Characterizing and Modelling RTN under real circuit bias conditions

Kean H. Tok, Jian F. Zhang, James Brown, Zhigang Ji, Weidong Zhang, and John S. Marsland

***Abstract*— Random Telegraph Noise (RTN) has attracted much attention, as it becomes higher for smaller devices. Early works focused on RTN in linear drain current, $I_{D,LIN}$, and there is only limited information on RTN in saturation current, $I_{D,SAT}$. As transistors can operate in either linear or saturation modes, lack of RTN model in $I_{D,SAT}$ prevents modelling RTN for real circuit operation. Moreover, circuit simulation requires both driving current and threshold voltage, V_{TH}. A common practice of early works is to evaluate the RTN in V_{TH} by $\Delta V_{TH}=\Delta I_{D,LIN}/gm$, where gm is transconductance. It has been reported that the ΔV_{TH} evaluated in this way significantly overestimates the real ΔV_{TH}, but there is little data for establishing the cumulative distribution function (CDF) of the real ΔV_{TH}. An open question is whether ΔV_{TH} and $\Delta I_{D,LIN}/I_{D,LIN}$ follow the same CDF. The objectives of this work are three-fold: to provide statistical test data for RTN in $I_{D,SAT}$; to measure the RTN in real ΔV_{TH} by pulse I_D-V_G; and, for the first time, to apply the integral methodology for developing the CDF per trap for all four key parameters needed by circuit simulation– $\Delta I_{D,LIN}/I_{D,LIN}$, $\Delta I_{D,SAT}/I_{D,SAT}$, $\Delta V_{TH,LIN}$, and $\Delta V_{TH,SAT}$. It is found that the Log-normal CDF is the best for $\Delta I_{D,LIN}/I_{D,LIN}$ and $\Delta I_{D,SAT}/I_{D,SAT}$, while the General Extreme Value CDF is the best for $\Delta V_{TH,LIN}$ and $\Delta V_{TH,SAT}$. Both $\Delta I_{D,SAT}/I_{D,SAT}$ and $\Delta V_{TH,SAT}$ are higher than their linear counterparts and separate modelling is required. Finally, the applicability of integral methodology in predicting the long term $\Delta I_{D,LIN}/I_{D,LIN}$ is demonstrated.**

***Index Terms*— Noise, Random telegraph noise (RTN), Jitters, Yield, Fluctuation, Device Variations, Time-dependent Variations.**

I. Introduction

Device-to-device variation (DDV) is a major challenge for designing modern CMOS circuits [1-27]. There are two types of DDV: time invariant DDV, such as line edge roughness [1], and time-dependent variation (TDV), such as age-induced DDV [2]. As transistor sizes downscale, random telegraph noise (RTN) has become one of the main sources of TDV [3-27]. A single trap in gate dielectric can cause substantial fluctuation of both driving current, ΔI_D, and threshold voltage, ΔV_{TH}, by capturing/emitting a charge carrier from/to the conduction channel [2-27]. The number of traps in a device follows the Poisson distribution [16,17] and RTN induces substantial stochastic TDV [3-27].

Despite of the efforts made by early works [3-27], accurately modelling and predicting RTN for circuit optimization remains a challenge for modern CMOS technologies and there are knowledge gaps that prevent achieving the optimization. One of them is that early works focused on $\Delta I_D/I_D$ under linear operation condition, i.e. $(V_G-V_{TH})>V_D$. Real circuits, however, can operate under not only linear, but also saturation modes. For example, Fig. 1 shows that the transistor M1 and M5 in a standard SRAM cell can operate in linear and saturation mode, respectively, during a read operation. The limited early works on RTN under saturation [18-21] mainly investigated the impact of drain bias on individual traps and the key information, such as the CDF per trap for $\Delta I_{D,SAT}/I_{D,SAT}$ and $\Delta V_{TH,SAT}$, is still missing. Without it, the simulation of RTN for real circuit operation cannot be completed.

In addition to driving current, circuit simulation also requires threshold voltage. Under linear mode, one common practice of early works [12, 22-24] is to evaluate the RTN in $V_{TH,LIN}$ by $\Delta V_{TH,LIN}=\Delta I_{D,LIN}/gm$, where $\Delta I_{D,LIN}$ is the fluctuation of linear I_D and gm the transconductance. When both $\Delta I_{D,LIN}$ and gm were measured at $V_G\approx V_{TH}$, $\Delta V_{TH,LIN}=\Delta I_{D,LIN}/gm$ is a good estimation of the real ΔV_{TH} [16],[25]. Most of circuits and early works, however, use an operation voltage of $V_G>V_{TH}$. Under $V_G>V_{TH}$, it has been reported that $\Delta V_{TH,LIN}=\Delta I_{D,LIN}/gm$ substantially overestimates the real $\Delta V_{TH,LIN}$ directly measured from the pulse I_D-V_G at $V_G\approx V_{TH}$ [25].

The impact of a trapped charge on device depends on the local charge carrier density beneath it [3,5,15,25]. Under $V_G\approx V_{TH}$, current flows through narrow paths and traps located away from these paths have small impact. In contrast, under $V_G>V_{TH}$, current distribution is more uniform and more traps can be charged and have large contribution to RTN. This explains why $\Delta V_{TH,LIN}=\Delta I_{D,LIN}/gm$ measured under $V_G>V_{TH}$ is larger than the $\Delta V_{TH,LIN}$ measured at $V_G\approx V_{TH}$. At present, there are insufficient data to establish the CDF of real $\Delta V_{TH,LIN}$ and there are even less data on the directly measured $\Delta V_{TH,SAT}$. As the current flow pattern under $V_G\approx V_{TH}$ is different from that under $V_G>V_{TH}$, one open question is whether the RTN in driving current and V_{TH} follows the same CDF.

The objectives of this work are:

- To provide experimental data for RTN under saturation condition;
- To directly measure ΔV_{TH} at $V_G\approx V_{TH}$ and obtain its statistical data;

Manuscript received MM DD, YYYY. This work was supported by the Engineering and Physical Science Research Council of UK under the grant no. EP/T026022/1. The review of this paper was arranged by Editor XXX (Corresponding author: J. F. Zhang, e-mail: j.f.zhang@ljmu.ac.uk)

K. H. Tok, J. F. Zhang, J. Brown, Z. Ji, W. Zhang, and J. S. Marsland are with School of Engineering, Liverpool John Moores University, Liverpool L3 3AF, UK. Z. Ji now is with School of Microelectronics, Shanghai Jiaotong University, Shanghai 200240, China.

- To extract the CDF per trap of the four parameters needed for circuit simulation: $\Delta I_{D,LIN}/I_{D,LIN}$, $\Delta I_{D,SAT}/I_{D,SAT}$, $\Delta V_{TH,LIN}$, and $\Delta V_{TH,SAT}$.

