the sum of λs of the underlying distributions. In this paper,



(a) $W = 120$ nm. (b) $W = 240$ nm. (c) $W = 360$ nm.

Fig. 18: Number of traps for pMOSFETs of three gate widths of 120 nm, 240 nm and 360 nm.



(a) $W = 240$ nm. (b) $W = 360$ nm.

Fig. 19: Number of traps for nMOSFETs of 240 nm and 360 nm gate widths.

we evaluate $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ distributions where RTN states are calculated by the peaks in the kernel density. Then, trap number N is calculated as follows.

$$N = \mathrm{ceil}(\log_2 \mathrm{n}), \tag{16}$$

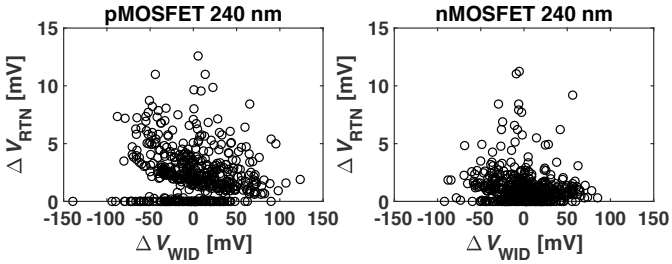
where n is the number of discrete states. λ of trap number distribution for ΔV_{T} can be approximated by dividing the λ of $\Delta V_{\mathrm{T1}} + c \cdot \Delta V_{\mathrm{T2}}$ by 2.

Fig. 18 shows the distributions of trap numbers for three different gate widths of pMOSFET. Fitted Poisson distribution is also shown in the plots. Poisson distribution fits the data well. Although λ should increase with the increase of channel area theoretically. ΔV_{T} amplitude per trap also decreases with the increase of channel area. Smaller amplitude causes smaller discrete fluctuations that remain undetected. Furthermore, a larger number of traps will result in the flicker noise as shown in Fig. 6 where only a single peak exists in the kernel density. As a result, the number of traps that can be detected depends on both the RTN amplitude and the channel area. In our characterization system, we find that the trap numbers are not monotonic with the gate area for the pMOSFET. Fig. 19 shows the distributions for two sizes of nMOSFET where larger gate area has smaller trap numbers. For further investigation, devices with a wider range in gate width need to be characterized.

F. Correlation between Process Variation and RTN

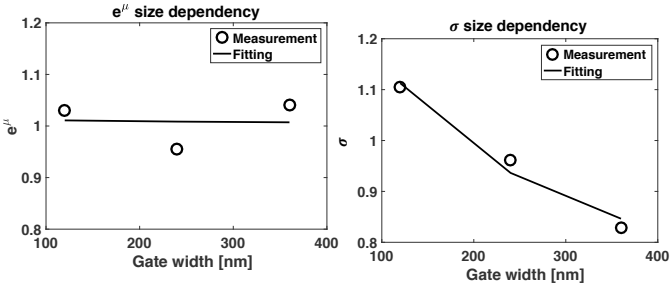
One key phenomenon of interest is the correlation between process variation and RTN. Strong correlation between them means that worst-case delay degradation will be severe. RTN-induced overall amplitude against process variation induced V_{T} fluctuation is plotted for each inhomogeneous stage. The

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing



(a) pMOSFET 240 nm. (b) nMOSFET 240 nm.

Fig. 20: Correlation between static process variation and RTN.



(a) $\hat{\mu} = \mathrm{e}^{\mu}$ of Lognormal distribution. (b) σ of Lognormal distribution.

Fig. 21: Gate area dependency of μ and σ parameters of ΔV_{T} distribution for lognormal distribution.

correlation for pMOSFET and nMOSFET of the gate width 240 nm is shown in Fig. 20. Figures show no correlation between the two variations. Literature reports also agree with our observation. So, for the heavily doped bulk process of 65 nm process, where process variation is strongly dominated by random dopant fluctuation, no correlation has been observed between RTN and process variation.

G. Gate area Dependency

We calculated the model parameters of μ and σ for the Lognormal distribution, and β for the Exponential distribution. Extracted μ and σ values are shown in Figs. 21(a) and 21(b). μ shows very small gate area dependency where the fitted model results in $1.2/W^{0.03}$. However, σ shows stronger dependency which fits the model $4.0/W^{0.26}$. β of Exponential distribution against each gate width is shown in Fig. 22. Fig. 22 shows the β values of the Exponential distribution against the gate widths of pMOSFET. β shows a size dependency of $4.7/W^{0.21}$ which is much smaller than expected. The reason for the smaller gate area dependency is from the fact that MOSFETs with gate widths of 240 nm and 360 nm show similar ΔV_{T} distributions due to RTN. Characterizing RTN for gate widths of a wider range will give us detailed understanding on gate area dependency. In this test chip, we have not implemented transistors with larger gate widths. The main focus of this paper is to show that delay based MOSFET RTN characterization is feasible.

