



Manut, A., Gao, R., Zhang, J. F., Ji, Z., Mehedi, M., Zhang, W. D., Vigar, D., Asenov, A. and Kaczer, B. (2019) Trigger-when-charged: a technique for directly measuring RTN and BTI-induced threshold voltage fluctuation under use-Vdd. IEEE Transactions on Electron Devices, 66(3), pp. 1482-1488.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/183480/>

Deposited on: 4 April 2019

Enlighten – Research publications by members of the University of Glasgow\_  
<http://eprints.gla.ac.uk>

# Trigger-When-Charged: A technique for directly measuring RTN and BTI-induced threshold voltage fluctuation under use-Vdd

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

**Abstract** — Low power circuits are important for many applications, such as IoT. Device variations and fluctuations are challenging their design. Random telegraph noise (RTN) is an important source of fluctuation. To verify a design by simulation, one needs assessing the impact of fluctuation in both driving current,  $\Delta I_d$ , and threshold voltage,  $\Delta V_{th}$ . Many early works, however, only measured RTN-induced  $\Delta I_d$ .  $\Delta V_{th}$  was not directly measured because of two difficulties: its average value is low and it is highly dynamic. Early works often estimated  $\Delta V_{th}$  from  $\Delta I_d/g_m(V_g=V_{dd})$ , where  $g_m$  is trans-conductance, without giving its accuracy. The objective of this work is to develop a new Trigger-When-Charged (TWC) technique for directly measuring the RTN-induced  $\Delta V_{th}$ . By triggering the measurement only when a trap is charged, measurement accuracy is substantially improved. It is found that there is a poor correlation between  $\Delta I_d/g_m(V_g=V_{dd})$  and the directly measured  $\Delta V_{th}(V_g=V_{th})$ . The former is twice of the latter on average. The origin for this difference is analyzed. For the first time, the TWC is applied to evaluate device-to-device variations of the directly measured RTN-induced  $\Delta V_{th}$  without selecting devices.

Index terms: Random telegraph noise (RTN), Fluctuations, Yield, Within-a-device-fluctuation, Jitters, Positive charges, NBTI.

## I. INTRODUCTION

As CMOS nodes scale down, the fluctuations induced by random charge-discharge of traps scale up. Smaller devices have larger statistical spread because of fewer traps per device and the larger impact of a single charge on them [1,2]. The increased number of devices per chip also leads to larger statistical spread [1,2] and high data transmission rate requires tight control of fluctuations [3]. Fluctuations have become a major concern for circuit design and have attracted many attentions recently [4-20]. It has been reported that current fluctuation in some fresh devices can be over the typical device lifetime criterion of 10% [5].

Fluctuations are commonly observed as the random telegraph noise (RTN) in the drain current,  $\Delta I_d$ , under a given

gate bias,  $V_g$ , and early works [5-13] have focused on them.  $\Delta I_d$  allows probing individual traps and an analysis of their mean capture and emission time dependence on  $V_g$  gives the trap energy and spatial locations [5, 6, 8, 10]. This has improved our understanding substantially. There are, however, little direct measurements of the RTN-induced fluctuation in threshold voltage,  $\Delta V_{th}$ . This is because its measurement is difficult: the charge-discharge of traps for RTN is highly dynamic and the average  $\Delta V_{th}$  is typically low. As a result, the RTN-induced  $\Delta V_{th}$  often was either not given [5,11] or estimated from dividing  $\Delta I_d$  by trans-conductance, i.e.  $\Delta V_{th} \approx \Delta I_d/g_m(V_{dd})$  [6-10]. The accuracy of the  $\Delta V_{th}$  evaluated in this way was not given in these works [6-10].

To model the impact of RTN on the margin of SRAM [15] and the timing error [14], one needs both  $\Delta I_d$  and  $\Delta V_{th}$ . For example, RTN in the pass transistor 1 in Fig. 1a can reduce the driving current by  $\Delta I_d$  and slow down the  $V_g$  rise of transistor 2 in reaching its threshold voltage,  $V_{th0}$ , by  $\Delta t(\Delta I_d)$ . RTN in the transistor 2 can increase its  $V_{th}$  by  $\Delta V_{th}$  and results in a further delay,  $\Delta t(\Delta V_{th})$ . There is a need to obtain both accurate  $\Delta I_d$  and  $\Delta V_{th}$ , therefore.

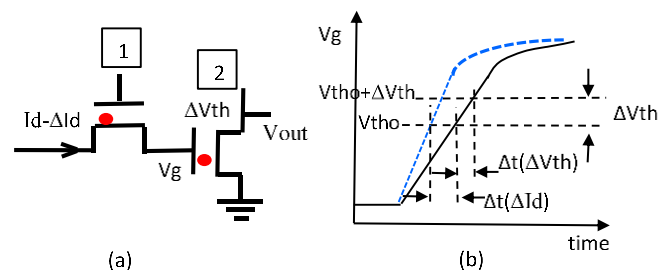


Fig. 1. A schematic illustration of the impact of  $\Delta I_d$  and  $\Delta V_{th}$  on timing: (a) circuits and (b) waveform.  $V_{out}$  switches when  $V_g \approx V_{th}$ , which is delayed by a lower charging current,  $I_d - \Delta I_d$ , supplied through the transistor 1 and a higher  $V_{th} = V_{th0} + \Delta V_{th}$  of the transistor 2.

The objective of this work is to develop a new Trigger-When-Charged (TWC) technique for directly measuring the RTN-induced  $\Delta V_{th}$ . By ensuring that the measurement is taken when traps are charged, the accuracy is substantially improved. It is found that the  $\Delta I_d/g_m(V_{dd})$  correlates poorly with the directly measured  $\Delta V_{th}$  and the former doubles the latter on

Manuscript received MMM DDD, YYYY. This work was supported by the Engineering and Physical Science Research Council of UK under the grant no. EP/L010607/1. The review of this paper was arranged by Editor Z. Çelik-Butler (Corresponding author: R. Gao and J. F. Zhang)

A. Manut, R. Gao, J. F. Zhang, Z. Ji, M. Mehedi, W. Zhang, are with the Department of Electronics and Electrical Engineering, John Moores University, Liverpool L3 3AF, UK. A. Manut now is with Faculty of Electrical Engineering, Universiti Teknologi MARA Selangor, 40450 Shah Alam, Malaysia.