In addition, based on the RTN measured at $V_G \approx V_{TH}$, an integral methodology for modelling RTN has been proposed and it can be used to predict the long term RTN [16],[17]. The predicative capability of this integral method for $\Delta I_{D,LIN}/I_{D,LIN}$ measured under $V_G > V_{TH}$ will be tested in this work.

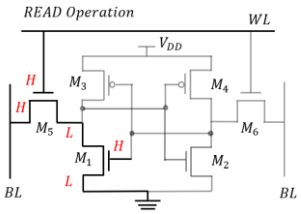

Fig. 1. Schematic 6 transistor SRAM cell under Read Operation. M1 and M5 operates under linear and saturation conditions, respectively.

II. Devices and Experiments

A. Devices

The nMOSFETs used in this work were fabricated by a commercial 28 nm CMOS process. The channel length and width are 27 and 90 nm, respectively. The gate stack consists of a metal layer, a Hf-high-k layer, and an interfacial SiON layer. The equivalent oxide thickness is 1.2 nm. The average threshold voltage, V_{TH}, is 0.45 V.

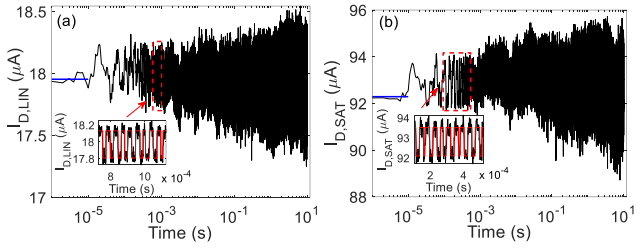

Fig. 2. The raw I_D recorded against log(time) under linear (a) and saturation (b) modes, respectively. The solid blue lines represent the $I_{D,REF}$. The insets show the two level RTN within relatively short time window. The red lines in the insets were fitted RTN by the hidden Markov model (HMM).

B. Experiments

The tests were carried out under either linear or saturation conditions. Linear condition has a constant gate bias, V_G, of 0.9 V and a constant drain bias, V_D, of 0.1 V. The saturation condition has $V_G = V_D = 0.9$ V. Under these biases, the hot carrier and PBTI ageing is insignificant, when compared with RTN fluctuation.

All tests were carried out under 125 °C. This is the temperature of thermal chuck and the self-heating effect is not included, since the test devices are bulk MOSFETs, rather than FinFETs. To assess RTN at different temperature, more tests are needed, which is out of the scope of this work.

The test was repeated on 402 devices with a time window of 10 sec for statistical analysis. To verify that the RTN model extracted from data in this short window can be used to predict long term RTN, tests with a time window up to 6×10^4 sec were also -carried out for 60 devices.

During RTN tests, drain current, I_D, was monitored continuously at a sampling rate of 1 Mpoint/sec [26] and an example of the raw I_D is given in Figs. 2(a) and 2(b) for linear and saturation mode, respectively. The reference I_D, I_{REF}, was obtained from the average of the first 10 points, as represented by the blue lines. It was used for evaluating the relative shift of I_D, $\Delta I_D/I_D = (I_{REF}-I_D)/I_D$. A positive $\Delta I_D/I_D$ represents a reduction of I_D, therefore. The insets in Figs. 2(a) and (b) show that I_D can follow the two level RTN within a relatively short time window. As time increases, slower traps start contributing, leading to increased fluctuation.

The V_G waveform for measuring ΔV_{TH} is given in Fig. 3(a). The reference V_{TH}, $V_{TH,REF}$, is first extracted by the maximum gm method from a reference pulse (3μs) I_D-V_G, taken before starting the RTN test for each device and an example is given in Fig. 3(b). The V_D used for the pulse I_D-V_G is 0.1 or 0.9 V for the linear or saturation tests, respectively. The current at $V_G = V_{TH,REF}$, I_{DTH}, is recorded. During the RTN tests, the pulse I_D-V_G were taken periodically and the instabilities in V_{TH} were extracted from the V_G shift at $I_D = I_{DTH}$, i.e. $\Delta V_{TH} = V_G (I_D = I_{DTH}) - V_{TH,REF}$, as shown by the inset of Fig. 3(b).

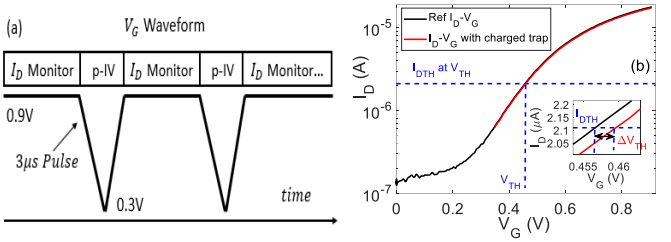

Fig. 3. (a) The V_G waveform for measuring ΔV_{TH}. (b) The ΔV_{TH} was extracted from the shift of V_G at $I_D = I_{DTH}$ when comparing the pulse I_D-V_G during RTN test with the reference I_D-V_G measured before starting the RTN test.

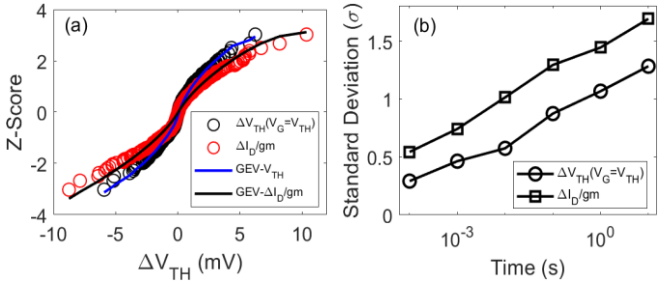

Fig. 4. A comparison between ΔV_{TH} measured at $V_G = V_{TH}$ and ΔI_D/gm. (a) The CDF of ΔV_{TH} and ΔI_D/gm at 10 sec. (b) Standard deviation at different time.

III. Results and Discussions

A. Linear mode: CDF of $\Delta V_{TH,LIN}$ and $\Delta I_{D,LIN}/I_{D,LIN}$

Integral methodology: The details of integral methodology were given in early works [16] and a brief description of its principle is included here for self-completeness. Early works [2,4,5,8-12] measured the impact of individual traps separately and collected a number of traps to build the CDF per trap. In contrast, integral method measures the total RTN in a device that can come from the cumulative contribution of multiple traps. One advantage of the integral method is that it does not require separating the measured fluctuation, such as that in Figs. 2(a) and 2(b), into the contributions of individual traps experimentally. This separation is carried out numerically

based on the maximum likelihood estimation [16],[17]. At a given time, each device will give one data point in Fig. 4(a) and 402 devices give the dataset for the CDF.