V. Conclusion

Gate delay based detailed RTN characterization is feasible. A topology-reconfigurable RO circuit facilitates delay to V_{T} conversion to evaluate RTN of pMOSFET and nMOSFET separately. The test circuit can be designed in a conventional cell-based design so that the layouts and operating conditions of the transistors are as close as to those in digital circuits. Reconfigurability reduces the circuit area significantly. Detailed RTN measurement results obtained from a 65 nm test chip are presented. The Lognormal distribution is found to fit the RTN-induced overall V_{T} fluctuation better than the Exponential distribution. We also find that for the same gate area, pMOSFET and nMOSFET show similar V_{T} fluctuations due to RTN. Furthermore, no correlation has been found between RDF induced V_{T} variation and RTN. Finally, RTNs for different gate widths are compared. The evaluated results do not give meaningful gate area dependency, partly because of the closeness of gate widths used in the test circuit. Implementing the circuit with a wider range of gate widths will allow deeper understanding of the gate area dependency.

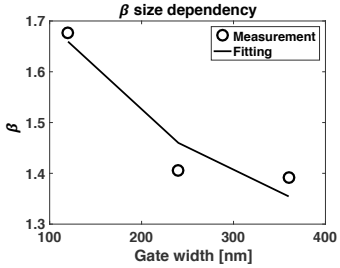


Fig. 22: β of Exponential distribution against gate width.

Delay variation due to RTN in the 65 nm process is found to have no serious impact on circuit performance. However, shrinking the gate area by 3 times may result in more than 40 % of delay fluctuation at the weak inversion operation. The characterization of overall ΔV_{T} distribution for different gate widths of pMOSFET and nMOSFET can be used to establish a compact statistical model that can be used for reliability analysis and optimization of digital circuits.

Acknowledgement

The VLSI chip in this study has been fabricated in the chip fabrication program of VDEC, the University of Tokyo in collaboration with STARC, e-Shuttle, Inc., and Fujitsu Ltd. This work was partly supported by JSPS KAKENHI numbers 25280014 and 16H01713.

References

[1] S. Borkar, "Design Perspectives on 22nm CMOS and Beyond," in *Design Automation Conference*, 2009, pp. 93–94.

[2] S. Jain, S. Khare, S. Yada, P. Salihundam, S. Ramani, S. Muthukumar, A. Kumar, S. K. Gb, R. Ramanarayanan, V. Erraguntla, J. Howard, S. Vangal, S. Dighe, G. Ruhl, P. Aseron, H. Wilson, N. Borkar, V. De, and S. Borkar, "A 280mV-to-1.2V Wide-Operating-Range IA-32 Processor in 32nm CMOS," in *International Solid State Circuits Conference*, 2012, pp. 202–203.

[3] 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*, 2008, pp. 156–157.

[4] N. Tega, H. Miki, and F. Pagette, "Increasing Threshold Voltage Variation due to Random Telegraph Noise in FETs as Gate Lengths Scale to 20 nm," in *Symposium on VLSI Technology*, 2009, pp. 50–51.

[5] M. Kirton and M. Uren, "Noise in solid-state microstructures: a new perspective on individual defects, interface states, and low-frequency noise," *Advances in Physics*, vol. 38, no. 4, pp. 367–468, 1989.

[6] K. Takeuchi, T. Nagumo, S. Yokogawa, K. Imai, and Y. Hayashi, "Single-Charge-Based Modeling of Transistor Characteristics Fluctuations Based on Statistical Measurement of RTN Amplitude," in *Symposium on VLSI Technology*, 2009, pp. 54–55.

[7] J. Zou, R. Wang, S. Guo, M. Luo, Z. Yu, X. Jiang, P. Ren, J. Wang, J. Liu, J. Wu, W. Wong, S. Yu, H. Wu, S. W. Lee, Y. Wang, and R. Huang, "New understanding of state-loss in complex RTN: Statistical experimental study, trap interaction models, and impact on circuits," in *IEEE International Electron Devices Meeting*, 2015, pp. 34.5.1–34.5.4.

[8] A. K. M. M. Islam, T. Nakai, and H. Onodera, "Statistical Analysis and Modeling of Random Telegraph Noise Based on Gate Delay Variation Measurement," in *International Conference on Microelectronic Test Structures*, 2016, pp. 82–87.