D. Vigar is with Qualcomm Technologies International Ltd., Cambridge CB4 0WZ, U.K.

A. Asenov is with the Department of Electronics and Electrical Engineering, University of Glasgow, Glasgow, G12 8QQ, U.K.

B. Kaczer is with Imec, B-3001 Leuven, Belgium.

average. The discrepancy originates partly from the device-to-device variation (DDV) of relative local current density beneath a trap at  $V_g=V_{th}$  [16-19] and partly from the charge-induced mobility degradation [20].

Some deeply scaled devices have analyzable RTN signals in terms of extracting mean capture/emission time [11], while others can have a complex within-a-device-fluctuation [12]. The latter was deselected in some early works [10,13,16,17], making the real DDV of fluctuation unobtainable. The TWC developed in this work is applicable to devices with or without analyzable RTN signals and it will be used to evaluate the DDV.

## II. DEVICES AND MEASUREMENT TECHNIQUE

### A. Devices

The MOSFETs used in this work were fabricated by a 28 nm commercial CMOS process with a use  $V_{dd}$  of 0.9 V. They have a metal gate and a high-k dielectric stack with an equivalent oxide thickness of 1.2 nm. The channel width and length are 135 nm and 27 nm, respectively. For comparison purpose, large devices of  $3 \times 1 \mu\text{m}$  were also used, which has insignificant DDV. All tests were performed at 125 °C.

### B. TWC technique

Difficulties with standard measure-stress-measure methods: For ageing-induced  $\Delta V_{th}$  under stresses such as negative bias temperature instability (NBTI) [21,22] and hot carriers [23,24], the degradation is commonly measured at preset time. This is acceptable, as the  $V_g$ -acceleration used in the stress generally leads to a large-enough  $\Delta V_{th}$  that is measurable and deterministic at a preset time. There are, however, two difficulties in applying this method to deeply scaled devices under use- $V_{dd}$ , where  $\Delta V_{th}$  mainly exhibits as Random Telegraph Noise (RTN). First, there are only a few active traps and the average  $\Delta V_{th}$  is typically low. Second, charge-discharge of these traps are highly dynamic: they are often neutral at the preset time for measurement, as shown by the red circle symbols in Fig. 2, and would be missed by the measurement.

One way to avoid these difficulties is selecting devices that only have one trap, which induces a high enough  $\Delta V_{th}$  (e.g. 20 mV) and its emission time is long enough (e.g. >1 sec) for completing the measurement [16,17]. This has improved our understanding of the interaction between a trap and the current. Such devices, however, are rare (e.g. ~10% [16]) and the required device selection precludes obtaining real DDV. The present work develops a new technique that removes the device selection and is applicable to all devices, so that the real DDV can be extracted.

Test procedure of TWC technique: Fig. 2a gives the  $V_g$  waveform. After recording the reference  $I_d$ - $V_g$  on a fresh device, the test starts by a ‘stabilization’ period of 40 sec under  $V_g=V_{dd}=-0.9$  V. If there are any traps at deep energy level in a device, they will be filled during this period [25].  $\Delta I_d$  under

$V_g=-0.9$  V is then monitored for a period, e.g. 100 sec, as marked by ‘Id monitor’ in Fig. 2a. A sampling rate of 1 M/sec was used [26]. The trapping-induced up-envelope (UE) of  $\Delta I_d$  is obtained.

To measure the trapping-induced  $\Delta V_{th}$ , one must ensure that the measurement was taken when the traps are charged. This is achieved by setting the trigger level of the oscilloscope and the pulse generator for  $V_g$  just below the UE, as shown in Figs. 2a&b. Once triggered, the pulse  $I_d$ - $V_g$  (p-IV) is recorded in 3  $\mu\text{s}$  to minimize discharge [25,26].

Although a sampling rate of 1 M/sec can be used to monitor  $\Delta I_d$  under a fixed  $V_g=-0.9$  V, it only gives 3 points in 3  $\mu\text{s}$  and is too slow for the p-IV. To have sufficient number of points for p-IV, a higher rate of 100 M/sec is used. The p-IV was repeatedly measured for 50 times and their average is used to reduce the system noise to ~1 mV.

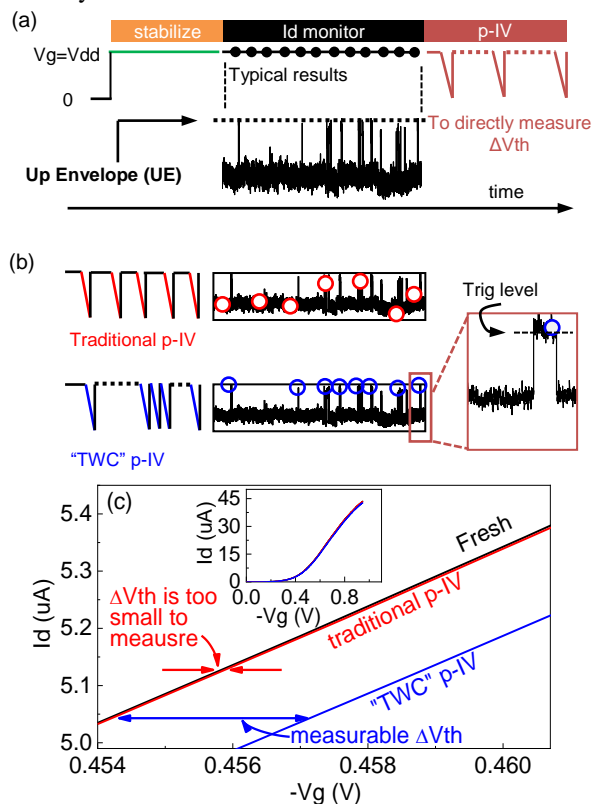


Fig. 2. The TWC technique. (a) Test procedure: After a stabilization period, the RTN-induced  $\Delta I_d$  is monitored under  $V_g=V_{dd}$  and the upper envelope (UE) is determined. The trigger-level for subsequent p-IV (3  $\mu\text{s}$ ) is then set just below UE to measure  $\Delta V_{th}$ . 50 p-IVs were measured in (b) and their average is given in (c). The TWC p-IV captures the RTN-induced  $\Delta V_{th}$ , while the traditional p-IV at pre-set time often misses the charge and is inapplicable. Both the TWC and traditional p-IVs were taken after the ‘stabilize’ period.