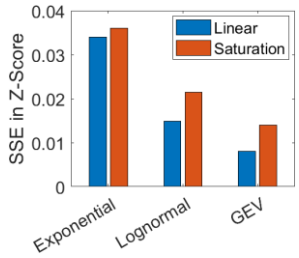

Fig. 5. The sum of squared errors (SSE) per device of the fitted CDFs on the $\Delta V_{TH}(V_G=V_{TH})$. Blue bars are for $\Delta V_{TH,LIN}$ in Fig. 4(a). Red bars are for $\Delta V_{TH,SAT}$ in Fig. 10(b).

It should be pointed out that the CDF in Fig. 4 is for $\Delta V_{TH,LIN}$ per device, which is different from the CDF per trap, as one device can have multiple traps. By assuming that the number of traps per device follows the Poisson distribution and the CDF per trap follows one of the distributions: Exponential, Log-normal, or General Extreme Value (GEV), the parameters in these distributions can be extracted through the Maximum likelihood method [16],[27]. In this way, the integral method decomposes $\Delta V_{TH,LIN}$ into the contribution of individual traps numerically and statistically, rather than experimentally.

A comparison between $\Delta V_{TH,LIN}$ and $\Delta I_{D,LIN}$ /gm: Figs. 4(a)&(b) compare the CDF and standard deviations of these two, respectively. In agreement with early work [25], the $\Delta I_{D,LIN}$/gm has substantially large deviation and longer distribution tails. As a result, using $\Delta I_{D,LIN}$/gm will overestimate the RTN in threshold voltage and $\Delta V_{TH,LIN}$ should be directly measured at $V_G=V_{TH}$, as shown in Fig. 3(b). We will use the directly measured $\Delta V_{TH,LIN}$ hereafter.

Selection of CDF per trap for $\Delta V_{TH,LIN}$: By assuming the impact per trap following a specific CDF: Exponential, Log-normal, or GEV, the parameters in these CDFs can be extracted through the maximum likelihood estimation [16],[17]. To find which CDF matches the test data best, the sum of squared error (SSE) per device between the extracted CDF and the test data is calculated. The blue bars in Fig. 5 show that GEV has the least SSE.

One possible explanation for the lowest error of GEV is that it has three fitting parameters, while Log-normal has two and Exponential has one. To further support the selection of GEV, we examine the dependence of SSE on the number of devices, N. If the data truly follows a CDF, SSE should reduce with increasing N. As N approaches infinite, SSE per data/device should approach zero [27].

For the red dashed line in Fig. 6(a), we used a theoretical Log-normal CDF to generate different number of hypothetical devices. These devices were then fitted with Log-normal CDF and the SSE per device is calculated from the difference between the fitted and the theoretical Log-normal CDFs. As expected, the SSE indeed reduces for higher number of devices. Repeating the same procedure for GEV gives the blue dashed line. When comparing the theoretical dashed lines with the corresponding solid lines of test data, GEV (blue) is the closest, supporting its selection for $\Delta V_{TH,LIN}$.

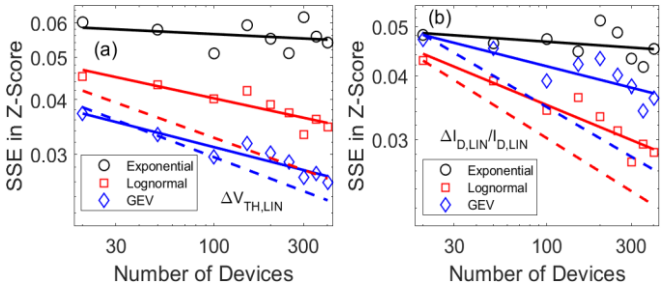

Fig. 6. Dependence of SSE per device on the number of devices used to extract the CDFs. The solid lines are fitted with test data. The dashed lines are results of hypothetical devices generated from theoretical Lognormal or GEV CDFs. (a) $\Delta V_{TH,LIN}$: GEV (blue) agrees best with theoretical line. (b) $\Delta I_{D,LIN}/I_{D,LIN}$: Log-normal (red) agrees best with theoretical line.

Selection of CDF per trap for $\Delta I_{D,LIN}/I_{D,LIN}$: The CDF of measured $\Delta I_{D,LIN}/I_{D,LIN}$ at V_G=0.9 and V_D=0.1 V at 10 sec is given in Fig. 7(a). By repeating the same procedure as that for $\Delta V_{TH,LIN}$, the SSE of extracted CDF per device is given in Fig. 7(b) as the blue bars. The Log-normal has the lowest error here, although it has less fitting parameter than GEV. Fig. 6(b) also shows that differences in the SSE dependence on device number between the theoretical and test data are smallest for the Log-normal (the red lines). Log-normal should be selected for $\Delta I_{D,LIN}/I_{D,LIN}$, therefore.

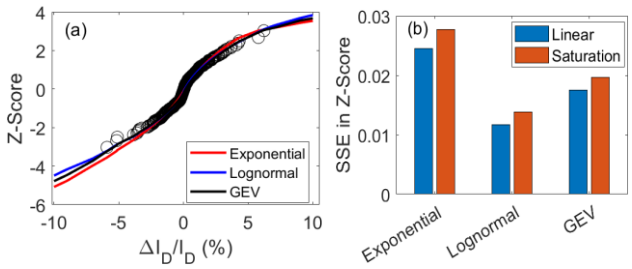

Fig. 7. (a) $\Delta I_{D,LIN}/I_{D,LIN}$ fitted with three different CDF per trap: Exponential, Log-normal, and GEV. (b) The errors. Blue bars are errors for $\Delta I_{D,LIN}/I_{D,LIN}$ in Fig. 7(a) and red bars are errors for $\Delta I_{D,SAT}/I_{D,SAT}$ in Fig. 10(a). In both cases, Log-normal gives the lowest errors.

The results show that different types of CDF per trap should be used for $\Delta I_{D,LIN}/I_{D,LIN}$ and $\Delta V_{TH,LIN}$: Log-normal for $\Delta I_{D,LIN}/I_{D,LIN}$ and GEV for $\Delta V_{TH,LIN}$. To explain this difference, it is noted that the GEV generally has larger statistical spread than Log-normal, when fitting the same set of data [16],[17],[27]. As mentioned earlier, the current follows narrow percolation path under $V_G=V_{TH}$, but is more uniformly distributed under $V_G>V_{TH}$. The trapped charges above the narrow current path will have profound impacts on the device under $V_G=V_{TH}$, generating the outlier and large statistical spread. GEV describes these outliers better.

Impact of V_G on RTN: This work measures RTN at V_G=0.9 V, since this is the operation voltage specified for this CMOS technology by the foundry. The desire to minimize power consumption can push V_G toward V_{TH}=0.45 V in future, so that it is of interest to compare the RTN at V_G=0.9 V with that at V_G=0.5 V. Fig. 8 shows that RTN has higher standard deviation, σ, at V_G=0.9 V. When compared with I_D at V_G=0.5 V, I_D at V_G=0.9 V is higher and more uniformly distributed. This allows more traps being charged, resulting in the higher RTN.

