Microelectronics Reliability 49 (2009) 642–649

Contents lists available at ScienceDirect

Microelectronics Reliability

journal homepage: www.elsevier.com/locate/microrel

Impacts of NBTI and PBTI on SRAM static/dynamic noise margins and cell failure probability

Aditya Bansal [a], Rahul Rao [a], Jae-Joon Kim [a], Sufi Zafar [a], James H. Stathis [a,*], Ching-Te Chuang [a,b]

[a] IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, United States
[b] National Chiao Tung University, Hsinchu, Taiwan

ARTICLE INFO

Article history:
Received 9 August 2008
Received in revised form 12 March 2009
Available online 22 April 2009

ABSTRACT

Negative and Positive Bias Temperature Instabilities (NBTI (in PFET) and PBTI (in NFET)) weaken MOSFETs with time. The impact of such device degradation can be severe in Static Random Access Memories (SRAMs) wherein stability is governed by relative strengths of FETs. Degradation in stability with time under 'worst case condition' gets more important because of reduced guard-banding due to process induced instability. In this work, circuit insights into worst-case conditions and effect of NBTI and PBTI, individually and in combination, on the stability of an SRAM cell are presented. It is shown that measurable quantities such as static noise-margin are not sufficient to completely understand the combined effect of NBTI and PBTI. Monte-Carlo simulations are performed in a 45 nm PDSOI technology to estimate the increase in cell failure probability with time. In worst case, NBTI and PBTI both degrade read stability (significantly) and writability (marginally). Further, we analyze the choice of optimal power supply considering the trade-off between short-term stability (due to process variations) and long-term stability (due to NBTI/PBTI) to achieve six-sigma confidence in SRAM cell robustness.

1. Introduction

Negative Bias Temperature Instability (NBTI) is a major concern since it deteriorates the drive-strength of p-channel FETs (PFETs), resulting in degraded circuit performance and robustness with time [1]. With the introduction of high-permittivity gate dielectrics to improve short-channel effects (SCE) and reduce gate tunneling leakage current, n-channel FETs (NFETs) are also becoming prone to time-dependent performance degradation due to charge trapping [2]. This phenomenon is called Positive Bias Temperature Instability (PBTI). The impact of NBTI and PBTI can be significant in circuits such as Static Random Access Memories (SRAMs) which are more susceptible to functional failure and demand good tolerance throughout the life of an integrated circuit (IC). Fig. 1 shows a conventional six transistor SRAM cell consisting of a cross-coupled inverter pair and two access transistors (AXR and AXL) that couple the inverter pair to the bit-lines (BL and BR). Very small sized FETs are used in SRAM cells to increase the integration density, making them more prone to process induced variability compared to logic circuits [4,5]. The FETs in an SRAM cell are precariously designed to achieve target stability and yield. However, time-dependent changes in FET characteristics (NBTI/PBTI) can potentially change the relative strengths of FETs (and make them asymmetric as shown in Fig. 1), resulting in varied stability (compared to as designed) during operation.

The impact of time-dependent degradation on SRAM cell stability has been investigated recently. Lin et al. [6,7] considered the simultaneous degradation in both the PFETs (PL/PR) and NFETs (NL/NR) and inferred that READ stability degrades with NBTI and PBTI while WRITE margin improves. They also looked at degradation in static noise-margin (SNM) due to asymmetric degradation in NFETs. Kang et al. [8], considered NBTI only, assuming asymmetric FET degradation (as in Fig. 1) during READ operation while symmetric FET degradation during WRITE operation. Some researchers have also talked about relaxation or recovery in FETs when stress voltage or temperature is removed. Grasser et al. [12] showed that permanent/slowly relaxing component increases with increase in stress time. Since an SRAM cell may store the same data for long period of time (read multiple times but data not flipped), asymmetric FET degradation can occur.

In this work, we analyze the worst-case condition for an SRAM cell due to NBTI and PBTI and deduce its noise immunity and failure probability with time. In particular,

- We provide the circuit insights into worst-case conditions affecting READ and WRITE operations.
- In worst case, NBTI and PBTI degrade the stability during READ (significantly) and WRITE (marginally) operations. Degradation is more sensitive to PBTI than NBTI.

* Corresponding author. Tel.: +1 914 945 2559; fax: +1 914 945 2141.
E-mail addresses: bansal@us.ibm.com (A. Bansal), stathis@us.ibm.com (J.H. Stathis).