[9] A. Ghetti, C. Compagnoni, A. Spinelli, and A. Visconti, "Comprehensive Analysis of Random Telegraph Noise Instability and Its Scaling in Deca-Nanometer Flash Memories," *IEEE Transactions on Electron Devices*, vol. 56, no. 8, pp. 1746–1752, 2009.

[10] S. Realov and K. L. Shepard, "Analysis of Random Telegraph Noise in 45-nm CMOS Using On-Chip Characterization System," *IEEE Transactions on Electron Devices*, vol. 60, no. 5, pp. 1716–1722, may 2013.

[11] J. Franco, B. Kaczer, M. Toledano-Luque, P. J. Roussel, J. Mitard, L. Å. Ragnarsson, L. Witters, T. Chiarella, M. Togo, N. Horiguchi, G. Groeseneken, M. F. Bukhori, T. Grasser, and A. Asenov, "Impact of single charged gate oxide defects on the performance and scaling of nanoscaled FETs," in *IEEE International Reliability Physics Symposium*, 2012, pp. 1–6.

[12] P. Ren, P. Hao, C. Liu, R. Wang, X. Jiang, Y. Qiu, R. Huang, S. Guo, M. Luo, J. Zou, M. Li, J. Wang, J. Wu, J. Liu, W. Bu, W. Wong, S. Yu, H. Wu, S. W. Lee, and Y. Wang, "New observations on complex RTN in scaled high-κ/metal-gate MOSFETs - The role of defect coupling under DC/AC condition," in *IEEE International Electron Devices Meeting*, 2013, pp. 778–781.

[13] E. G. Ioannidis, S. Haendler, C. G. Theodorou, N. Planes, C. A. Dimitriadis, and G. Ghibaudo, "Statistical analysis of dynamic variability in 28nm FD-SOI MOSFETs," in *IEEE European Solid-State Device Research Conference*, 2014, pp. 214–217.

[14] X. Wang, A. R. Brown, and A. Asenov, "Statistical Variability and Reliability in Nanoscale FinFETs," in *IEEE International Electron Devices Meeting*, dec 2011, pp. 5.4.1–5.4.4.

[15] K. R. Lakshmikumar, R. A. Hadaway, and M. A. Copeland, "Characterization and Modeling of Mismatch in MOS Transistors for Precision Analog Design," *IEEE Journal of Solid-State Circuits*, vol. sc-21, no. 6, pp. 1057–1066, 1986.

[16] M. D. Giles, N. A. 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 *IEEE Symposium on VLSI Technology*, 2015, pp. 150–151.

[17] M. Pelgrom, A. Duinmaijer, and A. Welbers, "Matching Properties of MOS Transistors," *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, oct 1989.

[18] J. Nishimura, T. Saraya, and T. Hiramoto, "Statistical comparison of random telegraph noise (RTN) in bulk and fully depleted SOI MOSFETs," in *International Conference on Ultimate Integration on Silicon*, mar 2011, pp. 1–4.

[19] T. Matsumoto, K. Kobayashi, and H. Onodera, "Impact of Random Telegraph Noise on CMOS Logic Delay Uncertainty under Low Voltage Operation," in *International Electron Devices Meeting*, dec 2012, pp. 25.6.1–25.6.4.

[20] J. P. Campbell, L. C. Yu, K. P. Cheung, J. Qin, J. S. Suehle, A. Oates, and K. Sheng, "Large random telegraph noise in sub-threshold operation of nano-scale nMOSFETs," in *IEEE International Conference on Integrated Circuit Design and Technology*, 2009, pp. 17–20.

[21] 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," in *IEEE Symposium on VLSI Technology*, 2015, pp. 38–39.

[22] T. Mizutani, T. Saraya, K. Takeuchi, M. Kobayashi, and T. Hiramoto, "Transistor-level characterization of static random access memory bit failures induced by random telegraph noise," *Japanese Journal of Applied Physics*, vol. 55, no. 4S, p. 04ED05, 2016.

[23] J. Martin-Martinez, J. Diaz, R. Rodriguez, M. Nafria, X. Aymerich, E. Roca, F. V. Fernandez, and A. Rubio, "Characterization of random telegraph noise and its impact on reliability of SRAM sense amplifiers," in *2014 5th European Workshop on CMOS Variability (VARI)*, 2014, pp. 1–6.

[24] M. Bhushan, A. Gattiker, M. B. Ketchen, and K. K. Das, "Ring Oscillators for CMOS Process Tuning and Variability Control," *IEEE Transactions on Semiconductor Manufacturing*, vol. 19, no. 1, pp. 10–18, 2006.

