IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version February 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.Doi Number

On the accuracy in modelling the statistical distribution of Random Telegraph Noise Amplitude

Mehzabeen Mehedi[1], Kean H. Tok[1], Zengliang Ye[1], Jian F. Zhang[1], Zhigang Ji[1,2], Weidong Zhang[1], and John S. Marsland[1]

[1]School of Engineering, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK.
[2]School of Microelectronics, Shanghai Jiaotong University, Shanghai 200240, P. R. China.

Corresponding author: Jian F. Zhang (e-mail: j.f.zhang@ljmu.ac.uk).

This work was supported by the Engineering and Physical Science Research Council of UK under the grant nos. EP/L010607/1 and EP/T026022/1.

ABSTRACT The power consumption of digital circuits is proportional to the square of operation voltage and the demand for low power circuits reduces the operation voltage towards the threshold of MOSFETs. A weak voltage signal makes circuits vulnerable to noise and the optimization of circuit design requires modelling noise. Random Telegraph Noise (RTN) is the dominant noise for modern CMOS technologies and Monte Carlo modelling has been used to assess its impact on circuits. This requires statistical distributions of RTN amplitude and three different distributions were proposed by early works: Lognormal, Exponential, and Gumbel distributions. They give substantially different RTN predictions and agreement has not been reached on which distribution should be used, calling the modelling accuracy into questions. The objective of this work is to assess the accuracy of these three distributions and to explore other distributions for better accuracy. A novel criterion has been proposed for selecting distributions, which requires a monotonic reduction of modelling errors with increasing number of traps. The three existing distributions do not meet this criterion and thirteen other distributions are explored. It is found that the Generalized Extreme Value (GEV) distribution has the lowest error and meet the new criterion. Moreover, to reduce modelling errors, early works used bimodal Lognormal and Exponential distributions, which have more fitting parameters. Their errors, however, are still higher than those of the monomodal GEV distribution. GEV has a long distribution tail and predicts substantially worse RTN impact. The work highlights the uncertainty in predicting the RTN distribution tail by different statistical models.

INDEX TERMS Random telegraph noise (RTN), Yield, Device Variations, Time Dependent Variations, Jitters, Traps, Statistical distributions.

I. INTRODUCTION

Random telegraph noise (RTN) is a step-like fluctuation of drain current under constant gate and drain voltages. It has received many attentions, as it adversely affects the operation of electronic circuits [1]-[15]. As MOSFETs become smaller, RTN becomes increasingly important, driven by an increased impact of a single charge on smaller devices and an increase in the number of devices in a system [1]-[8]. A large number of devices in a system will contain more devices in the tail of statistical distributions, which can cause errors. Moreover, low power is a key requirement for many Internet-of-Things edge units and this drives the operation voltage towards threshold voltage, Vth [16]-[18]. The minimization of overdrive voltage, (Vg-Vth), in the future leaves little room to tolerate the RTN induced jitter [5], [16], [18].

There have been many efforts to model RTN, both in the frequency domain [1], [19]-[21] and in the time domain [1]-[6]. It is widely accepted that RTN originates from trapping/detrapping charge carriers from/to the conduction channel [1]-[21]. The number of traps per device follows the Poisson distribution [3]-[6]. To perform Monte Carlo simulation in the time domain, one needs the capture-emission times and RTN amplitude of traps [5], [18], [22], [23]. We studied the statistical distribution of

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/ACCESS.2021.3065869, IEEE Access

capture/emission time constants in an early work [18] and focus on the amplitude distribution here.

The RTN amplitude can be measured as a Vth shift, ΔVth, or a normalized drain current fluctuation, ΔId/Id. ΔVth is the accumulative effect of multiple traps on a device and we use δVth to represent the RTN amplitude of one trap. δVth is stochastic and one feature of its cumulative distribution function (CDF) is a long tail, when compared with the Gaussian/Normal distribution, as shown in Fig. 1a [2], [6]. It has been proposed that this long tail originates from the uneven distribution of current [2], [6], [7] since the impact of a trapped charge in the oxide on the device depends on the local current density beneath it [7], [24]. As schematically illustrated in Fig. 1b, the current near threshold voltage flows through narrow percolation path. It is rare to have a trap located just above this percolation path and such a trap will cause a large δVth and result in the long distribution tail [2], [6], [7], [24].

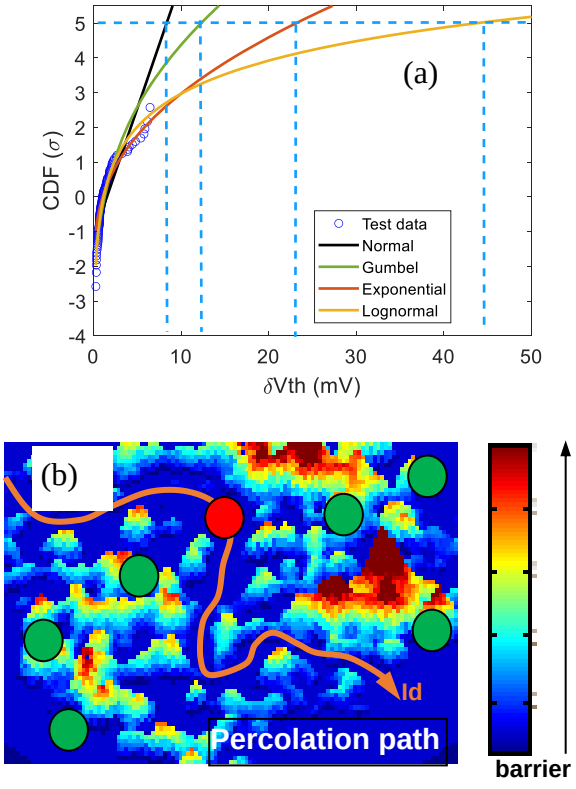

FIGURE 1. **(a) A comparison of different cumulative distribution functions (CDF) of threshold voltage shift, δVth. Each 'o' represents δVth induced by one trap and there are 100 traps here. Although both the Exponential and Lognormal CDFs describe the test data well, they give very different results when their tails were used to make predictions, for example at 5σ, as shown by the dashed lines. (b) A schematic illustration of the impact of traps (circles) on current path near threshold condition. The red circle represents a trap just above the percolation path of current, which has a large δVth and is in the distribution tail.**

Modelling the long tail in the CDF is a tall order and three statistical distributions have been proposed: Exponential [3]-[6], [23], [25]-[29], Lognormal [1], [5]-[8], [30]-[32] and Gumbel [9]-[11]. The success of RTN modelling in term of yield prediction for a system, such as SRAM, requires an accurate statistical distribution tail [2], [5], [7], [30], [31]. For a dataset of 100 traps, Fig. 1 shows that both Exponential and Lognormal CDFs agree well with the test data, but they have substantially different tails. For example, at 5σ where σ is the standard deviation, the δVth predicted by Exponential and Lognormal CDFs is 23 mV and 44 mV respectively. This uncertainty calls the accuracy of RTN modelling into question.

Agreement has not been reached on which distribution should be used. Many early works [1], [4], [9], [10], [25]-[28] only fitted their data with one statistical distribution. Different distributions were not compared and the reason for selecting a specific model is not given. For the works that compared the Exponential and Lognormal distributions [30], [31], it was reported that the Lognormal fitted the data better. There are, however, more fitting parameters in the Lognormal distribution than the Exponential distribution, so that it is not clear whether the improved fitting with the Lognormal originates from using extra fitting parameters.