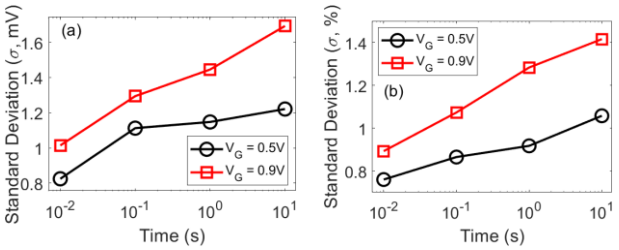

Fig. 8. A comparison of the RTN standard deviation under V_G=0.9 V with that under V_G=0.5 V. (a) ΔV_{TH} and (b) $\Delta I_{D,LIN}/I_{D,LIN}$.

B. Saturation mode: CDF of $\Delta I_{D,SAT}/I_{D,SAT}$ and $\Delta V_{TH,SAT}$

The CDF per trap for $\Delta I_{D,SAT}/I_{D,SAT}$: The measured $I_{D,SAT}$ and $\Delta I_{D,SAT}$ on 402 devices are given in Figs. 9(a) and (b), respectively. When compared with $I_{D,LIN}$, the mean $I_{D,SAT}$ is a factor of 4.8 higher. The $\Delta I_{D,SAT}$ in Fig. 9(b) has wider statistical spread than the $\Delta I_{D,LIN}$. This is mainly caused by the higher transconductance in saturation, so that the same shift in overdrive voltage (V_G-V_{TH}) results in larger shift in I_D.

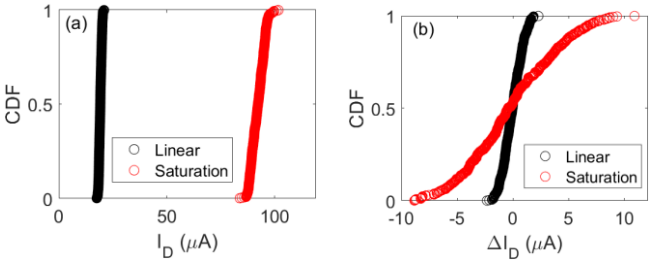

Fig. 9. A Comparison between linear (black) and saturation (red) measurements on 402 devices: (a) I_D and (b) ΔI_D.

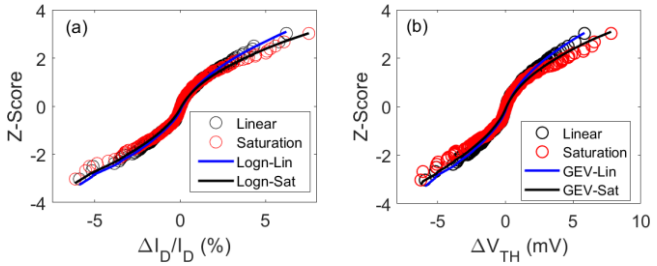

Fig. 10. A Comparison between linear and saturation measurements: (a) $\Delta I_D/I_D$ and (b) ΔV_{TH}. The lines are fitted results.

Fig. 10(a) compares $\Delta I_{D,SAT}/I_{D,SAT}$ with $\Delta I_{D,LIN}/I_{D,LIN}$. The normalization brings these two much closer, but $\Delta I_{D,SAT}/I_{D,SAT}$ clearly has larger statistical spread than $\Delta I_{D,LIN}/I_{D,LIN}$. The integral methodology is again used to extract the CDF per trap for $\Delta I_{D,SAT}/I_{D,SAT}$. As reported in Section III.A, GEV describes the RTN with narrow percolation current path best, while Log-normal CDF describes the RTN under relatively uniform current flow better. Under saturation, current flow should be relatively uniform near the source, but can be less uniform near the pinch-off point, where it is close to threshold condition. Fig. 7(b), however, shows that Log-normal CDF has the lowest error for $\Delta I_{D,SAT}/I_{D,SAT}$. This indicates that the inversion channel region has a larger impact on $\Delta I_{D,SAT}/I_{D,SAT}$ distribution.

To further compare $\Delta I_{D,SAT}/I_{D,SAT}$ with $\Delta I_{D,LIN}/I_{D,LIN}$ quantitatively, we extracted their statistical properties at different time windows. Fig. 11 shows that the number of average effective charged acceptor-like (N_A) and donor-like (N_D) traps is clearly larger for $\Delta I_{D,SAT}/I_{D,SAT}$. The procedure and formula used to extract the N_A and N_D for $\Delta I_D/I_D$ are the same as those used for ΔV_{TH} in early works [16],[17], but the unit for average impact per trap, μ, is % for $\Delta I_D/I_D$ and mV for ΔV_{TH}. Fig. 12 shows that μ is insensitive to time and in a range of 0.48 ~ 0.75% for N_A, and 0.46 ~ 0.6% for N_D, depending on the CDF per trap used.

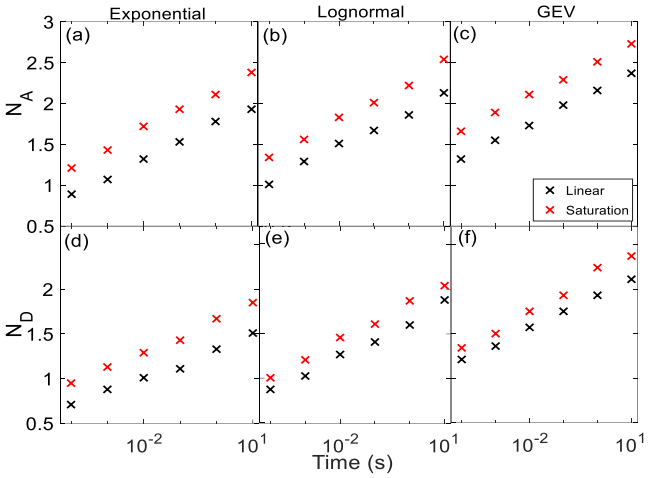

Fig. 11. $\Delta I_D/I_D$ at different time window: The extracted average number of acceptor-like (a)-(c) and donor-like (d)-(f) traps per device. The red and black symbols are for $\Delta I_{D,SAT}/I_{D,SAT}$ and $\Delta I_{D,LIN}/I_{D,LIN}$, respectively.