$\Delta V_{th}$  is evaluated from the difference between the TWC p-IV and the reference p-IV. The reference p-IV was obtained also from the average of 50 p-IV with the same sweep rate, performed on fresh devices before filling the energetically deep traps by applying the waveform in Fig. 2a. When measuring these 50 p-IV, it is possible that a trap can be filled during the measurement. These outlier p-IVs were excluded from the reference p-IV. This ensures capturing the  $\Delta V_{th}$  induced by both RTN and energetically deep traps, if they are present. In

case that one is interested in capturing RTN-induced  $\Delta V_{th}$  only, the reference p-IV should be taken after filling the energetically deep traps. Fig. 2c demonstrates that a single trap induced  $\Delta V_{th}$  of  $\sim 2$  mV is successfully captured by the TWC technique, which often would be missed by the traditional p-IV recorded at a preset time, as illustrated by the red circles in Fig. 2b. The measured  $\Delta V_{th}/\Delta I_d$  ratio is used to convert  $\Delta I_d$  to  $\Delta V_{th}$ .

**Measurement setup:** As the main objective of this work is to develop a technique for measuring the RTN-induced  $\Delta V_{th}$  under use  $V_{dd}$ , the detailed measurement setup is given in Fig. 3.  $I_d$  under  $V_d=0.1$  V was converted to a voltage,  $V_{out}$ , by a home-made operational amplifier circuit. During the ‘ $I_d$  monitor’ phase in Fig. 2a,  $V_{out}$  was monitored by both channels 2 and 3 of an oscilloscope and one example is given in Fig. 3b.

In the following ‘p-IV’ phase of Fig. 2a, when  $V_{out}$  is above the ‘trigger level’ in Fig. 3c, the oscilloscope triggers and simultaneously sends out a signal to trigger the pulse generator for  $V_g$ . Both the pulse applied to the gate and the corresponding  $V_{out}$  are captured, as shown in Fig. 3c. Two channels are needed here: channel 3 is at a fine scale to ensure capturing the small  $V_{out}$  fluctuation with good accuracy and channel 2 is switched to a coarse scale to capture the whole p-IV. As a comparison, Fig. 3d shows an example triggered at a preset time that missed the trapped charge.

The UE in Fig. 2a can be caused by either a single trap or multiple traps. In the latter case, the UE results from the combined charges of multiple traps. This removes the need for selecting devices of a single trap and makes the method applicable to all devices.

The differences of this work from the typical BTI tests are that the p-IVs are only triggered when traps being charged and  $V_g$ -acceleration is not used here.

### III. RESULTS AND DISCUSSIONS

#### A. A comparison between $\Delta I_d/gm(V_{dd})$ and $\Delta V_{th}(V_{th})$

As mentioned in the introduction, early works [6-10] often estimated  $\Delta V_{th}$  by  $\Delta I_d/gm(V_{dd})$ , where both  $\Delta I_d$  and  $gm$  were obtained under  $V_g=V_{dd}$ . This is effectively measuring the shift of IV at  $V_g=V_{dd}$ , as marked by the point ‘B’ in Fig. 4a and the corresponding inset. The real  $\Delta V_{th}$ , however, should be evaluated from  $V_g=V_{th}$  at the point ‘A’ in Fig. 4a. In this work,  $V_{th}$  is extracted by extrapolating from the maximum  $gm$  point and  $V_{th}=-0.45$  V in Fig. 4a. The shift in  $V_{th}$ ,  $\Delta V_{th}$ , at a given sensing  $V_g$  is evaluated from  $\Delta I_d/gm(V_{gsense})$ . We now compare the  $\Delta V_{th}$  evaluated at  $V_{gsense}=V_{th}$  (‘A’ in Fig. 4a) with that at  $V_{gsense}=V_{dd}$  (‘B’ in Fig. 4a).

Fig. 4b plots  $\Delta V_{th}(V_{th})$  against  $\Delta V_{th}(V_{dd})=\Delta I_d/gm(V_{dd})$  measured on 63 devices. Both of them have a large DDV, but the correlation between them is poor. For similar  $\Delta I_d/gm(V_{dd})$ ,  $\Delta V_{th}$  can spread from its minimum to its maximum approximately. As a result, errors are large if  $\Delta I_d/gm(V_{dd})$  is used as  $\Delta V_{th}$ , so that it is essential to measure  $\Delta V_{th}$  directly at  $V_g=V_{th}$ . Although both of them have maximum close to the typical device lifetime definition of 30~50 mV, the average  $\Delta I_d/gm(V_{dd})$  doubles that of  $\Delta V_{th}$ , as shown by the two dashed lines in Fig. 4b. This is because many devices have  $\Delta V_{th}(V_{th})$

close to zero, but  $\Delta I_d/gm(V_{dd})$  are above 10 mV. The origin of the differences between these two will be analyzed next.

#### B. Effects of sensing $V_g$ on $\Delta V_{th}$

In Fig. 4a, the sensing  $V_g$  for  $\Delta V_{th}$  is -0.9 V for the point B and  $V_{th}=-0.45$  V for the point A. Since the whole  $I_d\sim V_g$  was measured, one can also extract the ‘‘apparent  $\Delta V_{th}$ ’’ at other sensing  $V_g$  by using  $\Delta I_d/gm(V_{gsense})$ . The ‘‘apparent  $\Delta V_{th}$ ’’ here is referred to the  $\Delta V_{th}$  evaluated in this way under  $V_{gsense}\neq V_{th}$ . Typical examples obtained from different devices are given in Figs. 5a-e.