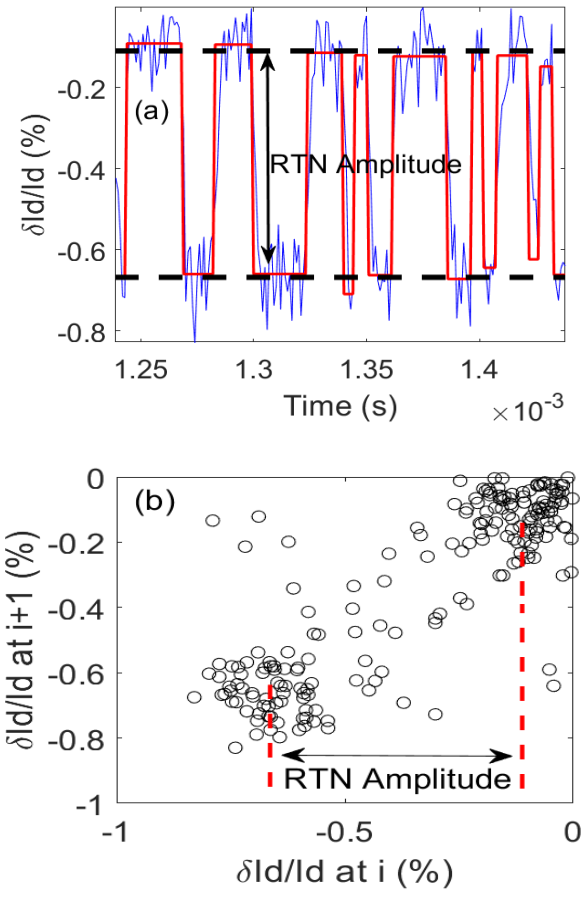

FIGURE 2. **(a) Extraction of RTN amplitude directly from the two discrete levels of Id used in this work. (b) The same data in (a) was used to extract the RTN amplitude by the conventional time-lag method. The RTN amplitudes extracted by these two methods agree well when there is only one trap in a device.**

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/ACCESS.2021.3065869, IEEE Access

The motivation of this work is to address the uncertainty in model selection for RTN amplitude through two ways. First, we attempt to find a statistical distribution that has lower error without using higher number of fitting parameters. In addition to the three distributions mentioned above, thirteen other distributions are evaluated. Second, we propose a new criterion for selecting statistical models. It will be shown that if the data truly follows a specific CDF, the error per trap should decrease when increasing the trap number.

We start by examining the three distributions mentioned above in terms of their errors both over the whole distribution and in the distribution tail. The number of traps used in some early works [7], [10], [25] is ~100, leaving too few traps in the tail (e.g. >95%) to evaluate the error reliably. To enable the tail evaluation, 1,178 traps were used here.

The CDF parameters are extracted by the Maximum likelihood estimation (MLE). Early works suggest that the accuracy can be improved by using either bimodal Exponential [29] or bimodal Lognormal distributions [6]. We will examine the impact of using bimodal distributions on the accuracy.

An analysis of the distributions proposed by early works [1]-[11] have not identified a clear winner. This leads us to search for new statistical distributions. Since there is little research on whether the RTN amplitude can be modeled better by other statistical distributions, apart from the three distributions mentioned above, a scoping study of different distributions are carried out. To emphasize the importance of the accuracy in the distribution tail for RTN modelling, the Z-score of corresponding CDF will be used to calculate errors, where $Z=(\delta Vth-\mu)/\sigma$ and μ is the average and σ the standard deviation. After comparing 16 distributions, it is found that Generalized Extreme Value (GEV) distribution [32] gives the lowest errors. GEV also meets the new criterion.

The last issue addressed in this work is the impact of trap number on the CDF accuracy in the distribution tail. The more traps used for extracting the distribution, the better the accuracy should be. In practice, however, the number of traps available is always limited. It is of importance to assess how reliable a CDF extracted from a limited number of traps can be used to predict the distribution tail at high sigma.

II. DEVICES, MEASUREMENT, AND METHODOLOGY

A. DEVICES

This work uses nMOSFETs fabricated by an industrial CMOS process, which has metal gate and a high-k/SiON stack. The channel length and width are 27×90 nm, respectively. The equivalent oxide thickness is 1.2 nm.

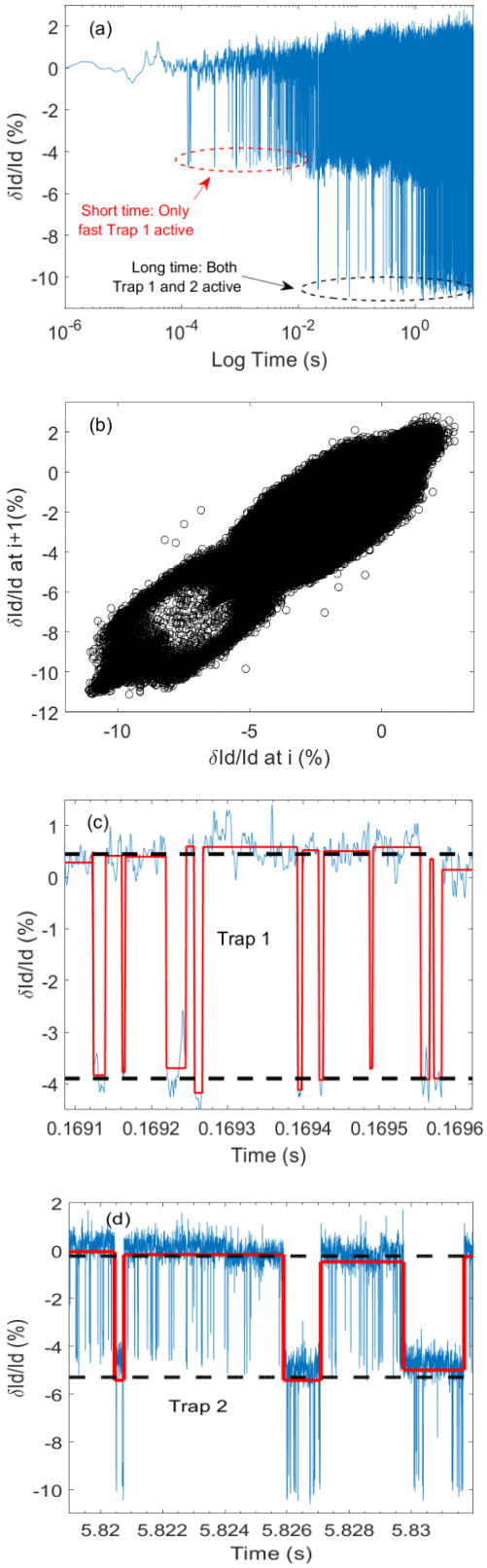

FIGURE 3. **(a) An example of two active RTN traps in a device. (b) It is difficult to use the time-lag method to extract the RTN amplitudes for the dataset in (a). (c) Extraction of the RTN amplitude of fast trap 1 by applying our method in the short time range. (d) Extraction of the RTN amplitude of slow trap 2 by applying our method in the long-time range.**

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/ACCESS.2021.3065869, IEEE Access

B. MEASUREMENTS

The test starts by applying a step voltage to the gate and drain. The Id is then monitored under a fixed Vg and Vd by an oscilloscope at a sampling rate of 1 Mpoint/sec [18], [33]. As low power requirement is driving operation voltage towards Vth and the average Vth of the devices used here is 0.45 V, we chose Vg=0.5 V and Vd=0.1 V for monitoring RTN. Unless otherwise specified, tests were carried out under 125 ºC.

Some typical results are given in Figs. 2 and 3, where the current fluctuation is plotted as δId/Id=(Id-Iref)/Id. The reference Id, Iref, was taken from the average of the first ten points of the measurement [18]. As Vg is close to Vth, δVth can be evaluated from -δId/gm, where gm is transconductance [24]. The gm is evaluated from a pulse (3 μs) Id-Vg, taken before the RTN test for each device [24].

Table I. The cumulative distribution functions (CDFs).