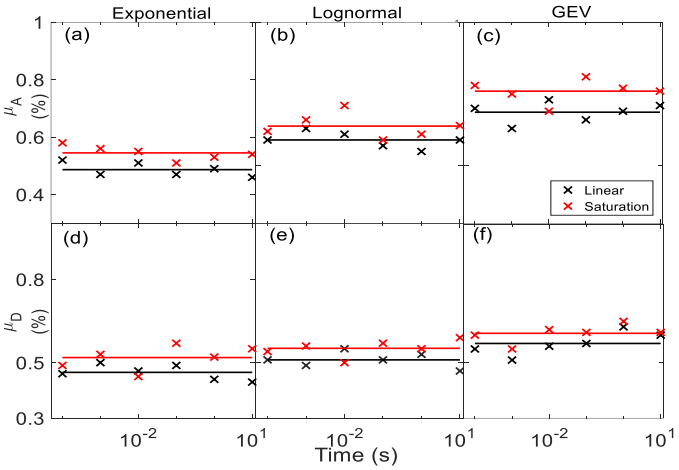

Fig. 12. $\Delta I_D/I_D$ at different time window: The extracted average impact per acceptor-like (a)-(c) and donor-like (d)-(f) traps. The red and black symbols are for $\Delta I_{D,SAT}/I_{D,SAT}$ and $\Delta I_{D,LIN}/I_{D,LIN}$, respectively.

Based on the extracted Log-normal CDF, the projected $\Delta I_{D,SAT}/I_{D,SAT}$ and $\Delta I_{D,LIN}/I_{D,LIN}$ at different multiples of standard deviation, σ, are given in Fig. 13(a). At 6σ, $\Delta I_{D,SAT}/I_{D,SAT}$ = 35.7% and $\Delta I_{D,LIN}/I_{D,LIN}$ = 27.5%. Typically, $\Delta I_D/I_D$ = 10% has been used to define ageing-induced device lifetime [28]. At 10%, Fig. 14(a) shows that the failure rates for $\Delta I_{D,SAT}/I_{D,SAT}$ and $\Delta I_{D,LIN}/I_{D,LIN}$ are 90 and 30 parts-per-million (ppm), respectively. At 25%, they become 0.17 for $\Delta I_{D,SAT}/I_{D,SAT}$ and 0.008 for $\Delta I_{D,LIN}/I_{D,LIN}$. These differences confirm the need to model $\Delta I_{D,SAT}/I_{D,SAT}$ and $\Delta I_{D,LIN}/I_{D,LIN}$ separately.

Selection of CDF per trap for $\Delta V_{TH,SAT}$: Similar to $\Delta V_{TH,LIN}$, $\Delta V_{TH,SAT}$ was measured from the shift of pulse I_D-V_G, but under V_D=0.9 V. Their distributions are compared in Fig. 10(b). Like $\Delta I_D/I_D$, $\Delta V_{TH,SAT}$ has higher statistical spread than $\Delta V_{TH,LIN}$. As both of them were measured under the condition where current

follows narrow percolation paths, it is not surprising to find that GEV also describes the CDF per trap best for $\Delta V_{TH,SAT}$.

The larger RTN in saturation mode should be explained. The occupancy of a trap in the gate dielectric should increase with (i) a higher number of mobile charge carriers beneath it in the substrate and (ii) the higher energy of these charge carriers. When an electron in silicon has higher energy, the potential barrier for its tunnelling will be lower. As a result, hot carriers are more efficient to fill a trap and increase the trap occupancy.

Under the same V_G=0.9 V, raising V_D from 0.1 to 0.9 V reduces the electrical field between gate and conduction channel away from the source. This leads to a reduction of the charge carrier density toward drain, which in turn reduces the trap occupancy. The reduced field between gate and channel does not support a higher $\Delta V_{TH,SAT}$, therefore. On the other hand, the lateral field increases when raising V_D, which makes carrier 'hot' and increase trap filling. The experimental results show that the rise in carrier energy overcompensates the reduction in carrier density, resulting in higher $\Delta V_{TH,SAT}$.

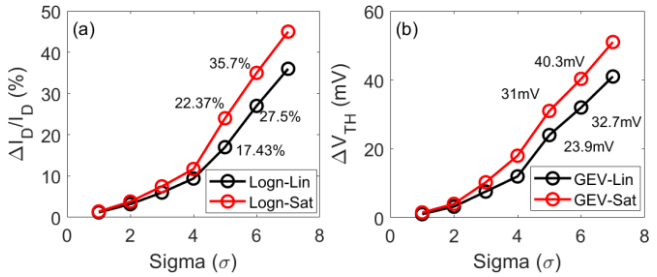

Fig. 13. A comparison of linear and saturation $\Delta I_D/I_D$ (a) and ΔV_{TH} (b) at different multiples of standard deviation, evaluated from Log-normal for (a) and GEV for (b).

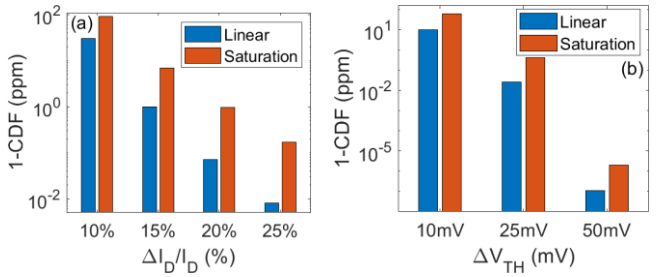

Fig. 14. The failure rates against failure criteria evaluated by log-normal CDF for $\Delta I_D/I_D$ (a) and GEV for ΔV_{TH} (b).

Table I. The CDF formula and their average parameter values extracted between 10^{-4} and 10 sec. The 'δ' is used in the formula to emphasize that it is the impact per trap, rather than per device.

	CDF	Accept or	Donor
Lognormal $\frac{\Delta I_{D,LIN}}{I_{D,LIN}}$	$\frac{1}{2}+\frac{1}{2}\mathrm{erf}\left(\frac{\ln(\delta I_{D,LIN}/I_{D,LIN})-\epsilon}{\sqrt{2}\theta}\right)$	$\epsilon = -0.53$ $\theta = 0.13$	$\epsilon = -0.68$ $\theta = 0.11$
Lognormal $\frac{\Delta I_{D,SAT}}{I_{D,SAT}}$	$\frac{1}{2}+\frac{1}{2}\mathrm{erf}\left(\frac{\ln(\delta I_{D,SAT}/I_{D,SAT})-\epsilon}{\sqrt{2}\theta}\right)$	$\epsilon = -0.47$ $\theta = 0.15$	$\epsilon = -0.61$ $\theta = 0.16$
GEV $\Delta V_{TH,LIN}$	$e^{-\left(\left(1+\xi\left(\frac{\delta V_{TH,LIN}-\alpha}{\beta}\right)\right)^{-\frac{1}{\xi}}\right)}$	$\xi = 0.54$ $\alpha = 0.46$ $\beta = 0.13$	$\xi = 0.4$ $\alpha = 0.37$ $\beta = 0.14$
GEV $\Delta V_{TH,SAT}$	$e^{-\left(\left(1+\xi\left(\frac{\delta V_{TH,SAT}-\alpha}{\beta}\right)\right)^{-\frac{1}{\xi}}\right)}$	$\xi = 0.51$ $\alpha = 0.49$ $\beta = 0.14$	$\xi = 0.38$ $\alpha = 0.4$ $\beta = 0.15$