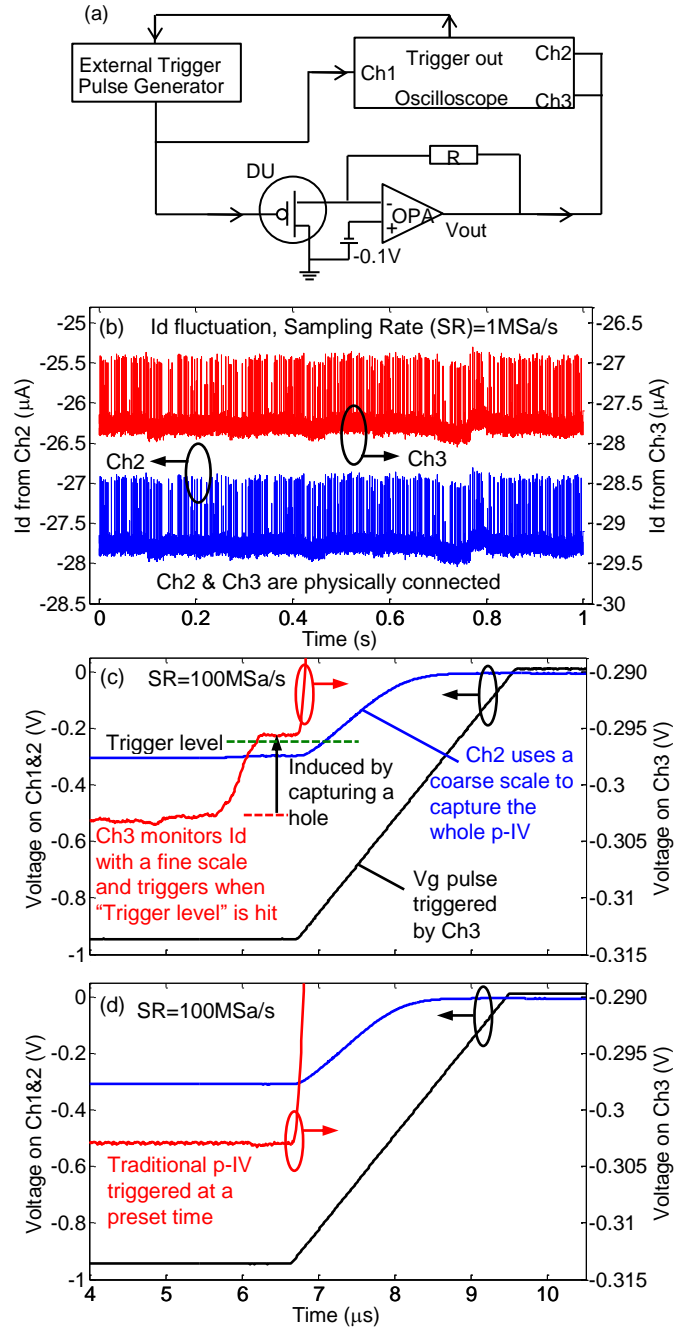


Fig. 3. (a) Test configuration for the TWC measurement technique. A high-speed operational amplifier based circuit is used to convert  $I_d$  to  $V_{out}$  that is connected to both channels 2 and 3. The ‘‘Trigger out’’ of the oscilloscope is

connected to the “External trigger in” of the pulse generator. (b) The  $V_{out}$  fluctuation is captured by both channels 2 and 3, as they are physically connected. (c) A screen-shot of the TWC p-IV measurement waveform. Channel 3 keeps its fine scale for accurate triggering, while channel 2 is switched to a coarse scale to capture the whole “TWC” p-IV. (d) A screen-shot of the traditional p-IV measurement at a preset time, where the trapped charge is missed.

The dependence of the apparent  $\Delta V_{th}$  on the sensing  $V_g$  has strong DDV, agreeing with that observed for single traps [16,17]. On one hand, Fig. 5a corresponds to Fig. 4a, where  $\Delta V_{th}$  increases monotonically with  $|V_g|$  and  $\Delta V_{th}$  at  $|V_g|=0.9$  V is 6 times of the real  $\Delta V_{th}(V_{th})$ . On the other hand,  $\Delta V_{th}$  can also reduce by almost half over the same voltage range, as shown in Fig. 5b. There are also cases where (i)  $\Delta V_{th}$  is almost a constant (Fig. 5c); (ii)  $\Delta V_{th}$  increases initially and then reduces (Fig. 5d); and (iii)  $\Delta V_{th}$  decreases initially and then increases (Fig. 5e).

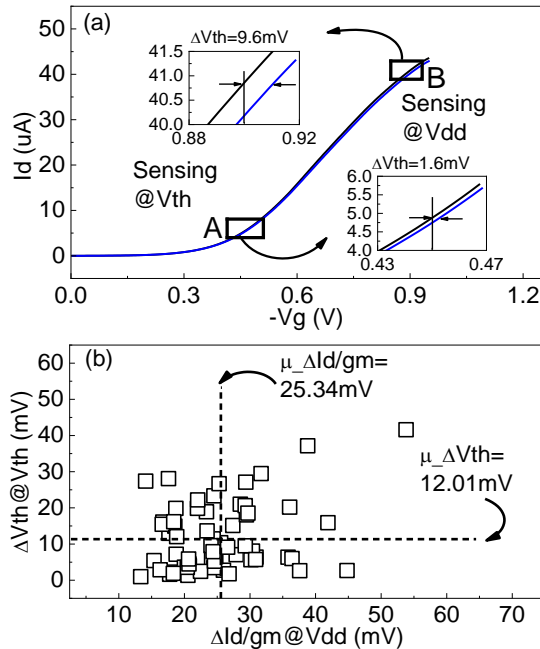


Fig. 4. (a) Early works estimated RTN-induced  $\Delta V_{th}$  from  $\Delta I_d/g_m$  at  $V_{dd}=0.9$  V (Point ‘B’), rather than directly measuring it at  $V_g=V_{th}$  (Point ‘A’). The two insets are enlarged p-IV around the two points. The black p-IV is reference and the blue p-IV is the TWC p-IV. (b) The poor correlation between  $\Delta I_d/g_m$  at  $V_{dd}$  and  $\Delta V_{th}$  at  $V_g=V_{th}$ . Each point was taken from a different device. The dotted lines mark the mean values.