CDF (Exponential)	CDF (Lognormal)
$1-e^{\left(-\frac{\delta Vth}{\eta}\right)}$	$\frac{1}{2}erfc\left(-\frac{\ln(\delta Vth)-\varepsilon}{\theta\sqrt{2}}\right)$
CDF (GEV)	**CDF (Gumbel)**
$e^{\left\{-\left[1+\xi\left(\frac{\delta Vth-\alpha}{\beta}\right)\right]^{-\frac{1}{\xi}}\right\}}$	$e^{(-e^{\frac{-(\delta Vth-\alpha)}{\beta}})}$

C. METHOD FOR EXTRACTING RTN AMPLITUDE

In the test of negative bias temperature instability (NBTI), the impact of one trap on Id is typically measured directly from its discharge induced step-change of Id [28], [34]. For RTN tests, Hidden Markov Method [13] and Factorial HMM [14]-[15] has been used, when both the RTN amplitude and time constants are needed. The time-lag plot has been often used to measure the RTN amplitude [35].

Table II. The estimated CDF parameters.

CDF (Lognormal)	CDF (Exponential)	CDF (Gumbel)	CDF (GEV)
$\epsilon = 0.143$	$\eta = 1.640$	$\alpha = 1.033$	$\xi = 0.540$
$\theta = 0.792$	-	$\beta = 0.843$	$\alpha = 0.831$
-	-	-	$\beta = 0.542$

Similar to the NBTI measurement [28], [34], we measured the RTN amplitude directly from the step-changes in Id in this work. As shown in Fig. 2a, once a step-like change is observed, the Id for each discrete level is taken from the average of that level to minimize the effect of thermal noise. Moreover, unlike NBTI tests where discharging one trap is often an one-off event [28], [34], we take advantage of the multiple charge-discharge events in RTN and use the average of step-heights to further improve measurement accuracy. The minimum detectable δId/Id is ~0.2%, corresponding to a δVth of ~0.2 mV.

Fig. 2b shows that the amplitude extracted by our method agrees well with that of time-lag method, when there is only one trap in a device. The time-lag method, however, uses data in the whole time window and is difficult to use when there are multiple traps and one example is given in Figs. 3a and 3b. The advantage of our method is that it can be applied to a selected time range where two-level RTN events are identified. For the same dataset in Fig. 3a, Figs. 3c shows that the amplitude of fast trap 1 can be measured in a short time window. For a longer time window in Fig. 3d, the slow trap 2 becomes active and its amplitude can be extracted from the difference in the two discrete levels after averaging out the impact of the fast trap.

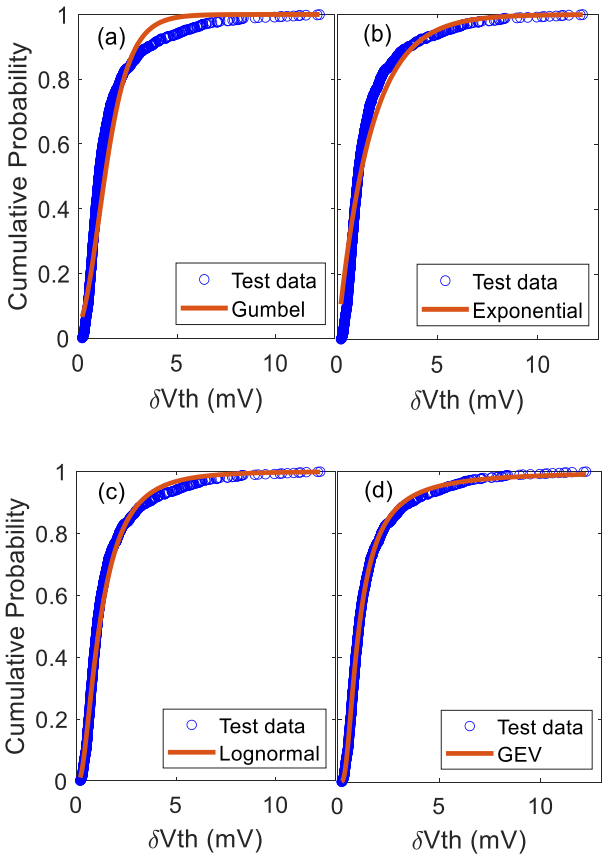

FIGURE 4. The CDFs (lines) extracted from 1,178 traps by the MLE method are compared with the test data (symbols) for: (a) Gumbel, (b) Exponential, (c) Lognormal, and (d) GEV, respectively.

III. RESULTS AND DISCUSSIONS

A. PROBLEMS WITH THE PROPOSED STATISTICAL DISTRIBUTIONS

For the RTN amplitude per trap, two popular statistical distributions used in early works are Lognormal [1], [5]-[8], and Exponential [3]-[6], [28], [29]. In addition, Gumbel distribution has been used to capture the long tail of RTN [9]-[11]. Their cumulative distribution functions (CDF) are summarized in Table I. Table I also gives the formula for the Generalized Extreme Value (GEV) distribution, which will

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/ACCESS.2021.3065869, IEEE Access

be investigated in Section III.C. In this section, we focus on the three distributions used in early works: Lognormal, Exponential, and Gumbel.

Using the equation in Table I, the parameters of different statistic distributions are extracted by the Maximum Likelihood Estimation (MLE) [29], [36]. MLE uses different weightings to different data to maximize the probability of test dataset occurrence [36]. Based on the 1,178 measured traps, the estimated parameters are given in Table II. The extracted CDFs are plotted together with test data in Fig. 4.

Following the early works [5], [6], [30], [31], [37], we use the error between the extracted CDFs and the test data to compare different statistical distributions. Fig. 5a shows the sum-square-error (SSE) per trap. Gumbel and Exponential CDFs have similar errors, while Lognormal CDF has lower error. For convenience, the result of GEV distribution is also given in Fig. 5, which will be discussed in Section III.C. If one uses the minimum error of whole dataset as a criterion, the Lognormal distribution should be better than the Exponential and Gumbel distributions.

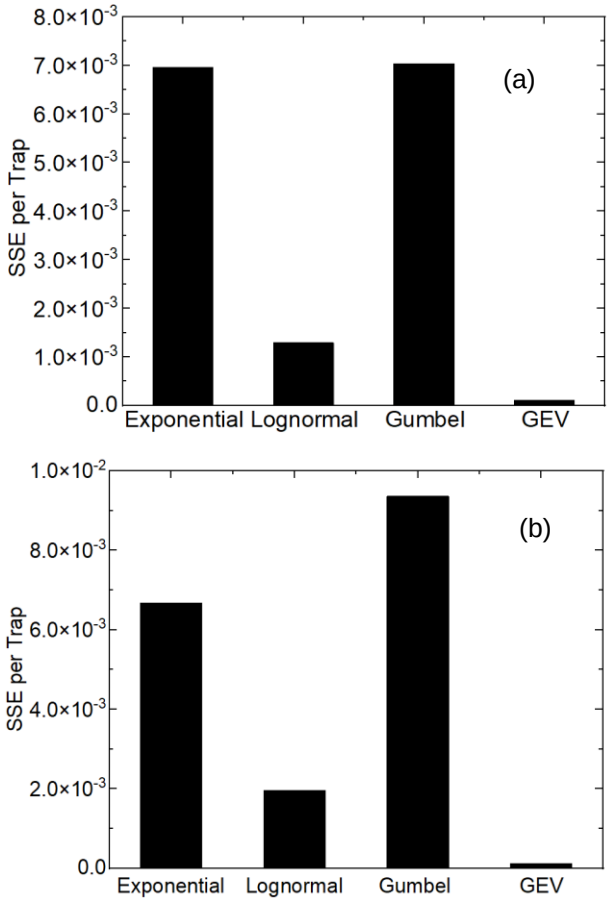

FIGURE 5. A comparison of the sum of square errors (SSE) per trap for CDFs extracted by Maximum likelihood estimation (MLE) at 125 °C (a) and 28 °C (b). The whole dataset were used in the error evaluation. Lognormal has smaller error than Exponential for the whole dataset.