0026-2714/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.microrel.2009.03.016

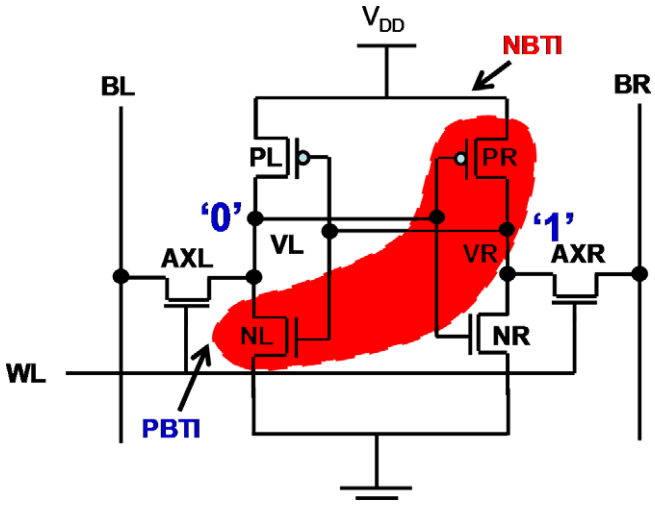

Fig. 1. Schematic of an SRAM cell showing degradation in NL and PR due to long-term storing of '0' at left node and '1' at right node resulting in an asymmetric cell.

- We compute the increase in number of faulty SRAM cells with time using intelligent Monte-Carlo simulations [10].
- We analyze the trade-off between short-term stability (process variations) and long-term stability (NBTI/PBTI) to achieve desired six-sigma confidence.

Circuit simulations are performed using SPICE in a 45 nm SOI technology.

2. Analysis framework

In this section, we present our analysis methodology to compute time, V_{DD} and temperature dependence of an SRAM design metric. We further discuss the different stress conditions assumed in this work.

2.1. Flow to compute time and stress dependent circuit metric

The analysis approach adopted in this work is shown in Fig. 2. The complete flow can be broadly classified into three steps:

Step 1 ($\Delta V_T = f(t, V_{DD}, T)$)

First, the dependence of change in threshold-voltage (ΔV_T) on time (t), voltage (V_{DD}) and temperature (T) is modeled and calibrated using measurement results for each FET type. Several models have been presented that predict the shift in device threshold voltage (ΔV_T) due to NBTI and PBTI. In this work, we use a simplistic model that has been derived from first principles [9]