To have a quantitative comparison, the number of effective charged traps per device is given in Fig. 15 for different time windows. The N_A and N_D are clearly higher for $\Delta V_{TH,SAT}$. Fig. 16 shows that the mean impact per trap, however, is similar for $\Delta V_{TH,SAT}$ and $\Delta V_{TH,LIN}$. Based on the extracted GEV per trap, Fig. 13(b) gives the predicted ΔV_{TH} at multiple σ at 10 sec. At 6σ, $\Delta V_{TH,SAT}$ and $\Delta V_{TH,LIN}$ reaches 40.3 mV and 32.7 mV, respectively. Fig. 14(b) shows that the failure rates for $\Delta V_{TH,SAT}$ and $\Delta V_{TH,LIN}$ are 2×10^{-6} and 0.11×10^{-6} ppm at 50 mV. As a result, they should be modelled separately as well.

Table I summarizes the best CDF per trap for the four parameters: $\Delta I_{D,LIN}/I_{D,LIN}$, and $\Delta I_{D,SAT}/I_{D,SAT}$, $\Delta V_{TH,SAT}$ and $\Delta V_{TH,LIN}$.

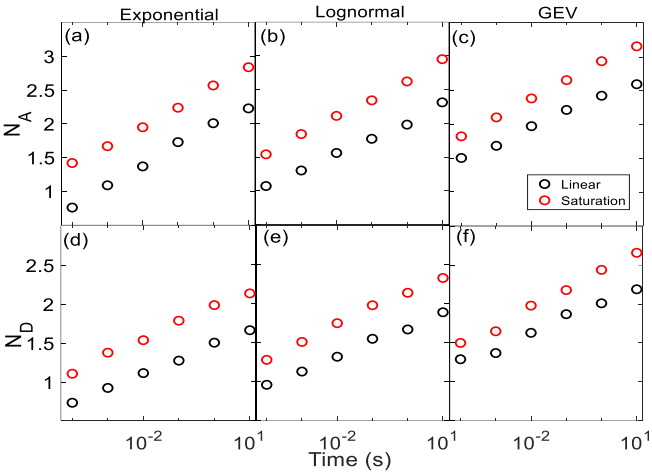

Fig. 15. ΔV_{TH} at different time window: The extracted average number of acceptor-like (a)-(c) and donor-like (d)-(f) traps per device. The red and black symbols are for $\Delta V_{TH,SAT}$ and $\Delta V_{TH,LIN}$, respectively.

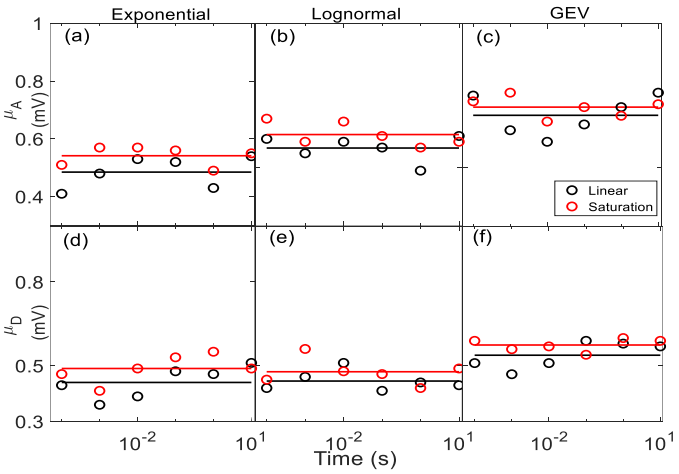

Fig. 16. ΔV_{TH} at different time window: The extracted average impact per acceptor-like (a)-(c) and donor-like (d)-(f) trap. The red and black symbols are for $\Delta V_{TH,SAT}$ and $\Delta V_{TH,LIN}$, respectively.

C. *Predicting long term $\Delta I_{D,LIN}/I_{D,LIN}$*

The principle: One advantage of the integral methodology is that it does not require the statistical distributions of trap's capture and emission times, which are difficult to establish [7,11,12]. At a given time window, the statistical distribution of ΔV_{TH} such as that in Fig. 10(b) is assumed originating from a set of 'effectively charged traps (ECT)'. As ECTs are always charged, there is no need to specify their capture and emission

times. As time window increases, slower traps become active, leading to higher N_A and N_D in Figs. 11 and 15. In this way, the impact of trap's time constants on ΔV_{TH} is transformed to the dependence of N_A and N_D on time. If the kinetics of N_A and N_D can be established based on short time test data, their long term values can be estimated through extrapolation, similar to predicting the ageing-induced device lifetime. Once N_A and N_D is known, the CDF of ΔV_{TH} at long term can be directly evaluated, without using the time consuming Monte Carlo simulation.

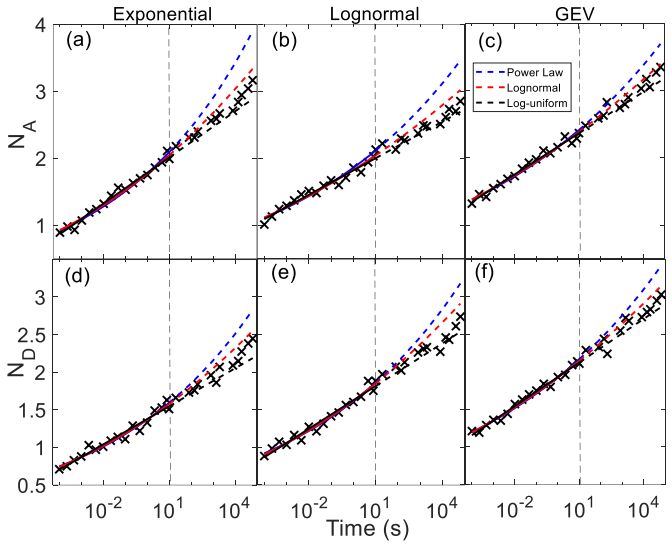

Fig. 17. $\Delta I_{D,LIN}/I_{D,LIN}$: Prediction of the average number of acceptor-like traps per device, N_A, in (a)-(c) and donor-like traps per device, N_D, in (d)-(f). Symbols are extracted from three different distributions: Exponential, Lognormal and GEV. Solid lines within 10s represent the fitted kinetics with power law, lognormal and log-uniform. Dashed lines beyond 10s represent the extrapolated kinetics. The test data beyond 10s were not used for fitting.