The results in Fig. 5a were obtained at 125 °C. To show that the observation is independent of test conditions, Fig. 5b gives the results at 28 °C. RTN is generally sensitive to temperature and 814 traps were measured under 28 °C. The error of Lognormal distribution again is lower than that of Exponential and Gumbel distributions.

The minimum SSE per trap for the whole dataset should not be the only criterion for selecting statistical distribution functions. As the loss of yield is mainly caused by traps in the distribution tail, the SSE in the tail region should also be examined. To see the tail clearly, the corresponding Z-score of CDF is plotted linearly in Fig. 6. Some early works used ~100 traps [7], [10], [25] so that there are too few traps to evaluate the SSE reliably in the >95% tail. With 1,178 traps here, their SSE in the >95% tail is compared in Fig. 7. Although the Lognormal CDF has lower SSE for the whole dataset in Figs. 5a&b, Figs. 7a&b show that the Exponential CDF actually matches the test data better in this tail region. The choice between Lognormal and Exponential is not straightforward, therefore.

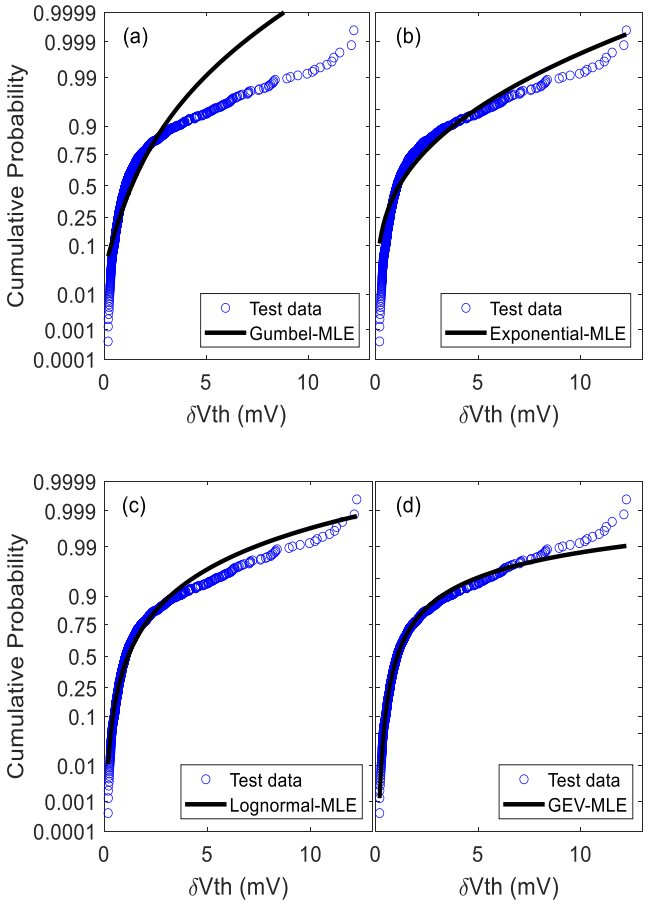

FIGURE 6. A comparison of the tail region between the test data (symbols) and the CDFs (lines) extracted by Maximum likelihood estimation (MLE) for (a) Gumbel, (b) Exponential, (c) Lognormal, and (d) GEV. The vertical axis is plotted linearly for the Z-score corresponding to the cumulative probability.

B. BIMODAL STATISTICAL DISTRIBUTIONS

To improve the accuracy of CDFs, bimodal CDF, BCDF, has been proposed:

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/ACCESS.2021.3065869, IEEE Access

$$BCDF = p * CDF1 + (1 - p) * CDF2 \qquad (1)$$

where 0≤p≤1 is an adjustable parameter that can be fitted by using the MLE method [36]. CDF1 and CDF2 are two monomodal distributions. Both bimodal Exponential [29] and Lognormal [6] CDFs have been used. It has been suggested that CDF1 and CDF2 originate from traps in different layers of gate dielectric that have different statistical properties [29].

Figs. 8a-f show the bimodal CDFs extracted by the MLE method for Lognormal, Exponential, and Gumbel, respectively. The bimodal Lognormal in Figs. 8a&b is dominated by the first Lognormal CDF and the contribution of the second Lognormal CDF is weak with a p value of only 0.033. For bimodal Exponential CDFs, the second CDF only counts for 9% in Figs. 8c&d. This increases to 25% for the bimodal Gumbel in Figs. 8e&f.

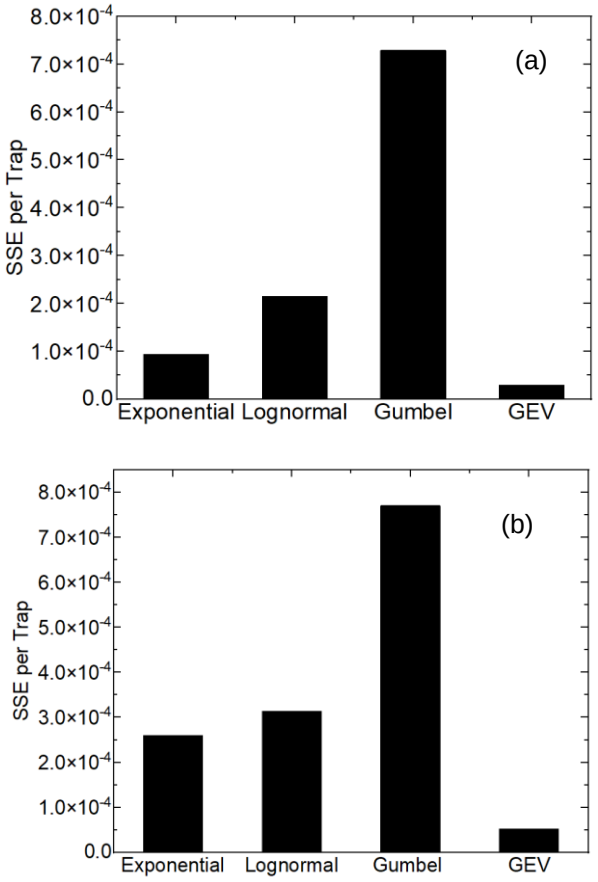

FIGURE 7. The error per trap of CDFs in the >95% tail extracted by MLE at 125 °C (a) and 28 °C (b). Exponential has smaller error than Lognormal in the tail.

To compare the bimodal CDFs with their monomodal counterparts, we calculate their errors from their Z-score plot in Figs. 8b, 8d, and 8f. This places more weightings on the distribution tails where accuracy is important for RTN modelling. It is more appropriate than the error calculation directly from CDF values in Fig. 5, therefore.