It is known that channel current can have a narrow percolation path near  $V_{th}$  and the impact of a charged trap on a deeply scaled device depends on the relative local current density beneath the trap [16-19]. This can explain the device-specific dependence observed in Fig. 5. As schematically illustrated in Fig. 6, for the device in Fig. 5a, the trap is located far away from the current percolation path at  $V_{th}$ , so that it has little impact and  $\Delta V_{th}(V_{th})$  is low. The many close-to-zero  $\Delta V_{th}(V_{th})$  points in Fig. 4b indicates that this is often the case. As  $V_g$  increases, the current becomes more evenly spread and its relative density under this trap rises, leading to the increase of  $\Delta V_{th}$  with  $V_g$ . As there is current flowing beneath each trap

at  $V_{dd}$ , there is no close-to-zero apparent  $\Delta V_{th}$  in Fig. 4b, when evaluated by  $\Delta I_d/g_m(V_{dd})$ .

For the device in Fig. 5b, however, the trapped charge is on top of the current percolation path at  $V_{th}$ , resulting in a large  $\Delta V_{th}$  at  $V_{th}$ . As  $V_g$  increases, the current path is widened, so that the impact of the same charge on the device reduces and the  $\Delta V_{th}$  decreases with  $|V_g|$  in Fig. 5b. Similarly, the relative current density under the trap in Fig. 5c changes little with  $V_g$  and  $\Delta V_{th}$  is insensitive to  $V_g$ . The dependence of relative current density under a trap on  $V_g$  may not be monotonic, which can explain the behavior in Figs. 5d&e. For instance, in Fig. 5d, it may increase initially and then decrease. Alternatively, when there are multiple traps, some can behave like Fig. 5a and some like Fig. 5b. A combination of them can give the complex dependence in Figs. 5d&e.

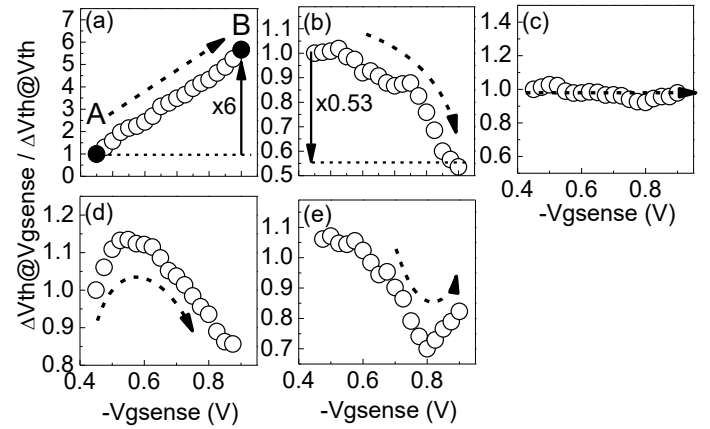


Fig. 5. Examples of the device specific dependence of the apparent  $\Delta V_{th}$  on the sensing  $V_g$ . (a)-(e) were obtained from five different devices. The apparent  $\Delta V_{th}$  at a  $V_{gsense}$  was obtained from the shift of TWC p-IV from the reference at  $V_{gsense}=V_{th}$ . The  $\Delta V_{th}$  is normalized against its value at  $V_{gsense}=V_{th}$ . As the lowest  $|V_{gsense}|$  is close to  $V_{th}$ , the data starts from  $\sim 1$  in all devices.

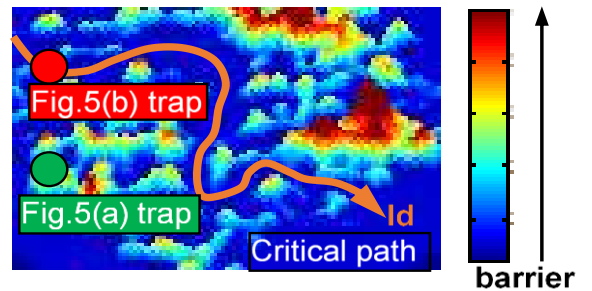


Fig. 6. A schematic illustration of different impacts of traps at different locations on a device at threshold condition. The current can follow a percolation path under  $V_g=V_{th}$ . The trap in green corresponds to the device in Fig. 5a: it is away from the critical current path, so that it only has a small effect on the device at  $V_{th}$ . The trap in red corresponds to the device in Fig. 5b: it is on top of the current critical path and has a large effect on the device at  $V_{th}$ .

Although the deeply scaled device-specific dependence of  $\Delta V_{th}$  on sensing  $V_g$  can be explained by the interaction between the trap and the relative local current density beneath it, there is also a device independent  $\Delta V_{th}$  dependence on the sensing  $V_g$ . For a large  $3 \times 1 \mu m$  device where DDV is insignificant, Fig. 7a shows that  $\Delta V_{th}$  also increases with

$|V_{gsense}|$ . On one hand, a more evenly distributed  $I_d$  at higher  $|V_{gsense}|$  allows more traps making an effective impact. On the other hand, the charge induced Coulombic scattering causes mobility degradation [27,28], which lead to  $\Delta I_d(\text{mobility})$ . When the apparent  $\Delta V_{th}$  is evaluated from  $\Delta I_d(\text{measured})/g_m$ , the  $\Delta I_d(\text{mobility})$  is treated as if it was caused by  $\Delta V_{th}$ . In other words, the apparent  $\Delta V_{th} = \Delta I_d(\text{measured})/g_m$  includes the contribution from mobility degradation to  $\Delta I_d$ . As the effect of mobility degradation increases with  $|V_{gsense}|$ , it contributes to the increase in the apparent  $\Delta V_{th}$  for higher  $|V_{gsense}|$ .

### C. Statistics

As there is hardly any information on the statistical properties of the directly measured RTN-induced  $\Delta V_{th}$ , especially in terms of its dependence on  $V_{gsense}$ , we report the DDV of this dependence here. Each line in Fig. 7b represents one device and the first impression is that the apparent  $\Delta V_{th}$  broadly increases for higher  $|V_{gsense}|$ . Although the  $\Delta V_{th}$  for some devices can reduce for higher  $|V_{gsense}|$  as shown in Fig. 5b, it is rare for a trap to be above a localized percolation path. As a result, the average (symbols in Fig. 7b) increases monotonically for higher  $|V_{gsense}|$ , which is partly driven by the mobility degradation.