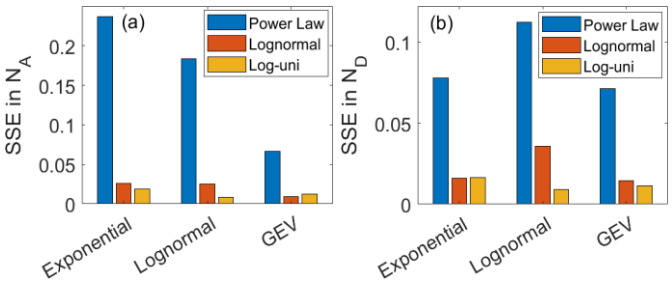

Fig. 18. The sum of squared errors of the prediction in Fig. 17: (a) for acceptor-like and (b) for donor-like traps. The lowest errors were obtained for Log-uniform kinetics with Lognormal CDF.

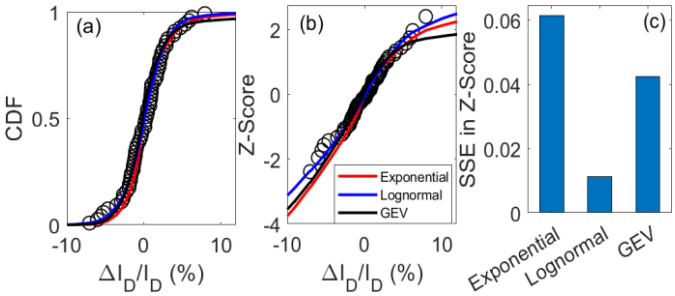

Fig. 19. A comparison of the measured and predicted CDF of $\Delta I_{D,LIN}/I_{D,LIN}$ at 6×10^4 sec. The CDF is plotted in linear (a) and Z-score (b) scales. (c) shows that the lowest error was obtained with Lognormal CDF per trap.

Applicability to $\Delta I_{D,LIN}/I_{D,LIN}$: Early works [16],[17] showed that the integral method was successfully used to predict the long term CDF of ΔV_{TH}. Its applicability to predict the long term $\Delta I_{D,LIN}/I_{D,LIN}$ measured at $V_G>V_{TH}$ is investigated here. Two statistical distributions have been proposed for trap's time constants by early works: Log-uniform or Log-normal [5,7-9]. Their applicability for the kinetics of N_A and N_D will be tested. In addition, power law has been widely used as the kinetics for ageing [29],[30] and will also be tested.

In Fig. 17, the data within 10 sec are used to fit the kinetics. The extracted kinetics were then extrapolated to 6×10^4 sec and compared with the test data. Fig. 18 shows that, for both N_A and N_D, the power law gives the highest errors, and the Log-uniform kinetics gives the lowest errors within the time window used in this work. If extrapolated further into future, Log-uniform kinetics will not saturate, but Lognormal kinetics will. Within a typical device lifetime (e.g. 10 years), further works are needed to determine which one is more accurate.

With the predicted N_A and N_D at 6×10^4 sec and the mean impact per trap given in Fig. 12, we can evaluate the CDF of $\Delta I_{D,LIN}/I_{D,LIN}$ based on the Poisson distribution of number of traps per device and log-normal CDF of the impact per trap. The predicted CDF of $\Delta I_{D,LIN}/I_{D,LIN}$ is compared with test data in Fig. 19 and good agreement has been obtained for the Log-normal CDF per trap. The integral methodology can be applied to predict the long term $\Delta I_{D,LIN}/I_{D,LIN}$, therefore.

IV. CONCLUSIONS

The key parameters required for circuit simulation include driving current and threshold voltage under both linear and saturation modes. For the first time, this work used the integral methodology to assess their RTN-induced statistical distributions per trap based on test data. Unlike early works that overestimates ΔV_{TH} by using $\Delta I_D/gm$, ΔV_{TH} was directly measured from the pulse I_D-V_G at $V_G=V_{TH}$ here. It is shown that $\Delta I_D/I_D$ and ΔV_{TH} follow different CDF: Log-normal describes $\Delta I_D/I_D$ best, while GEV is the best for ΔV_{TH}, under both linear and saturation conditions. Despite the reduced field between gate and conduction channel under saturation, both $\Delta I_{D,SAT}/I_{D,SAT}$ and $\Delta V_{TH,SAT}$ are actually higher than their linear counterparts, indicating hot carriers assisting trap-filling. The results show that the RTN in each of the key parameters should be modelled separately. The work also shows that the integral methodology can be used to predict the long term $\Delta I_{D,LIN}/I_{D,LIN}$, based on the CDF extracted from short test time window and Log-uniform kinetics.

ACKNOWLEDGMENT

The authors thank D. Vigar of Qualcomm Technologies International Ltd for supplying test samples.

REFERENCES

[1]. A. Asenov *et al.*, "Modeling and simulation of transistor and circuit variability and reliability," in *Proc. IEEE Custom Integrated Circuits Conf.*, Sep. 2010, pp. 1-8. doi: 10.1109/CICC.2010.5617627.

[2]. B. Kaczer *et al.*, "Origin of NBTI variability in deeply scaled pFETs," in *IEEE Proc. Int. Rel. Phys. Symp. (IRPS)*, May 2010, pp. 26–32, doi: 10.1109/IRPS.2010.5488856.

[3]. A. Asenov, R. Balasubramaniam, A.R. Brown, J.H. Davies "RTS amplitudes in decananometer MOSFETs: A 3D simulation study," *IEEE Trans. Electron Devices*, vol. 50, no. 3, pp. 839-845, 2003, doi: 10.1109/TED.2003.811418.