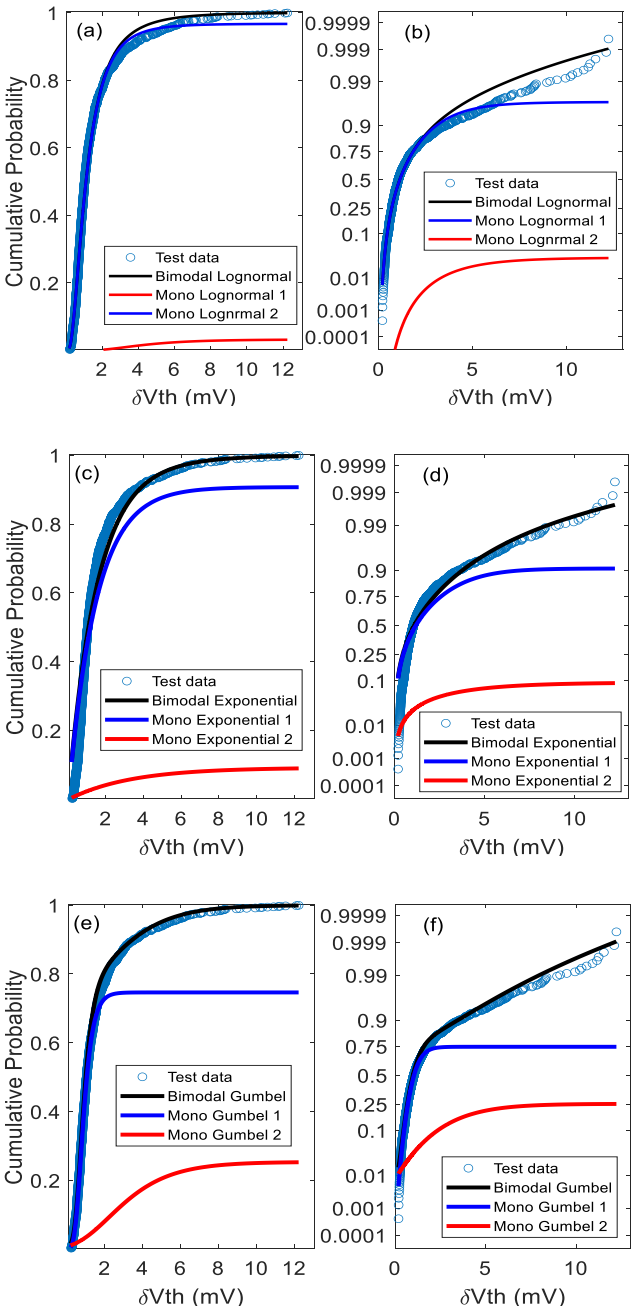

FIGURE 8. Bimodal CDFs for Lognormal (a) and (b), Exponential (c) and (d), and Gumbel (e) and (f). (b), (d) and (f) are the Z-score to enlarge the tail region. The symbols are test data. The black lines are the sum of two CDFs. The blue and red lines are the monomodal CDF1 and CDF2, respectively.

Fig. 9 shows that, although bimodal Gumbel has less error than monomodal Gumbel, it is still well above the error of monomodal Lognormal. The impact of using bimodal CDFs on the errors is modest for both Lognormal and Exponential. When compared with monomodal CDFs, bimodal CDFs more than double the number of fitting parameters. According to the Bayesian Information Criterion [38], penalty should be applied to models with more fitting parameters, so that using bimodal CDFs is not strongly supported by the data in this work. The question is whether

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/ACCESS.2021.3065869, IEEE Access

there is a monomodal CDF that can give similar or even smaller error than the lowest error achieved by the bimodal Lognormal CDF in Fig. 9. This will be investigated next.

C. GENERALIZED EXTREME VALUE (GEV) DISTRIBUTION

Driven by the desire to find a statistical distribution that has the lowest SSE per trap without using its bimodal CDF, we evaluated 13 other distributions [39] and their SSE per trap is compared in Fig. 10, together with the three distributions used in early works. Among the 16, the Generalized Extreme Value (GEV) distribution has the lowest error. It is worth of exploring this distribution further, therefore.

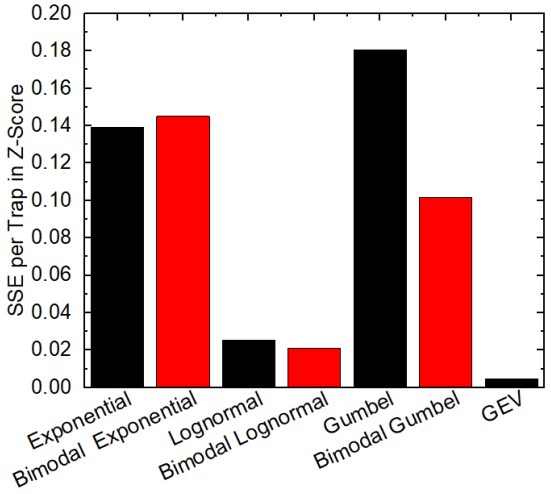

FIGURE 9. A comparison of errors in bimodal CDFs with their monomodal CDFs for Exponential, Lognormal, and Gumbel: The SSE per trap is calculated from the Z-score for the whole dataset. The use of bimodal CDFs has not reduced errors below the level achieved with a monomodal GEV.

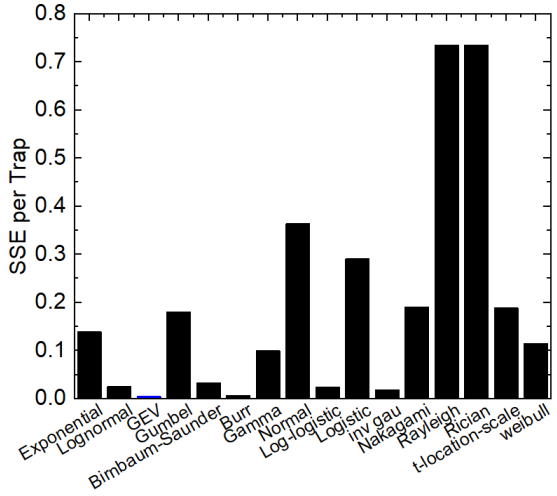

FIGURE 10. A comparison of the SSE per trap for 16 CDFs [39]. The error is calculated from the Z-score for the whole dataset. The Generalized extreme value (GEV) distribution has the lowest error.

The equation for GEV is included in Table I and the extracted parameter values are given in Table II. Fig. 4d shows that the CDF of GEV agrees well with the test data overall. Although Fig. 6d shows that the difference between GEV and the highest few data points appear increasing, this is an artifact, as the last few points of test data is always lifted upwardly by the limitation in the size of dataset. The Z-score approaches infinity when CDF approaches 1. As the last data point has CDF=1, its Z-score would be infinity. To avoid this, it is a common practice to calculate the CDF of test data by [40],

$$\mathrm{CDF}(\delta Vth,i)=(i-0.5)/N,$$

where i=1 has the lowest δVth and i=N=1,178 has the highest δVth in our test dataset. This brings the last CDF point from 1 to 0.999576 and their corresponding Z-score from infinity to 3.34. It, however, cannot completely eliminate the artificial up-swing of the last few data points.

Figs. 5 and 7 show that the GEV has the lowest error for both the whole dataset and the tail region when compared with other CDFs. Fig. 9 shows that the error of monomodal GEV CDF is also lower than that of the bimodal Lognormal, Exponential, and Gumbel CDFs. The number of fitting parameters is 5, 3, and 5 for the bimodal Lognormal, Exponential, and Gumbel CDFs, respectively. It is 3 for the GEV in Fig. 9, so that the better accuracy of GEV was not gained from using larger number of fitting parameters.

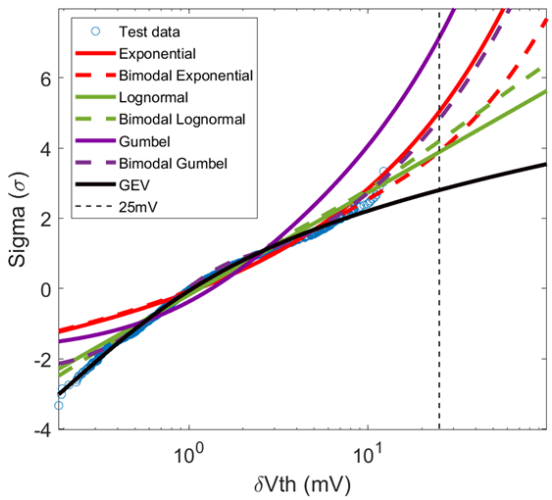

FIGURE 11. A comparison of different CDFs extracted from the same dataset (symbols). The solid lines are the monomodal CDFs and the dashed lines are their bimodal counterparts for the same color.

Fig. 11 compares the CDFs of different distributions extracted from the same dataset. The predicted distribution tail is sensitive to model selection. The δVth at high σ increases in the order of Gumbel, Exponential, Lognormal, and GEV. In another word, Gumbel has the shortest tail and gives the optimistic prediction, while the GEV has the longest tail and gives the pessimistic prediction. Fig. 12 compares the probability for δVth≥25 mV predicted by different CDFs. Quantitatively, it is 4.5×10^{-7}, 0.24, 52, and 2553 parts-per-million (ppm) for Gumbel, Exponential,

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/ACCESS.2021.3065869, IEEE Access

Lognormal, and GEV, respectively. This highlights the uncertainty of RTN prediction when using different CDFs.