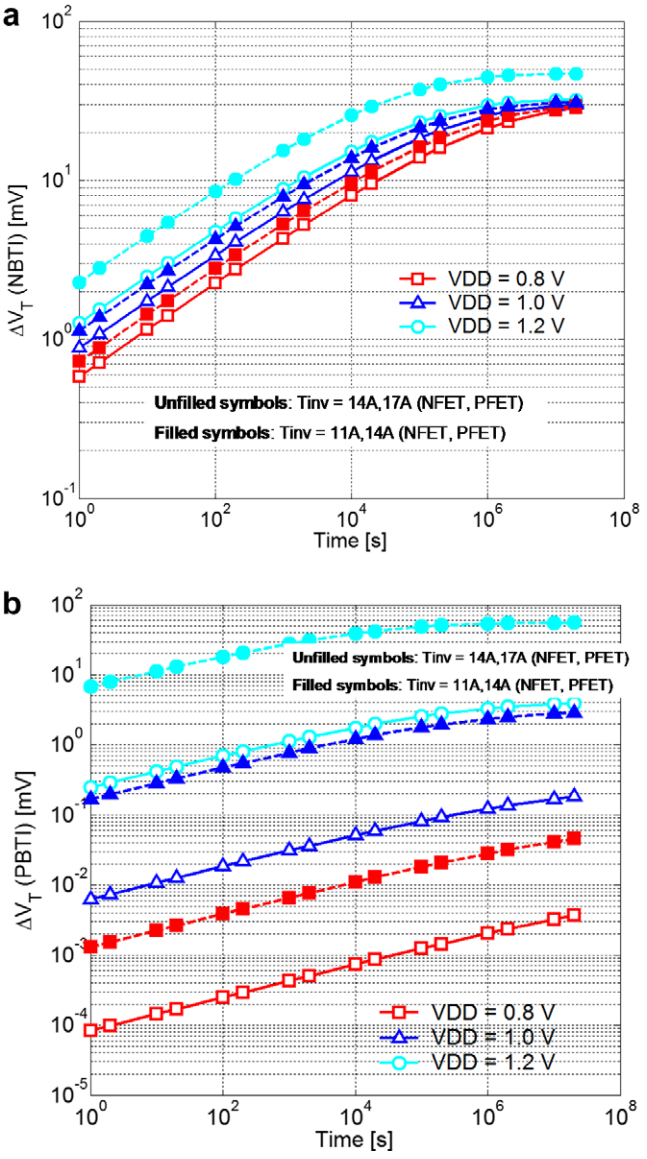

Fig. 3. Time-dependence of threshold voltage in TiN and Re gated devices with SiO_2/HfO_2 as dielectric stack due to (a) NBTI (on PFET), and (b) PBTI (on NFET). V_T shift is obtained using Eq. (1). Two cases of inversion thicknesses (T_{inv}) are shown to analyze the impact of scaling. PBTI has found to be more sensitive to T_{inv} than NBTI.

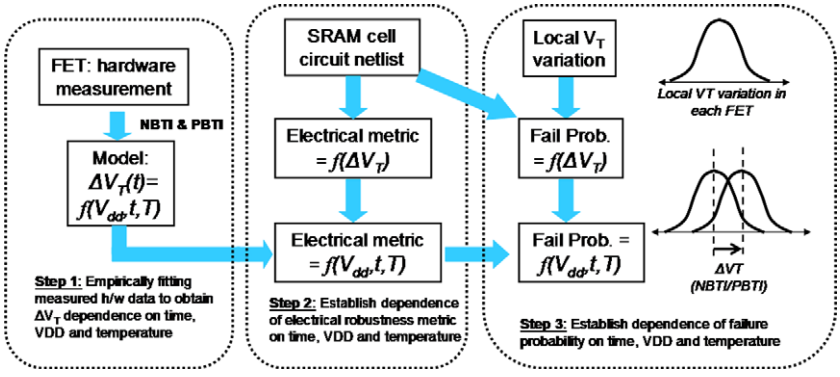

Fig. 2. Analysis approach to estimate time-dependent circuit parameters (noise-margins) and failure probability (READ/WRITE).

$$\Delta V_T(t) = \Delta V_{MAX}(1 - e^{(-t/\tau)\beta}) \quad (1)$$

Here, ΔV_{MAX} represents the maximum shift that could occur under prolonged stress conditions. τ represents the time required for ΔV_T to reach 63% of ΔV_{MAX} and is thus a measure of degradation rate. β is a measure of hydrogen dispersion. ΔV_{MAX} and τ are dependent on the stress electric field across the oxide, and hence on the supply voltage, while β does not depend on stress electric field across the oxide. The derivation and explanation of the various model coefficients can be found in [9].

Using the above model, Fig. 3a and b show the time dependent V_T increase due to NBTI and PBTI, respectively, in TiN and Re gated devices with SiO_2/HfO_2 as dielectric stack [3]. Typical inversion thicknesses (T_{inv}) used in [3] are 1.4 nm in NFETs and 1.7 nm in PFETs. It can be seen that at these T_{inv} values, V_T degradation due to NBTI is significantly larger than PBTI at typical operating voltages (~0.8–1.2 V). With technology scaling, T_{inv} needs to be scaled without (or little) scaling V_{DD} to increase performance at every technology generation. Hence, stress electric field, $E_{stress} = (V_{DD} - V_T)/T_{inv}$ will increase thereby increasing V_T degradation [1–3]. Using the above model, we anticipate the V_T degradation to increase if T_{inv} is reduced by 0.3 nm (shown as filled symbols in Fig. 3a and b). The model exhibits more sensitivity of PBTI on E_{stress} compared to NBTI. Please note that several technology features such as gate dielectric stack, and gate material are embedded in the fitting parameters of Eq. (1). In this paper, our aim is to set up an analysis methodology independent of technology. All the SRAM simulations in this work are performed with T_{inv} = 1.4 nm (NFET), 1.7 nm (PFET).

Step 2 ($\Delta SNM = f(\Delta V_T) = f(t, V_{DD}, T)$)

Next we establish the dependence of electrical reliability metrics, such as READ static noise margin (SNM), writability and cell failure probability (P_F) on the change in V_T due to NBTI and PBTI. To calculate a circuit's electrical robustness metric (say SNM), its dependence on V_T shifts due to NBTI and PBTI is obtained for a large range of ΔV_T values. Then, the dependence of ΔV_T on time, V_{DD} and T is inserted using step 1.