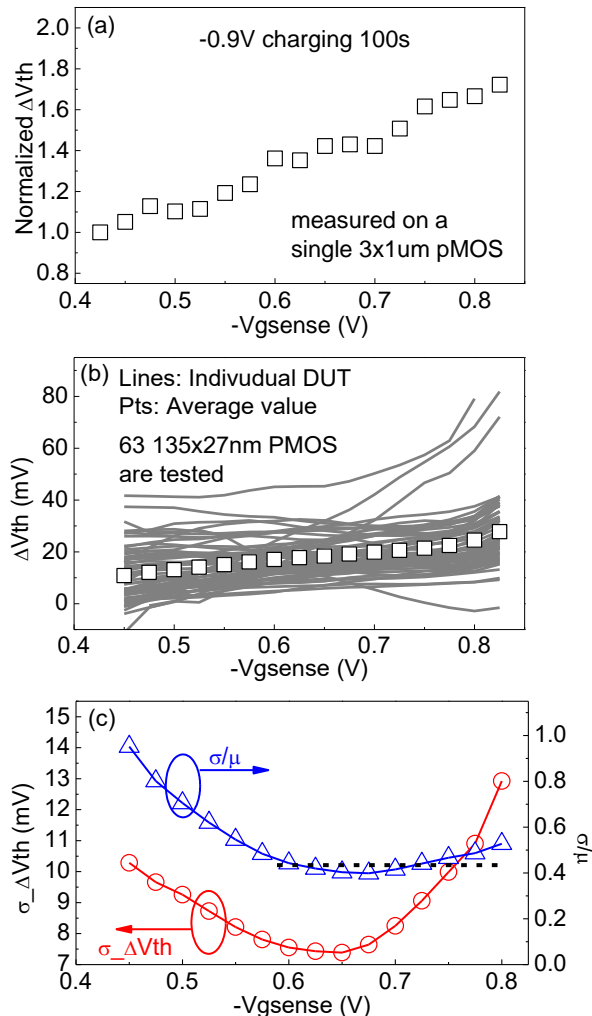


Fig. 7. (a) The impact of sensing  $V_g$  on the apparent  $\Delta V_{th}$  of a large device of  $3 \times 1 \mu\text{m}$ . The deep level traps were filled under  $V_g = -0.9 \text{ V}$  for 100 sec before

the measurement.  $\Delta V_{th}$  is normalized against its value at  $V_g = V_{th}$ . (b) The stochastic variation of  $135 \times 27 \text{ nm}$  devices. Each line is from one device. The symbols are the average. (c) Dependence of  $\sigma$  and  $\sigma/\mu$  on the sensing  $V_g$ .

The standard deviation,  $\sigma$ , is plotted against  $V_{gsense}$  in Fig. 7c. It can be divided into two regions: as  $|V_{gsense}|$  increases,  $\sigma$  decreases first and then increases. The minimum point is around 0.65 V. To explore this further, the relative variation,  $\sigma/\mu$ , is also plotted in Fig. 7c. When  $|V_{gsense}| > 0.65 \text{ V}$ ,  $\sigma/\mu$  only rises modestly, so that the higher  $\sigma$  is mainly caused by the higher  $\mu$ , as shown by the symbols in Fig. 7b. Below 0.65 V, however,  $\sigma$  increases and  $\mu$  decreases for lower  $|V_{gsense}|$ , resulting in a rising  $\sigma/\mu$ . When  $|V_{gsense}|$  lowers towards  $|V_{th}|$ , the current path becomes increasingly localized, leading to higher statistical variations, even though the trapped charges remain the same.

The cumulative distribution probability of  $\Delta V_{th}$  is given in Fig. 8a and  $\sigma$  is plotted against  $\mu$  in Fig. 8b for  $V_{gsense} = V_{th}$ . The RTN of nMOSFETs is smaller than that of pMOSFETs.  $\sigma$  follows  $\mu$  by a power law with an exponent of  $\sim 0.5$ , agreeing with the prediction of Defect-Centric model [2, 16, 29]. According to this model, the average  $\Delta V_{th}$  induced by a trap,  $\eta$ , is,

$$\eta = \frac{\sigma^2}{2\mu}$$

Using the fitted line in Fig. 8b,  $\eta \sim 3.2 \text{ mV}$  is obtained for pMOSFETs. This  $\eta$  is  $\sim 2 \times q/C_{ox}$  approximately, where  $q$  is one electron charge and  $C_{ox}$  the gate oxide capacitance. This agrees well with the value reported for the recoverable component of NBTI of pFinFETs [16], although the test samples used here are planar pMOSFETs from a different supplier. The average number of traps,  $N$ , per device is,

$$N = \frac{\mu}{\eta}$$

For pMOSFETs, a  $\mu \sim 12 \text{ mV}$  in Fig. 4b gives  $N \sim 4$ .

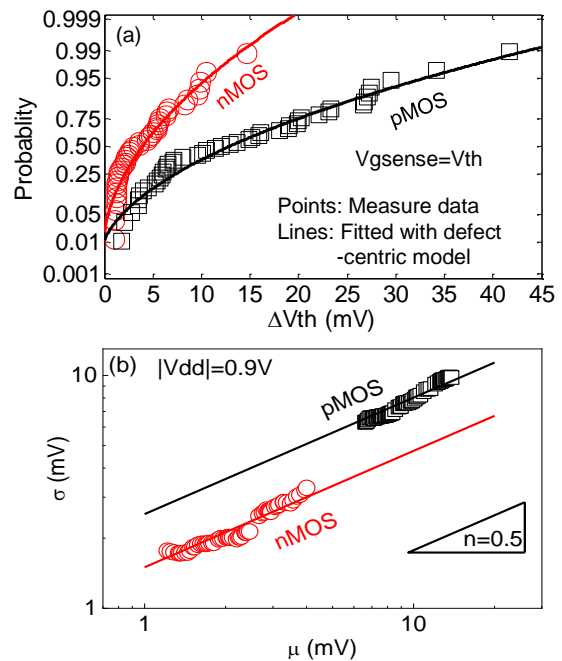


Fig. 8. (a) The cumulative distribution of  $\Delta V_{th}$ . The symbols are test data and the lines are fitted with the Defect-Centric model that assumes the number of traps per device following the Poisson's distribution and the  $\Delta V_{th}$  induced by a trap following

exponential distribution. (b) Standard deviation versus mean. Lines show that the data follow the prediction of Defect-Centric model well with a power exponent of 0.5. The different pairs of  $(\mu, \sigma)$  are obtained by varying the time window of “Id monitor” from 10  $\mu$ s to 100 sec in Fig. 2a.