[25] A. K. M. M. Islam, A. Tsuchiya, K. Kobayashi, and H. Onodera, "Variation-sensitive Monitor Circuits for Estimation of Global Process Parameter Variation," *IEEE Transactions on Semiconductor Manufacturing*, vol. 25, no. 4, pp. 571–580, 2012.

[26] S. Fujimoto, A. K. M. M. Islam, T. Matsumoto, and H. Onodera, "Inhomogeneous Ring Oscillator for Within-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSM.2017.2715168, IEEE Transactions on Semiconductor Manufacturing

Die Variability and RTN Characterization," *IEEE Transactions on Semiconductor Manufacturing*, vol. 26, no. 3, pp. 296–305, 2013.

[27] M. Ketchen, M. Bhushan, and R. Bolam, "Ring Oscillator Based Test Structure for NBTI Analysis," in *IEEE International Conference on Microelectronic Test Structures*, 2007, pp. 42–47.

[28] T. Matsumoto, K. Kobayashi, and H. Onodera, "Impact of random telegraph noise on CMOS logic circuit reliability," in *IEEE Custom Integrated Circuits Conference*. Ieee, sep 2014, pp. 1–8.

[29] A. K. M. M. Islam and H. Onodera, "Sensitivity-independent Extraction of Vth Variation Utilizing Log-normal Delay Distribution," in *International Conference on Microelectronic Test Structures*, 2015, pp. 212–217.

[30] Z. Zhang, S. Guo, X. Jiang, R. Wang, and R. Huang, "Investigation on the Amplitude Distribution of Random Telegraph Noise (RTN) in Nanoscale MOS Devices," in *IEEE International Nanoelectronics Conference*, 2016, pp. 5–6.

[31] H. Awano, H. Tsutsui, H. Ochi, and T. Sato, "Bayesian Estimation of Multi-Trap RTN Parameters Using Markov Chain Monte Carlo Method," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E95.A, no. 12, pp. 2272–2283, 2012.

[32] J. Martin-Martinez, J. Diaz, R. Rodriguez, M. Nafria, and X. Aymerich, "New weighted time lag method for the analysis of random telegraph signals," *IEEE Electron Device Letters*, vol. 35, no. 4, pp. 479–481, 2014.

[33] S. Dongaonkar, M. D. Giles, A. Kornfeld, B. Grossnickle, and J. Yoon, "Random Telegraph Noise (RTN) in 14nm Logic Technology : High Volume Data Extraction and Analysis," in *IEEE Symposium on VLSI Technology*, 2016, pp. 176–177.

[34] L. F. Fenton, "The Sum of Log-Normal Probability Distributions in Scatter Transmission Systems," *IRE Transactions on Communications Systems*, vol. 8, no. 1, pp. 57–67, 1960.

[35] T. Sato, H. Ueyama, N. Nakayama, and K. Masu, "Accurate Array-Based Measurement for Subthreshold-Current of MOS Transistors," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 11, pp. 2977–2986, nov 2009.

[36] N. A. Weiss, *A Course in Probability*, 2005.

A.K.M. Mahfuzul Islam received the B.E. degree in electrical and electronics engineering, the M.E. degree in communications and computer engineering, and the Ph.D. degree in informatics, all from Kyoto University, Kyoto, Japan, in 2009, 2011, and 2014, respectively.

He has been a Research Fellow of Japan Society for the Promotion of Science from 2013 to 2015. Since 2015, he has been a Research Associate at the Institute of Industrial Science, the University of Tokyo. His current research interests include development of low-cost and low-power on-chip sensor circuits, transistor variability characterization and modeling, and low-power design techniques.

Tatsuya Nakai received the B.E. degree in electrical and electronics engineering in 2015, and the M.E. degree communications and computer engineering in 2017, all from Kyoto University, Kyoto, Japan. His research interests include transistor noise characterization and modeling.

Hidetoshi Onodera received the B.E., and M.E., and Dr. Eng. degrees in Electronic Engineering, all from Kyoto University, Kyoto, Japan. He joined the Department of Electronics, Kyoto University, in 1983, and currently a Professor in the Department of Communications and Computer Engineering, Graduate School of Informatics, Kyoto University. His research interests include design technologies for Digital, Analog, and RF LSIs, with particular emphasis on low-power design, design for manufacturability, and design for dependability.

Dr. Onodera served as a Program Chair and a General Chair of ICCAD and ASP-DAC. He was a Chairman of the IPSJ SIG-SLDM (System LSI Design Methodology), the IEICE Technical Group on VLSI Design Technologies, the IEEE SSCS Kansai Chapter, the IEEE CASS Kansai Chapter, and IEEE Kansai Section. He served as an Editor-in-Chief of IEICE Transactions on Electronics and IPSJ Transactions on System LSI Design methodology.