On the applicability of the conclusions drawn here to other fabrication processes, ideally we should compare the results of samples fabricated by different processes. However, we only have one wafer from one company. The test samples used here were fabricated by the 28 nm CMOS process, which has been widely used commercially. The results reported here should be a typical representation of industrial processes, but further work will be needed to confirm this.

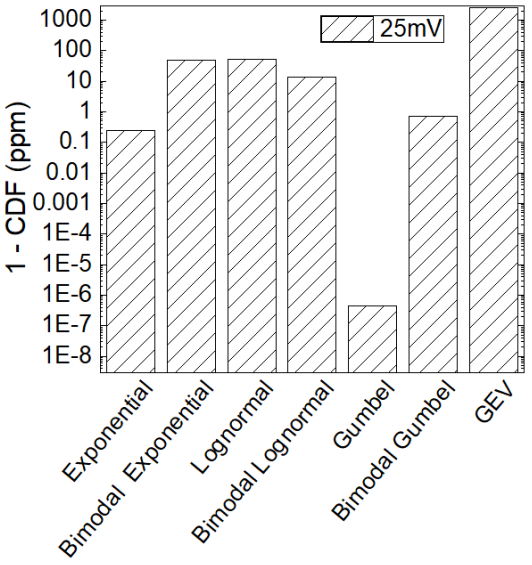

FIGURE 12. A comparison of the probability of occurrence for δVth≥25 mV predicted by different CDFs. The CDF values at δVth=25 mV were taken from Fig. 11, as marked out by the vertical dashed line.

D. NEW MODEL SELECTION CRITERION

Given the large uncertainties in the RTN predicted by different CDF models, further work is needed to justify their selection, in addition to their errors. Ideally, the selected model should be justified by device physics. Unfortunately, we could not link the Exponential, Gumbel, and GEV with a physical process, as these models are empirical [1]. GEV is developed from the extreme value theory to capture the long distribution tails, with Gumbel, Fréchet, and Weibull distributions as its special cases [32]. The number of traps in a device is minimized in a modern commercial CMOS process through quality control and one may consider that having a trap right above the narrow percolation current path in Fig. 1b is extremely rare.

The Lognormal CDF has been interpreted physically [6], [7]. As the number of charge carrier in the channel depends on (Vg-Vth) exponentially in the subthreshold region, a local Vth fluctuation spatially leads to an exponential fluctuation of local density of charge carrier, n. If Vth varies spatially by following Normal distribution, Log(n) will vary by following Normal distribution. The impact of a trapped charge on the channel is proportional to n, so that Log(δVth) will also follow the normal distribution, i.e. δVth follows Lognormal distribution [7].

There are, however, two difficulties with this interpretation. One is that Id was monitored above threshold in typical RTN tests, where n no longer depends on (Vg-Vth) exponentially. The other is that the impact of trapping on carrier mobility is neglected here [7]. It has been reported that the contribution of charge-induced mobility degradation is similar to that of carrier number reduction [41].

In searching for further criterion for model selection, we examine the dependence of error per trap on the number of traps. If the test data truly follow a specific CDF, we expect that the error per trap decreases with increasing number of traps, because an infinite number of data should produce this specific CDF perfectly. To support this statement, we used a theoretical Lognormal CDF to randomly generate a number of data and then treat them as 'test data'. These 'test data' were used to extract the Lognormal CDF and the errors were evaluated in the same way as that for the real test data. Fig. 13a shows that the SSE per trap indeed reduces for higher number of traps, despite of the statistical scattering. The same also applies to the GEV CDF.

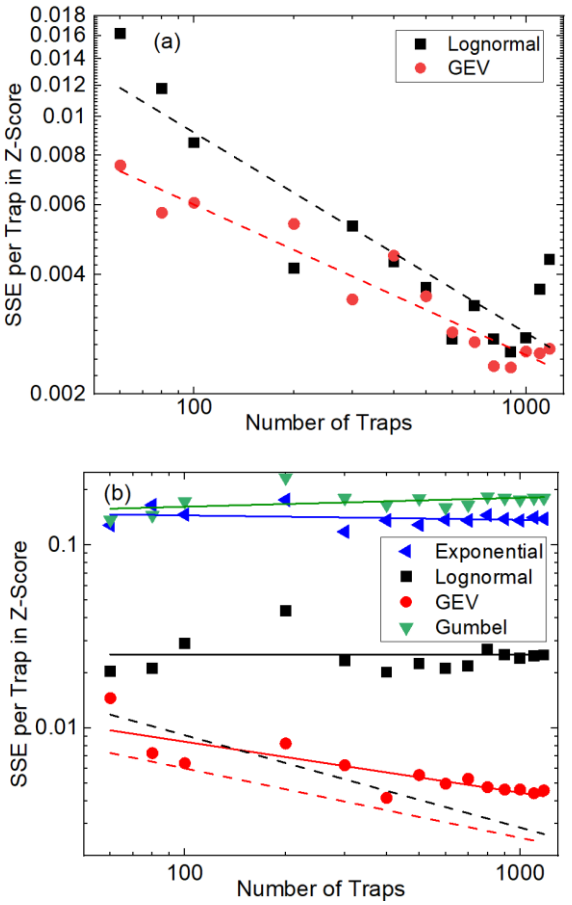

FIGURE 13. Dependence of SSE per trap on the number of traps used to extract the CDFs. (a) The data are generated randomly from the theoretical Lognormal (■) and GEV (●). They were treated as the 'test data' and used to extract Lognormal and GEV CDFs, respectively. Their SSE per trap decreases with increasing trap number, as shown by the fitted dashed lines. (b) The real test data were used to extract CDFs and calculate SSE per trap. The solid lines are fitted. For comparison, the two dashed lines in (a) were replotted in (b). Only GEV clearly shows the expected decrease of errors with increasing trap number.

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/ACCESS.2021.3065869, IEEE Access

Fig. 13b shows the dependence of SSE per trap on the trap number for the real test data. Only the error of GEV exhibits a clear decrease for higher number of traps. To quantitatively compare the error of theoretical and real test data, the two fitted lines in Fig. 13a were reproduced as the two dashed lines in Fig. 13b. The difference in the error between the theoretical and real test data is substantially larger for Lognormal, when compared with GEV distribution. This supports the GEV model.

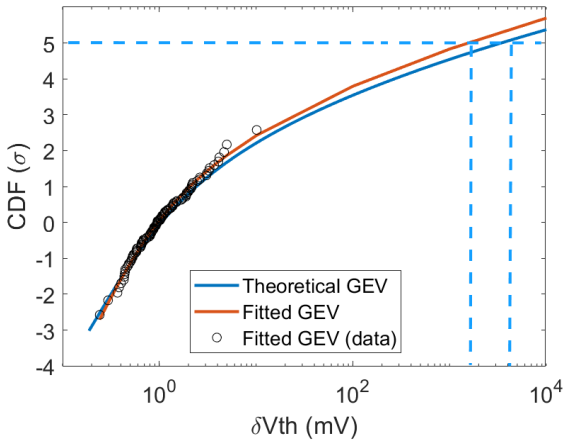

FIGURE 14. A comparison between theoretical GEV by using parameters in Table II and the GEV fitted by using 100 hypothetical traps, which were randomly generated by the theoretical GEV. The difference between theoretical and fitted GEV at 5σ is shown by the dashed lines.

E. IMPACT OF TRAP NUMBERS ON PREDICTION ACCURACY

Fig. 13a shows that the error per trap reduces for higher number of traps. In practice, the number of available traps is always limited. When the CDFs extracted from a limited number of traps is used to predict the RTN in the long tail, an important question is how accurate it is.