Step 3 ($\Delta P_F = f(\Delta V_T) = f(t, V_{DD}, T)$)

An SRAM cell can fail at time $t = 0$ due to process variations [4]. Process induced variability can affect an SRAM cell locally as well as globally [5]. Local variations such as random dopant fluctuations (RDF) and line-edge roughness (LER) affect each FET in a cell differently resulting in mismatch in drive strengths. In this work, we only consider the impact of local V_T variations on SRAM failure as they are dominant culprit behind SRAM cell failure [4]. Assuming random V_T variations to be Gaussian, a probability distribution function for ΔV_T is obtained. For each set of randomly assigned ΔV_Ts to each FET, an SRAM cell is tested to determine whether it passes the design metric. Assuming n random samples of ΔV_T (for each FET), cell failure probability (P_F) at time t can be given as

$$P_F(t) = 1 - \frac{1}{n}\sum_{j=1}^{n} I_j, \quad I = \text{'1' (if pass) } else \text{ '0'} \quad (2)$$

Typically n has to be large; hence, we perform Monte-Carlo simulations using mixture importance sampling technique [10] to reduce the computational complexity. Local V_T mismatch due to process variations is superimposed on the time-dependent V_T

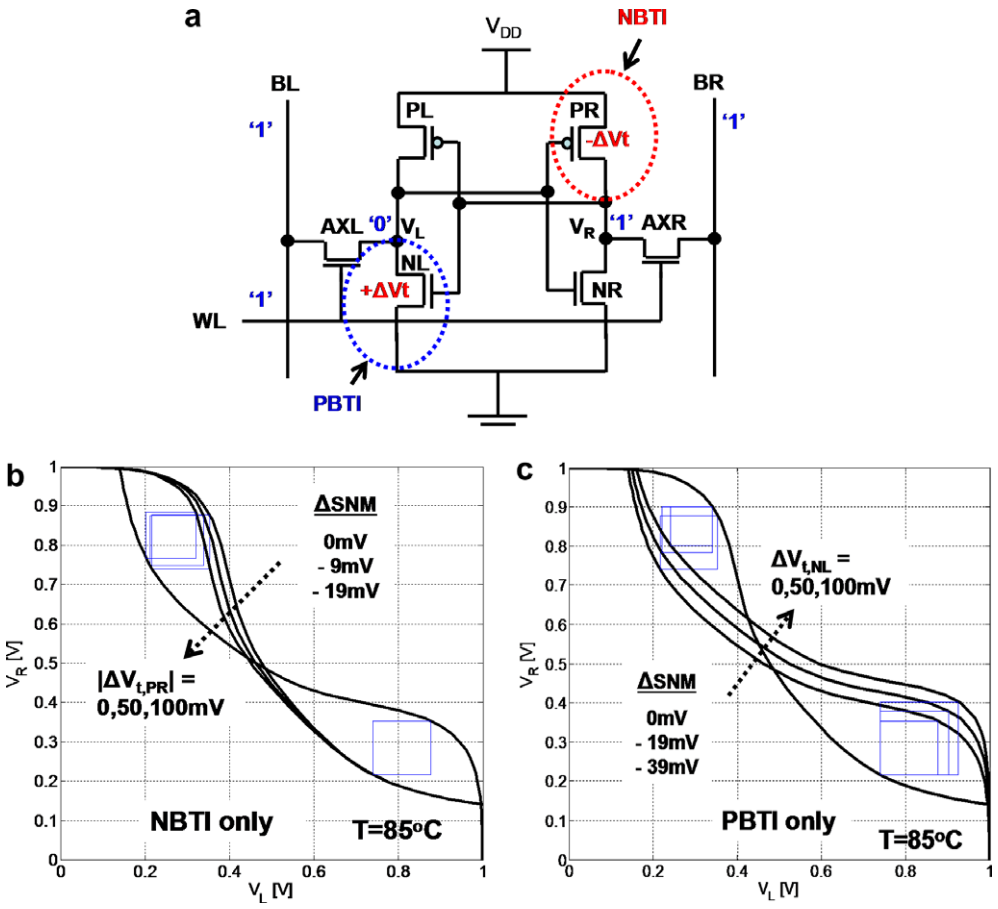

Fig. 4. (a) Worst-case scenario for READ stability; (b) & (c) Static stress condition: calculation of Static Noise Margin by fitting the largest square in the voltage transfer characteristics of two inverters due to NBTI only (b), and PBTI only (c).

increase due to NBTI and PBTI (for the worst case). Finally, the dependence of failure probability on V_T is translated to the dependence on time, V_{DD} and T using calibrated NBTI/PBTI models in step 1. For comparison, we translate the failure probability in terms of number of faulty cells (N_F) in a 100 MB memory. N_F can be expressed as

$$N_F(t) = 100 \times 1024 \times 1024 \times P_F(t) \quad (3)$$

The assumptions made in this work are:

- Only the FETs in the cross-coupled inverters (NL, NR, PL, PR) experience NBTI/PBTI. The access NFETs (AXL, AXR) do not experience PBTI since the transistors are *on* only when the cell is accessed and the total access time for the cell is much smaller than stand-by time.
- The recovery of devices [11,12] during non-stress periods is not considered.