For nMOSFETs, the corresponding values are  $\mu \sim 6.5$  mV,  $\eta \sim 1.1$  mV, and  $N \sim 6$ . When compared with pMOSFETs, the lower RTN in nMOSFETs is caused by smaller  $\eta$ . Although there are more traps in nMOSFETs, they are in the high-k layer and further away from the conduction channel and induce a smaller  $\Delta V_{th}$  [30].

#### IV. CONCLUSIONS

The conventional method of ‘Measure-Stress-Measure’ at preset time is inapplicable for the RTN-induced  $\Delta V_{th}$ , since the trap can be neutral when pulse IVs are taken. Early works estimate the RTN-induced  $\Delta V_{th}$  by  $\Delta I_d/g_m$  at  $V_g=V_{dd}$  and its accuracy is not known. In this paper, we propose a new TWC method for directly measuring the real  $\Delta V_{th}$  at  $V_g=V_{th}$ . By setting the trigger level close to the upper envelope of trapping-induced  $\Delta I_d$ , it ensures that the pulse IV is taken when traps are charged.

Results show that there is no unique relationship between  $\Delta I_d/g_m$  at  $V_g=V_{dd}$  and the directly measured  $\Delta V_{th}$  and their correlation is poor. The device-specific dependence of the apparent  $\Delta V_{th}$  on the sensing  $V_g$  originates from the DDV of relative local current density under a trap at  $V_{th}$ . Moreover, on average,  $\Delta I_d/g_m(V_{dd})$  doubles  $\Delta V_{th}(V_{th})$  and the charge-induced mobility degradation through Columbic scattering plays a role.

The TWC is applicable to devices with or without analyzable RTN signals. For the first time, it is used for assessing the statistical properties of the directly measured RTN-induced  $\Delta V_{th}$ . For the same trapped charges, it is found that  $\sigma$  has a minimum around  $|V_{gsense}|=0.65$  V. The increase in  $\sigma$  when  $|V_g|$  lowers toward  $|V_{th}|$  is explained by an increased localization of current path. The DDV follows the Defect-Centric model. For the  $135 \times 27$  nm devices used in this work, the average  $\Delta V_{th}$  induced per trap is  $\sim 3.2$  mV for pMOSFETs and  $\sim 1.1$  mV for nMOSFETs.