To assess the impact of trap number on this accuracy, one needs a reference distribution as the benchmark. Here, we use the GEV distribution extracted in Fig. 11 as the reference and their parameters are given in Table II. One set of 'N' data is randomly generated according to this distribution, as shown in Fig. 14. These N data are then used to extract the statistical distribution, which gives the orange curve in Fig. 14. The difference between the fitted and the reference distributions (the blue curve in Fig. 14) at a given σ can then be determined, as illustrated by the dashed lines in Fig. 14. By repeating this process 1000 times and each time with a different and randomly generated set of N data, we can obtain the confidence for the accuracy of statistical distributions extracted from a set of N data [42].

Fig. 15a shows the error at 3σ for different N. For N=100, the error at 90% confidence is -58.35% and 57.42%, respectively. For N=1,000, these two errors are reduced to -13.68% and 13.18%. If one targets an accuracy of 15% at 3σ with 90% confidence, 1,000 traps can be used.

Fig. 15b shows the errors for N=1000 at different sigma. The error increases from -13.68% and 13.18% at 3σ to -37.15% and 26.5% at 5σ (a probability of 0.57 parts per million) for 90% confidence. To be conservative, the guide-band for RTN induced δVth at 5σ should be increased by 26.5% from the value predicted by the statistical distribution extracted from 1000 traps, therefore.

With 1000 traps, the probability of occurrence for δVth≥25 mV is between 1584.2 and 3537.6 with 90% confidence. This uncertainty is substantially smaller than that from using different CDF models shown in Fig. 12. We conclude that the uncertainty in RTN amplitude prediction is dominated by model selection.

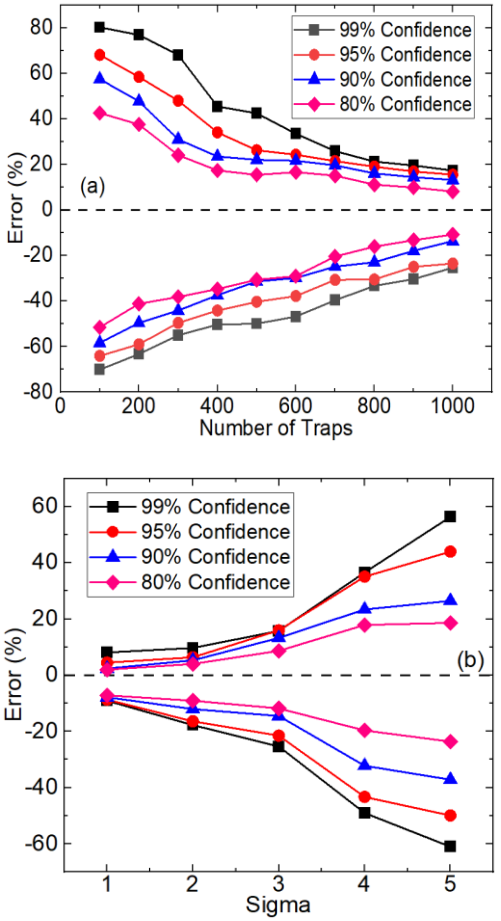

FIGURE 15. (a) shows the errors of prediction at 3σ by the CDFs extracted from different number of traps with 80%, 90%, 95% and 99% confidence. (b) shows the errors at different σ for 1,000 traps.

IV. CONCLUSION

This work assesses the accuracy of the statistical distributions for the RTN amplitude per trap. Its novelty includes proposing a new model selection criterion based on the relation between error and trap number, exploring the applicability of a wide range of statistical distributions to RTN amplitudes, and finding that the Generalized Extreme Value (GEV) distribution has the least Z-score based error. The new model selection criterion requires a monotonic error decrease for

higher number of traps. The GEV meets this criterion, while the Exponential, Lognormal, and Gumbel distributions do not. Based on our data, using bimodal Exponential and Lognormal CDFs only has a modest impact on the error, despite the increased fitting parameters.

The accuracy of CDF extracted from a limited number of traps is also assessed. For 90% confidence, the guide-band for RTN induced δVth at 5σ should be increased by 26.5% from the value predicted by the statistical distribution extracted from 1,000 traps. The uncertainties caused by using a limited number of traps is relatively small and the selection of CDF model dominates the uncertainty in RTN amplitude prediction and modelling.

ACKNOWLEDGMENT

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

REFERENCES

[1]. M. J. Kirton and M. J. Uren, "Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency ($1/f$) noise", *Advances in Physics*, vol. 38, no. 4, pp. 367-468, 1989, doi: 10.1080/00018738900101122.

[2]. A. Asenov, R. Balasubramaniam, A. R. Brown, and 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.

[3]. 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.

[4]. T. H. Both, G. Firpo Furtado, and G. I. Wirth,, "Modeling and simulation of the charge trapping component of BTI and RTS," *Microelectron. Rel.*, vol. 80, pp. 278-283, 2018, doi:10.1016/j.microrel.2017.11.009.

[5]. R. Wang, S. Guo, Z. Zhang, Q. Wang, D. Wu, J. Wang, and R. Huang, "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]. Z. Zhang S. Guo, X. Jiang, R. Wang, Z. Zhang, P. Hao, Y. Wang, and R. Huang, "New Insights into the Amplitude of Random Telegraph Noise in Nanoscale MOS Devices," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, April 2017, pp.3C-3.1 – 3C-3.5, doi: 10.1109/IRPS.2017.7936288.

[7]. K. Sonoda, K. Ishikawa, T. Eimori, and 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.

[8]. B. Zimmer, O. Thomas, S. O. Toh, T. Vincent, K. Asanovic, and B. Nikolic, "Joint Impact of Random Variations and RTN on Dynamic Writeability in 28nm Bulk and FDSOI SRAM," in *Proc. European Solid State Device Research Conference (ESSDERC)*, Sept. 2014, pp.98-101, doi:10.1109/ESSDERC.2014.6948767.

[9]. C. Y. P. Chao, H. Tu, T. M. H. Wu, K. Y. Chou, S. F. Yeh, C. Yin, and C. L. Lee, "Statistical Analysis of the Random Telegraph Noise in a 1.1 μm Pixel, 8.3 MP CMOS Image Sensor Using On-Chip Time Constant Extraction Method," *Sensors*, vol. 17, p.2704, 2017, doi: 10.3390/s17122704.

[10]. K. Ito, T. Matsumoto, S. Nishizawa, H. Sunagawa, K. Kobayashi, and 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.

[11]. S. Ichino, T. Mawaki, A. Teramoto, R. Kuroda, S. Wakashima, T. Suwa, and S. Sugawa, "Statistical Analyses of Random Telegraph Noise in Pixel Source Follower with Various Gate Shapes in CMOS Image Sensor," *ITE Trans. on Media Technology and Applications*, vol. 6, no. 3, pp.163-170, 2018, doi: 10.3169/mta.6.163.

[12]. F. M. Puglisi, F. Costantini, B. Kaczer, L. Larcher, and P. Pavan, "Monitoring stress-induced defects in HK/MG FinFETs using random telegraph noise," *IEEE Electron Device Lett.*, vol. 37, no. 9, pp. 1211-1214, 2016, doi: 10.1109/LED.2016.2590883.

[13]. L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989, doi: 10.1109/5.18626

[14]. F.M. Puglisi and P. Pavan, "Factorial hidden Markov model analysis of random telegraph noise in resistive random access memories," *ECTI Trans. Electr. Eng. Electr. Commun.*, vol. 12, no. 1, pp. 24–29 2014.

[15]. F. M. Puglisi, L. Larcher, A. Padovani, and P. Pavan, "A complete statistical investigation of RTN in HfO2-based RRAM in high resistive state," *IEEE Trans. Electron Devices*, vol. 62, no. 8, pp. 2606-2613, 2015, doi: 10.1109/TED.2015.2439812.