[4]. T. Grasser, "Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities," *Microelectron. Rel.*, vol. 52, pp. 39-70, 2012, doi:10.1016/j.microrel.2011.09.002.
[5]. R. Wang *et al.*, "Too Noisy at the Bottom? -Random Telegraph Noise (RTN) in Advanced Logic Devices and Circuits," in *IEDM Tech. Dig.*, Dec. 2018, pp. 388–391, doi: 10.1109/IEDM.2018.8614594.
[6]. M. Duan *et al.*, "Time-dependent variation: A new defect-based prediction methodology," *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers*, 2014, pp. 1-2, doi: 10.1109/VLSIT.2014.6894373.
[7]. M. Mehedi *et al.*, "An assessment of the statistical distribution of Random Telegraph Noise Time Constants," *IEEE Access*, vol. 8, no. 10, pp.1496-1499, 2020, doi: 10.1109/ACCESS.2020.3028747.
[8]. M. Tanizawa *et al.*, "Application of a Statistical Compact Model for Random Telegraph Noise to Scaled-SRAM Vmin Analysis," in *Proc. Symp. VLSI Technol.*, Jun. 2010, pp. 95–96, doi: 10.1109/VLSIT.2010.5556184.
[9]. K. Ito, T. Matsumoto, S. Nishizawa, H. Sunagawa, K. Kobayashi, H. Onodera, "Modeling of Random Telegraph Noise under Circuit Operation-Simulation and Measurement of RTN-induced delay fluctuation," in *Proc. 12th Int'l Symp. on Quality Electronic Design*, pp.22-27, 2011, doi: 10.1109/ISQED.2011.5770698.
[10]. K. Sonoda, K. Ishikawa, T. Eimori, O. Tsuchiya, "Discrete Dopant Effects on Statistical Variation of Random Telegraph Signal Magnitude," *IEEE Trans. Electron Devices*, vol. 54, no. 8, pp.1981-1925, 2017, doi: 10.1109/TED.2007.900684.
[11]. H. Miki *et al.*, "Understanding Short-term BTI Behavior through Comprehensive Observation of Gate-voltage Dependence of RTN in Highly Scaled High-k / Metal-gate pFETs," in *Proc. Symp. Very Large Scale Integr. (VLSI) Technol.*, Jun. 2011, pp. 148–149.
[12]. T. Nagumo, K. Takeuchi, T. Hase, and Y. Hayashi, "Statistical Characterization of Trap Position, Energy, Amplitude and Time Constants by RTN Measurement of Multiple Individual Traps," in *IEDM Tech. Dig.*, Dec. 2010, pp. 628–631, doi: 10.1109/IEDM.2010.5703437.
[13]. M. Duan *et al.*, "New analysis method for time-dependent device-to-device variation accounting for within-device fluctuation," *IEEE Trans. Electron Devices*, vol. 60, no. 8, pp. 2505–2511, Aug. 2013, doi: 10.1109/TED.2013.2270893.
[14]. P. Weckx *et al.*, "Defect-based compact modeling for RTN and BTI variability," in *Proc. Int. Rel. Phys. Symp.*, 2017, pp. CR-7.1-CR-7.6.
[15]. M. Luo, R. Wang, S. Guo, J. Wang, J. Zou, and R. Huang, "Impacts of Random Telegraph Noise (RTN) on Digital Circuits," *IEEE Trans. Electron Devices*, vol. 62, no. 6, pp. 1725-1732, 2015, doi: 10.1109/TED.2014.2368191.
[16]. K. H. Tok *et al.*, "An Integral Methodology for Predicting Long-Term RTN," in *IEEE Transactions on Electron Devices*, vol. 69, no. 7, pp. 3869-3875, July 2022, doi: 10.1109/TED.2022.3176585.
[17]. K. H. Tok *et al.*, "AC RTN: Testing, Modeling, and Prediction," in *IEEE Transactions on Electron Devices*, vol. 69, no. 10, pp. 5780-5786, Oct. 2022, doi: 10.1109/TED.2022.3195690.
[18]. R. Akimoto *et al.*, "Effect of Drain-to-Source Voltage on Random Telegraph Noise Based on Statistical Analysis of MOSFETs with Various Gate Shapes," *2020 IEEE International Reliability Physics Symposium (IRPS)*, 2020, pp. 1-6, doi: 10.1109/IRPS45951.2020.9128341.
[19]. Chia-Hao Pao *et al.*, "Impacts of random telegraph noise on the analog properties of FinFET and trigate devices and Widlar current source," *2012 IEEE International Conference on IC Design & Technology*, 2012, pp. 1-4, doi: 10.1109/ICICDT.2012.6232841.
[20]. A. Yonezawa *et al.*, "Statistical analysis of Random Telegraph Noise reduction effect by separating channel from the interface," *2012 IEEE International Reliability Physics Symposium (RIPS)*, 2012, pp. 3B.5.1-3B.5.7, doi: 10.1109/IRPS.2012.6241809.
[21]. K. Ota *et al.*, "Random Telegraph Noise after Hot Carrier Injection in Tri-gate Nanowire Transistor," *2019 Electron Devices Technology and Manufacturing Conference (EDTM)*, 2019, pp. 169-171, doi: 10.1109/EDTM.2019.8731025.
[22]. C. Liu, K. T. Lee, H. Lee, Y. Kim, S. Pae, and J. Park, "New observations on the random telegraph noise induced Vth variation innano-scale MOSFETs," in Proc. IEEE Int. Rel. Phys. Symp., Jun. 2014, pp. XT.17.1–XT.17.5, doi: 10.1109/IRPS.2014.6861194.
[23]. 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 *Proc. Symp. VLSI Technol.*, Jun. 2009, pp. 54–55.
[24]. K. Ota, M. Saitoh, C. Tanaka, D. Matsushita, and T. Numata, "Systematic study of RTN in nanowire transistor and enhanced RTN by hot carrier injection and negative bias temperature instability," in *Symp. VLSI Technol. (VLSI-Technol.) Dig. Tech. Papers*, Jun. 2014, pp. 1–2, doi: 10.1109/VLSIT.2014.6894417.
[25]. A. Manut *et al.*, "Trigger-When-Charged: A Technique for Directly Measuring RTN and BTI-Induced Threshold Voltage Fluctuation Under Use-Vdd," *IEEE Trans. Electron Devices*, vol. 66, no.3, pp. 1482-1488, 2019, doi: 10.1109/TED.2019.2895700.
[26]. M. Duan *et al.*, "Development of a technique for characterizing bias temperature instability-induced device-to-device variation at SRAM relevant conditions," *IEEE Trans. Electron Devices*, vol. 61, no. 9, pp. 3081–3089, Sep. 2014, doi: 10.1109/TED.2014.2335053.
[27]. M. Mehedi, K.H. Tok, Z. Ye, J.F. Zhang, Z. Ji, W. Zhang, J.S. Marsland, "On the accuracy in modelling the statistical distribution of Random Telegraph Noise Amplitude," *IEEE Access*, vol. 9, pp.43551-43561, 2021, doi: 10.1109/ACCESS.2021.3065869.
[28]. Z. Ji *et al.*, "Negative bias temperature instability lifetime prediction:Problems and solutions," in *IEDM Tech. Dig*., Dec. 2013, pp. 15.6.1–15.6.4. doi: 10.1109/IEDM.2013.6724638.
[29]. J. F. Zhang, Z. Ji, and W. Zhang, "As-grown-generation (AG) model of NBTI: A shift from fitting test data to prediction," Microelectron. Rel., vol. 80, pp. 109–123, Jan. 2018. doi: 10.1016/j.microrel.2017.11.026.
[30]. R. Gao *et al.*, "Predictive as-grown-generation (A-G) model for BTI induced device/circuit level variations in nanoscale technology nodes," in *IEDM Tech. Dig*., Dec. 2016, pp. 31.4.1–31.4.4. doi: 10.1109/IEDM.2016.7838520.