2.2. Stress conditions

We look at two different stress conditions.

A. Static stress: A cell storing the same data for long period of time becomes asymmetric due to degradation of two of the four devices in the cross-coupled pair. For example, in Fig. 1a, if a cell is storing '0' ('1') at the left (right) storage node, the left pull-down NFET (NL) and right pull-up PFET (PR) will be affected. Note that reading a cell or writing the same data multiple times does not flip the data and results in static stress.

B. Alternating stress: Data is regularly flipped resulting in deterioration of all four devices. Note that the cell remains symmetric.

Note that in some SRAM bit-cells, data may be flipping at regular intervals resulting in an asymmetric degradation condition somewhere between static stress and alternating stress. For worst case, extreme asymmetric condition has been considered in static stress.

3. SRAM: how unstable can it get?

In this section, we discuss the READ and WRITE modes of operation of an SRAM cell and the worst-case conditions that result in the cell becoming less robust and ultimately failing to operate as desired.

3.1. READ operation

With reference to Fig. 4a, let us consider a scenario where the left node (V_L) is storing '0' and right node (V_R) is storing '1'. During a READ operation, first bit-lines (BL and BR) are pre-charged to V_{DD}. Then the word-line (WL) is raised to '1' resulting in AXL and AXR becoming *on*. The bit-line BL starts discharging through AXL and NL. The bit-line voltage on BL starts to drop, which is detected by the sense-amplifier. During this discharge process, a small voltage is generated at V_L that is dependent on the relative strengths of AXL and NL. If this node voltage increases beyond the trip point of the inverter pair (PR–NR), the cell flips, resulting in an erroneous operation. The gate of NL should remain close to '1' during the whole process, keeping it *on* which helps to maintain V_L close to '0'. If NL becomes weak due to PBTI, it will not be able to maintain '0' at V_L and AXL will start increasing the voltage at V_L. As the voltage at V_L increases, NR starts turning *on* while PR starts turning *off*. Note that PR is holding the value '1' at the right node (V_R) and as it turns *off*, cell will flip data. Hence, weakening of PR due to NBTI will also reduce READ stability. If cell flips during READ operation, the stored data gets corrupted and READ fail occurs.

A worst case condition for READ operation can occur due to static stress, i.e., if a cell is storing the same value for long time. Say if left node is storing '0' and right node '1' for long time, it will result in PBTI shift in NL and NBTI shift in PR. This is illustrated in Fig. 4a and such a scenario is quite possible in relatively dormant sections of a memory. Thus, a worst case condition for READ operation after a long time (t_1) can be given as

$$\begin{aligned}&\text{Worst condition for READ :}\\ &t = 0 \text{ to } t_1 : \quad V_L = \text{'0' and } V_R = \text{'1'}\\ &\text{(NL and PR weaken due to static stress)}\\ &t = t_1 + T_{CLK} : \text{ READ operation is performed}\end{aligned} \quad (4)$$

Note that if left node is storing '1' (and right node '0') the degradation in corresponding FETs (PL and NR) will also result in worst case.

3.1.1. Static Noise Margin (SNM)

Stability of a cell during READ can be measured as static noise margin (SNM) of the cross-coupled inverters. SNM is the side of the largest fitted square in the wings of butterfly curve obtained by plotting voltage transfer characteristics of cross-coupled inverters. Some key observations from SNM curves in Fig. 4a and c:

- Assuming static stress degrades only NL while NR remains unaffected, transfer characteristics of left inverter (NL–PL) will shift resulting in reduced SNM (Fig. 4c).
- Alternating stress will result in PBTI-induced V_T shift in both the NFETs – NL and NR. If V_T shifts equally in NL and NR, both the lobes of SNM butterfly curve (Fig. 4c) will shift, resulting in negligible impact due to PBTI.
- Whereas, NBTI only impacts one lobe of the SNM butterfly curve (Fig. 4b). Note that degradation in PL will affect one lobe, whereas PR will affect second lobe. Hence, static as well as alternating stress will impact the noise margin in similar fashion.