[16]. A.K.M. M. Islam and H. Onodera, "Worst-case Performance Analysis Under Random Telegraph Noise Induced Threshold Voltage Variability," in *Proc. 28th Int. Symp. Power and Timing Modeling, Optimization and Simulation (PATMOS)*, pp.140-146, 2018, doi: 10.1109/PATMOS.2018.8464147.

[17]. A. Bal, S. Roy, K. Chakraborty, "Trident: Comprehensive Choke Error Mitigation in NTC Systems," *IEEE Trans. VLSI*, vol. 26, no. 11, pp. 2195-2204, 2018, doi: 10.1109/TVLSI.2018.2863954.

[18]. M. Mehedi, K. H. Tok, J. F. ZHANG, Z. JI, Z. Ye, W. Zhang, and J. S. Marsland, "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.

[19]. K. K. Hung, P. K. Ko, C. Hu, and Y. C. Cheng, "A Unified Model for the Flicker Noise in Metal-Oxide-Semiconductor Field-Effect Transistors," *IEEE Trans. Electron Devices*, vol. 37, no. 3, pp. 654-665, 1990, doi: 10.1109/16.47770.

[20]. Z. Çelik-Butler, S. P. Devireddy, H. Tseng, P. Tobin, and A. Zlotnicka, "A low-frequency noise model for advanced gate-stack MOSFETs," *Microelectron. Rel.*, vol. 49 pp. 103–112, 2009, doii:10.1016/j.microrel.2008.12.005.

[21]. M. Nour, Z. Çelik-Butler, A. Sonnet, F. C. Hou, S. Tang, and G. Mathur, "A stand-alone, physics-based, measurement-driven model and simulation tool for random telegraph signals originating from experimentally identified MOS gate-oxide defects," *IEEE Trans. Electron Devices*, vol. 63, no. 4, pp. 1428–1436, Apr. 2016, doi: 10.1109/TED.2016.2528218.

[22]. R. Gao Z. Ji, A. B. Manut, J. F. Zhang, J. Franco, S. W. M. Hatta, W. Zhang, B. Kaczer, D. Linten, and G. Groeseneken, "NBTI-Generated Defects in Nanoscaled Devices: Fast Characterization Methodology and Modeling," *IEEE Trans. Electron Devices*, vol. 64, no. 10, pp.4011-4017, 2017, doi:10.1109/TED.2017.2742700.

[23]. 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.

[24]. A. Manut, R. Gao, J. F. Zhang , Z. Ji , M. Mehedi, W. Zhang, D. Vigar, A. Asenov, and B. Kaczer, "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.

[25]. K. Fukuda, Y. Shimizu, K. Amemiya, M. Kamoshida, and C. Hu, "Random Telegraph Noise in Flash Memories—Model and Technology Scaling", in *IEDM Tech. Dig.*, Dec. 2007, pp. 169–172, doi: 10.1109/IEDM.2007.4418893

[26]. T. Nagumo, K. Takeuchi, S. Yokogawa, K. Imai, and Y. Hayashi, "New Analysis Methods for Comprehensive Understanding of

Random Telegraph Noise," in *IEDM Tech. Dig.*, Dec. 2009, pp. 759–762, doi: 10.1109/IEDM.2009.5424230.

[27]. H. Qiu, K. Takeuchi, T. Mizutani, T. Saraya, J. Chen, M. Kobayashi, and T. Hiramoto, "Statistical Analyses of Random Telegraph Noise Amplitude in Ultra-Narrow (Deep Sub-10nm) Silicon Nanowire Transistors," in *Proc. Symp. VLSI Technol.*, Jun. 2017, pp. 50–51.

[28]. B. Kaczer, T. Grasser, Ph. J. Roussel, J. Franco, R. Degraeve, L.-A. Ragnarsson, E. Simoen, G. Groeseneken, and H. Reisinger "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.

[29]. M. Toledano-Luque, B. Kaczer, E. Simoen, Ph. J. Roussel, A. Veloso, T. Grasser, and G. Groeseneken, "Temperature and voltage dependences of the capture and emission times of individual traps in high-k dielectrics," *Microelectronic Eng.*, vol. 88, pp. 1243-1246, 2011, doi: doi:10.1016/j.mee.2011.03.097.

[30]. S. Realov and K. L. Shepard, "Analysis of Random Telegraph Noise in 45-nm CMOS Using On-Chip Characterization System," *IEEE Trans. Electron Devices*, vol. 60, no. 5, pp.1716-1722, 2013, doi: 10.1109/TED.2013.2254118.

[31]. Z. Zhang S. Guo, X. Jiang, R. Wang, R. Huang, and J. Zou, "Investigation on the amplitude distribution of random telegraph noise (RTN) in nanoscale MOS devices," in *Proc. IEEE Int. Nanoelectron. Conf.*, Chengdu, China, 2016, pp. 5–6.

[32]. D. McFadden, "Modeling the Choice of Residential Location," *Transportation Research Record*, vol. 673, pp. 72–77, 1978.

[33]. M. Duan, J. F. Zhang, Z. Ji, J. G. Ma, W. Zhang, B. Kaczer, T. Schram, R. Ritzenthaler, G. Groeseneken, and A. Asenov, "Key issues and techniques for characterizing Time-dependent Device-to-Device Variation of SRAM," in *IEDM Tech. Dig.*, Dec. 2013, pp. 774–777, doi: 10.1109/IEDM.2013.6724730.

[34]. T. Grasser, K. Rott, H. Reisinger, M. Waltl, J. Franco, and B. Kaczer, "A Unified Perspective of RTN and BTI," in *IEEE Proc. Int. Rel. Phys. Symp. (IRPS)*, April 2014, pp.4A.5.1-4A.5.7, doi: 10.1109/IRPS.2014.6860643.

[35]. X. Zhan, C. Shen, Z. Ji, J. Chen, H. Fang, F. Guo, and J. F. Zhang, "A Dual-Point Technique for the Entire ID–VG Characterization Into Subthreshold Region Under Random Telegraph Noise Condition," *IEEE Electron Device Lett.*, vol. 40, no. 5, pp.674-677, 2019, doi: 10.1109/LED.2019.2903516.

[36]. G. C. Chow, "Maximum-likelihood estimation of misspecified models," *Economic Modelling*, vol.1, Issue. 2, pp.134-138, 1984, doi: 10.1016/0264-9993(84)90001-4.

[37]. S. Realov and K. L. Shepard, "Random telegraph noise in 45-nm MOS: Analysis using an on-chip test and measurement system," in *IEDM Tech. Dig.*, Dec., 2010, pp. 28.2.1-28.2.4, doi: 10.1109/IEDM.2010.5703436

[38]. E. Wit, E. van den Heuvel, and J. W. Romeijn, "'All models are wrong...': an introduction to model uncertainty," *Statistica Neerlandica*, vol.66, no.3, pp.217–236, 2012, doi:10.1111/j.1467-9574.2012.00530.x.

[39]. The MathWorks, I., 2019. *Statistics and Machine Learning Toolbox*, Natick, Massachusetts, United State. Available at: https://www.mathworks.com/help/stats/

[40]. J. Heyes, "Basic Tools, Normal Probability Plots," *ASQC Statistics Division Newsletter*, vol. 13, no. 1, 1992.

[41]. K. K. Hung P. K. Ko, C. Hu, and Y. C. Cheng, "Random Telegraph Noise of Deep-Submicrometer MOSFET's," *IEEE Electron Device Lett.*, vol. 11, no. 2, pp.90-92, 1990, doi: 10.1109/55.46938.

[42]. M. Duan, J. F. Zhang, Z. Ji, W. Zhang, B. Kaczer, and A. Asenov, "Key issues and solutions for characterizing hot carrier aging of nano-meter scale nMOSFETs," *IEEE Trans. Electron Devices*, vol. 64, no. 6, pp.2478-2484, 2017, doi: 10.1109/TED.2017.