#### REFERENCES

- [1]. C. Prasad, M. Agostinelli, J. Hicks, S. Ramey, C. Auth, K. Mistry, S. Natarajan, P. Packan, I. Post, S. Bodapati, M. Giles, S. Gupta, S. Mudanai, K. Kuhn, “Bias Temperature Instability Variation on SiON/Poly, HK/MG and Trigate Architectures,” in Proc. Int. Rel. Phys. Symp., 2014, pp. 6A.5.1-6A.5.7, doi: 10.1109/IRPS.2014.6861101.
- [2]. B. Kaczer, T. Grasser, P. 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 Proc. Int. Rel. Phys. Symp., May 2010, pp. 26–32, doi: 10.1109/IRPS.2010.5488856.
- [3]. M. K. Li, “Jitter Challenges and Reduction Techniques at 10 Gb/s and Beyond,” IEEE Trans. Advanced Packaging, vol. 32, no. 2, pp. 290-297, 2009, doi: 10.1109/TADVP.2009.2012432.
- [4]. J. Zou, R. Wang, N. Gong, R. Huang, X. Xu, J. Ou, C. Liu, J. Wang, J. Liu, J. Wu, S. Yu, P. Ren, H. Wu, S. W. Lee, and Y. Wang, “New insights into AC RTN in scaled high- $\kappa$ /metal-gate MOSFETs under digital circuit operations,” in Proc. Symp. Very Large Scale Integr. (VLSI) Technol., Jun. 2012, pp. 139–140, doi: 10.1109/VLSIT.2012.6242500.
- [5]. H. Miki, M. Yamaoka, N. Tega, Z. Ren, M. Kobayashi, C. P. D’Emic, Y. Zhu, D. J. Frank, M. A. Guillorn, D.-G. Park, W. Haensch, and K. Torii, “Understanding Short-term BTI Behavior through Comprehensive Observation of Gate-voltage Dependence of RTN in Highly Scaled High- $\kappa$  / Metal-gate pFETs,” in Proc. Symp. Very Large Scale Integr. (VLSI) Technol., Jun. 2011, pp. 148–149.
- [6]. H. Miki, M. Yamaoka, D. J. Frank, K. Cheng, D.-G. Park, E. Leobandung, and K. Torii, “Voltage and Temperature Dependence of Random Telegraph Noise in Highly Scaled HKMG ETSOI nFETs and its Impact on Logic Delay Uncertainty,” in Proc. Symp. Very Large Scale Integr. (VLSI) Technol., Jun. 2012, pp. 137–138, doi: 10.1109/VLSIT.2012.6242499.
- [7]. 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. Very Large Scale Integr. (VLSI) Technol., Jun. 2009, pp. 54–55.
- [8]. 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 Proc. Symp. Very Large Scale Integr. (VLSI) Technol., Jun. 2014, pp. 200–201, doi: 10.1109/VLSIT.2014.6894417.
- [9]. C. Liu, K. T. Lee, H. Lee, Y. Kim, S. Pae, and J. Park, “New Observations on the Random Telegraph Noise induced  $V_{th}$  Variation in Nano-scale MOSFETs,” in Proc. Int. Rel. Phys. Symp., 2014, pp. XT.17.1-XT.17.5, doi: 10.1109/IRPS.2014.6861194.
- [10]. 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., pp. 628-631, 2010, doi: 10.1109/IEDM.2010.5703437.
- [11]. C. Chen, Q. Ran, H. Cho, A. Kerber, Y. Liu, M. Lin, and R. W. Dutton, “Correlation of  $I_d$ - and  $I_g$ -Random Telegraph Noise to Positive Bias Temperature Instability in Scaled High- $\kappa$ /Metal Gate n-type MOSFETs,” in Proc. Int. Rel. Phys. Symp., May 2011, pp. 190–195, doi: 10.1109/IRPS.2011.5784475.
- [12]. M. Duan, J. F. Zhang, Z. Ji, W. Zhang, B. Kaczer, S. De Gendt, and G. Groeseneken, “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.
- [13]. A. Teramoto, T. Fujisawa, K. Abe, S. Sugawa, and T. Ohmi, “Statistical Evaluation for Trap Energy Level of RTS Characteristics,” in Proc. Symp. Very Large Scale Integr. (VLSI) Technol., Jun. 2010, pp. 99–100, doi: 10.1109/VLSIT.2010.5556186.
- [14]. J. W. Lu, C. Vaz, J. P. Campbell, J. T. Ryan, K. P. Cheung, G. F. Jiao, G. Bersuker, and C. D. Young, “Device-Level PBTI-induced Timing Jitter Increase in Circuit-Speed Random Logic Operation,” in Proc. Symp. Very Large Scale Integr. (VLSI) Technol., Jun. 2014, pp. 102–103, doi: 10.1109/VLSIT.2014.6894387.
- [15]. M. Duan, J. F. Zhang, Z. Ji, W. Zhang, B. Kaczer, T. Schram, R. Ritzenthaler, G. Groeseneken, and A. Asenov, “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, Sept. 2014, doi: 10.1109/TED.2014.2335053.
- [16]. J. Franco, B. Kaczer, M. Toledano-Luque, Ph. J. Roussel, J. Mitard, L.-A. 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 Proc. Int. Rel. Phys. Symp., 2012, pp. 5A.4.1-5A.4.6, doi: 10.1109/IRPS.2012.6241841.
- [17]. B. Kaczer, J. Franco, M. Toledano-Luque, Ph. J. Roussel, M. F. Bukhori, A. Asenov, B. Schwarz, M. Bina, T. Grasser, and G. Groeseneken, “The Relevance of Deeply-Scaled FET Threshold Voltage Shifts for Operation Lifetimes,” in Proc. Int. Rel. Phys. Symp., 2012, pp. 5A.2.1-5A.2.6, doi: 10.1109/IRPS.2012.6241839.
- [18]. A. Asenov, R. Balasubramaniam, A. R. Brown, and J. H. Davies, “RTS Amplitudes in Decanometer MOSFETs: 3-D Simulation Study,” IEEE Trans. Electron Devices, vol. 50, no. 3, pp. 839–845, Mar. 2003, doi: 10.1109/TED.2003.811418.
- [19]. M. F. Bukhori, S. Roy, and A. Asenov, “Simulation of Statistical Aspects of Charge Trapping and Related Degradation in Bulk MOSFETs in the Presence of Random Discrete Dopants,” IEEE Trans. Electron Devices, Vol. 57, No. 4, pp. 795-803, 2010, doi: 10.1109/TED.2010.2041859.
- [20]. 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, 2016, doi: 10.1109/TED.2016.2528218.
- [21]. J. F. Zhang, M. H. Chang, Z. Ji, L. Lin, I. Ferain, G. Groeseneken, L. Pantisano, S. De Gendt, M. M. Heyns, "Dominant layer for stress-induced positive charges in Hf-based gate stacks," IEEE Electron Dev. Lett., Vol.29, No.12, pp.1360-1363, 2008, doi: 10.1109/LED.2008.2006288.
- [22]. R. Gao, Z. Ji, S. M. Hatta, J. F. Zhang, J. Franco, B. Kaczer, W. Zhang, M. Duan, S. De Gendt, D. Linten, G. Groeseneken, J. Bi and M. Liu, "Predictive As-grown-Generation (A-G) model for BTI-induced device/circuit level variations in nanoscale technology nodes," in IEDM Tech. Dig., pp. 778-781, 2016, doi: 10.1109/IEDM.2016.7838520.
- [23]. M. Duan, J. F. Zhang, Z. Ji, W. Zhang, D. Vigar, A. Asenov, L. Gerrer, V. Chandra, R. Aitken, and B. Kaczer, "Insight into Electron Traps and Their Energy Distribution under Positive Bias Temperature Stress and Hot Carrier Aging," IEEE Trans. Electron Devices, Vol. 63, No. 9, pp. 3642-3648, 2016, doi: 10.1109/TED.2016.2590946.
- [24]. 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.2017.2691008.
- [25]. J. F. Zhang, Z. Ji, and W. Zhang, "As-grown-generation (AG) model of NBTI: A shift from fitting test data to prediction," Microelectronics Reliability, Vol. 80, pp. 109-123, 2018, doi: 10.1016/j.microrel.2017.11.026.
- [26]. 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., pp.774-777, 2013, doi: 10.1109/IEDM.2013.6724730.
- [27]. Z. Çelik-Butler, S. P. Devireddy, H.-H. Tseng, P. Tobin, A. Zlotnicka, "A low-frequency noise model for advanced gate-stack MOSFETs," Microelectronics Rel. Vol. 49, pp.103-112, 2009, doi: 10.1016/j.microrel.2008.12.005.
- [28]. W. Zhu, J. P. Han, and T. P. Ma, "Mobility measurement and degradation mechanisms of MOSFETs made with ultrathin high-k dielectrics," IEEE Trans. Electron Dev., vol. 51, No.1, pp. 98-104, 2004, doi: 10.1109/TED.2003.821384.
- [29]. L. M. Procel, F. Crupi, J. Franco, L. Trojman, and B. Kaczer, "Defect-Centric Distribution of Channel Hot Carrier Degradation in Nano-MOSFETs," IEEE Electron. Dev. Lett., vol. 35, no. 12, pp. 1167-1169, Dec, 2014, doi: 10.1109/LED.2014.2361342.
- [30]. M. Toledano-Luque, B. Kaczer, J. Franco, Ph.J. Roussel, T. Grasser, T.Y. Hoffmann, and G. Groeseneken, "From mean values to distributions of BTI lifetime of deeply scaled FETs through atomistic understanding of the degradation," in Proc. Symp. Very Large Scale Integr. (VLSI) Technol., Jun. 2011, pp. 152-153.