Fig. 4b and c show the individual impacts of NBTI and PBTI on SNM, respectively, due to static stress. It can be seen that SNM reduces due to both NBTI and PBTI and their combined effect is additive as can be seen from Fig. 5a. It is observed that SNM, under static stress, varies linearly with V_T shifts in FETs and can be approximated as

$$\Delta SNM(t) = \left.\frac{\partial SNM}{\partial V_T}\right|_{PBTI} \Delta V_{T,NL}(t) + \left.\frac{\partial SNM}{\partial V_T}\right|_{NBTI} \Delta V_{T,PR}(t) \quad (5)$$

Fig. 5a shows that SNM is more sensitive to V_T increase due to PBTI compared to NBTI. This difference in relative SNM sensitivity to V_T shift ($\partial SNM/\partial V_T$) due to PBTI/NBTI can be explained from the relative driving strengths of pull-down and pull-up FETs. Strength can be measured as drain-to-source current (I_{DS}) during saturation ($V_{GS} = V_{DS} = V_{DD}$). A first order expression for I_{DS} can be given as [13]

$$I_{DS} = \mu\left(\frac{W}{L}\right) C_{OX}(V_{GS} - V_T)^{\alpha} \quad (6)$$

where μ is the effective carrier mobility, W is the FET width, L is the FET gate length, C_{OX} is the effective gate capacitance and α is a technology-dependent constant, typically ranging between 1 and 2. Typically, pull-down and pull-up FETs have same C_{OX}, L and same magnitude of applied voltage biases. Hence, relative sensitivities of their drive current's strength to V_T can be simplified as

$$\frac{(\delta I_{DS}/\delta V_T)|_{\text{pull-down}}}{(\delta I_{DS}/\delta V_T)|_{\text{pull-up}}} = \left(\frac{\mu_{\text{electron}}}{\mu_{\text{hole}}}\right)\left(\frac{W_{\text{pull-down}}}{W_{\text{pull-up}}}\right) \quad (7)$$

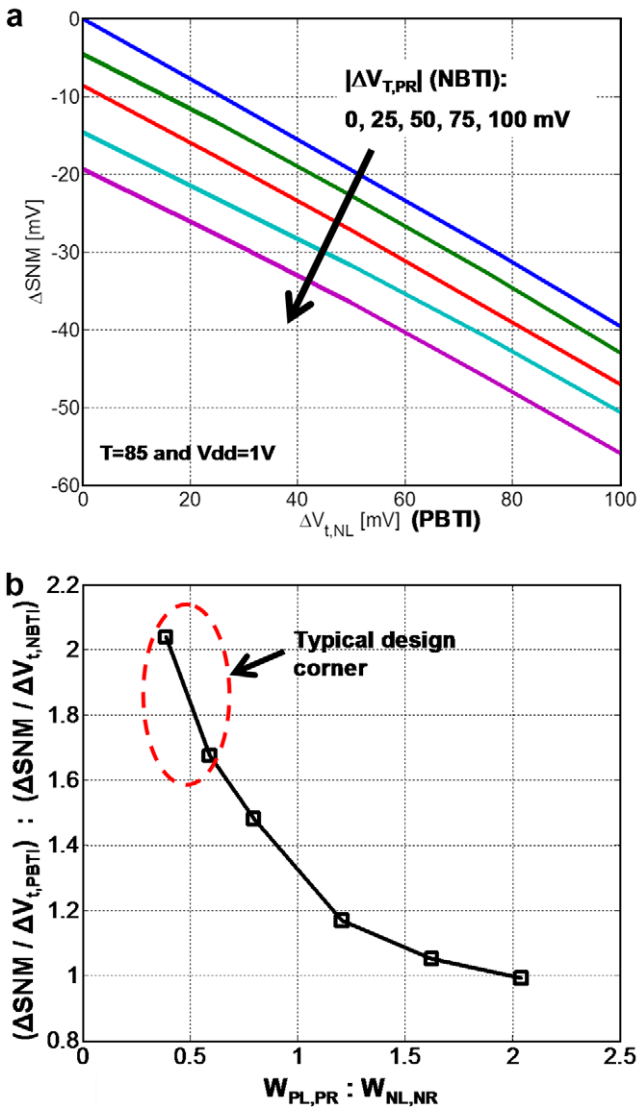

Fig. 5. (a) Combined effect of NBTI and PBTI on READ SNM under static stress, and (b) ratio of sensitivities of SNM to NBTI and PBTI against varying width ratios of PFET and NFET. Typically, NFET is wider than PFET, resulting in higher sensitivity of SNM to PBTI than NBTI.

Typically, electron mobility ($\mu_{electron}$) in pull-down NFET is twice of hole mobility (μ_{hole}) in pull-up PFET. Further, in SRAMs, $W_{pull\text{-}down}$ is larger than $W_{pull\text{-}up}$ for stability [4]. Hence, strength of pull-down NFET is more sensitive to change in V_T than pull-up PFET. Therefore, SRAM characteristics depend more on V_T degradation in NFETs due to PBTI. Fig. 5b shows the ratio of $\partial SNM/\partial V_T$ under static stress, due to NBTI and PBTI for varying widths (W) of PR (pull-up PFET) and NL (pull-down NFET). Typically, the width of NFET is larger than PFET for improving READ stability; hence the sensitivity of PBTI on SNM will be larger than NBTI. If we increase the width of PR to almost twice that of NL, the sensitivity of SNM to NBTI and PBTI also becomes comparable.

3.1.2. Number of faulty cells ($N_{F,READ}$) in 100 MB memory

Six-sigma confidence requirement translates to approx. 2 faulty cells among one billion cells i.e., $P_{F,READ} = 2 \times 10^{-9}$. This is equivalent to approximately 0.2 faulty cells in a memory of size 100 MB. For comparison, the FET sizes and V_{DD} are chosen so as to meet six-sigma requirement at $t = 0$. Fig. 6a shows the number of faulty cells due to READ failure ($N_{F,READ}$) in an array of size 100 MB due to NBTI and PBTI induced V_T shifts. We are showing the READ failure in two cases – (1) cell becomes asymmetric i.e., worst case under static stress as discussed above, and (2) symmetric degradation in both the NFETs (and/or PFETs) keeping the cell symmetric due to alternating stress. In the symmetric degradation case, PBTI has negligible effect on READ failure while the impact of NBTI remains unaffected. This is explained with the help of SNM butterfly curves in previous section.

$N_{F,READ}$ due to worst-case static stress can be similarly expressed as SNM (Eq. (5)),

$$\log_{10}\left(\frac{N_{F,READ}(t)}{N_{F,READ}(0)}\right) = \left.\frac{\partial(\log_{10}(N_{F,READ}(t)/N_{F,READ}(0)))}{\partial V_T}\right|_{PBTI} \Delta V_{T,NL}(t) + \left.\frac{\partial(\log_{10}(N_{F,READ}(t)/N_{F,READ}(0)))}{\partial V_T}\right|_{NBTI} \Delta V_{T,PR}(t) \quad (8)$$

where $N_{F,READ}(0)$ is the number of faulty cells at time $t = 0$. In the worst case, relative sensitivities of cell failure to NBTI and PBTI follow similar trend as READ SNM. Please see the explanation in Section 3.1.1 on SRAM characteristic sensitivities to NBTI and PBTI.

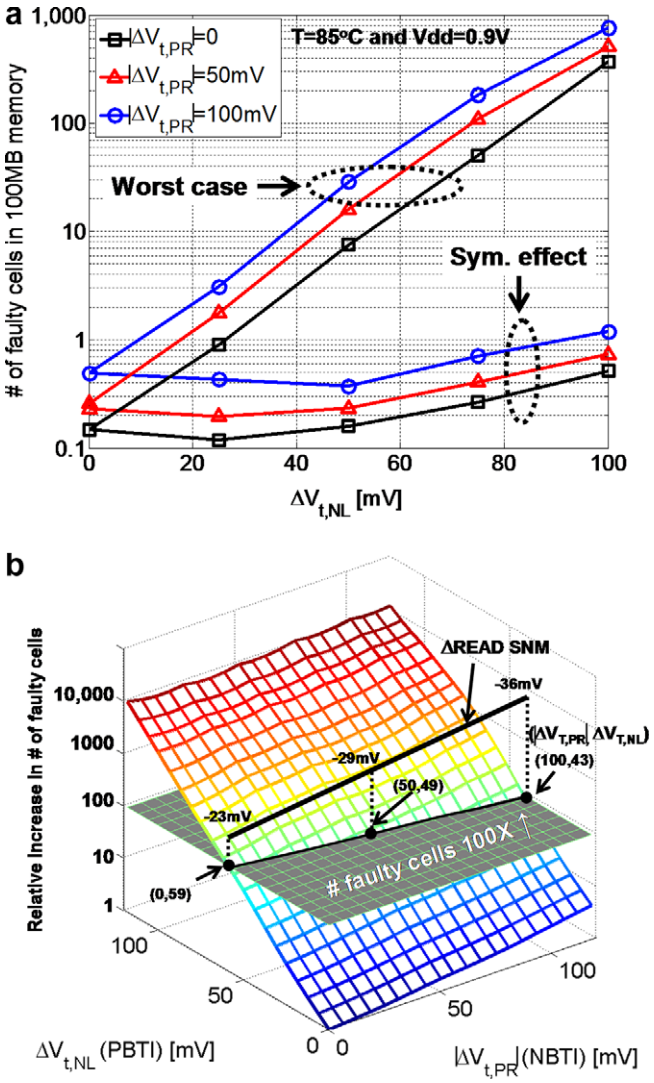

Fig. 6. (a) Number of faulty cells (unable to READ) in 100 MB memory; (b) Contour plot showing worst-case increase (under static stress) in no. of faulty cells. Combinations resulting in 100× increase in faulty cells are shown along the intersection of horizontal plane. Several combinations of NBTI and PBTI (hence, READ SNMs) result in certain increase in failure.

Hence, $N_{F,READ}(t)$ is more sensitive to PBTI than NBTI due to asymmetric degradation under static stress. However, in the case of symmetric degradation (all the four transistors are equally impacted) in both the PFETs and NFETs, $N_{F,READ}(t)$ still remains sensitive to NBTI while impact of PBTI becomes negligible.

Traditional practice is to consider READ SNM as a metric to determine the robustness of a cell against READ failures. Higher SNM implies a greater tolerance to local V_T mismatch, supply voltage variations and other dynamic noises. However, there is no one-to-one correlation between $N_{F,READ}$ (due to local V_T variation) and SNM (due to NBTI and PBTI induced V_T shifts). For example, Fig. 6b shows the dependence of cell failure on V_T shifts due to NBTI and PBTI in a 3-D contour plot. If the desired bound on increase in faulty cells is say 100× after certain time, there is a range of acceptable V_T shifts due to NBTI and PBTI. For each combination of ΔV_T (due to NBTI/PBTI), we get different READ SNM (which is represented as a shift from the nominal value along the line of intersection of the desired plane and the contour). This is because the coefficients of $\Delta V_{T,PR}$ (NBTI) and $\Delta V_{T,NL}$ (PBTI) affecting READ SNM (Eq. (5)) and $N_{F,READ}$ (Eq. (5)) are different. In particular:

- Relative impact of NBTI and PBTI depends upon the relative strengths of PFET and NFET.
- To restrict the number of time-dependent cell failures, technology designers can focus on reducing the bigger cause of instability during technology ramp-up cycles.
- On the other hand, for a fixed technology, to reduce the time-dependent degradation of circuit yield, circuit designers need to re-size the cell accordingly.

3.2. WRITE operation

Fig. 7a illustrates the WRITE operation. Let's assume that left node is storing '1' and right node is storing '0'. During WRITE, bit-line BR and word-line (WL) are raised to '1' and bit-line BL is pulled down to 0 V. Both the access FETs (AXL and AXR) are *on*. The left node starts discharging through AXL (in contention with PL) while right node starts charging through AXR (in contention with NR). For successful WRITE operation, cell nodes should flip in the time when word-line WL is high (at V_{DD}). As the cell node starts to flip, the discharging of the left node and the charging of the right node are aided by the turning *on* of NL and PR, respectively.

If a cell is in the state as shown in Fig. 7a for long time, NR and PL will weaken resulting in easier flipping of the cell. Hence, a bad case for READ is a good case for WRITE. The worst case to oppose the flipping, thus, is for NL and PR to become weak. Note that for NL (PR) to become weak, its gate should be at V_{DD} (0 V) for long time. Hence, a worst case for WRITE operation after a long time (t_1) can be given as

$$\begin{aligned}&\text{Worst condition for WRITE}:\\ &t = 0 \text{ to } t_1: \ V_L = 0 \text{ and } V_R = 1 \ (\text{PR and NL weaken})\\ &t = t_1 + T_{CLK}: \ \text{cell flips and } V_L = 1 \text{ and } V_R = 0\\ &t = t_1 + n_{CLK}T_{CLK}: \ \text{cell is written again after } n_{CLK}\\ &\quad \text{clock cycles to have } V_L = 0 \text{ and } V_R = 1\end{aligned} \tag{9}$$

At time $t = t_1$, the cell became asymmetric (i.e., $V_{T,PR} > V_{T,PL}$ and $V_{T,NL} > V_{T,NR}$). If n_{CLK} is large (Note: V_L = '1' *and* V_R = '0' for time $n_{CLK}T_{CLK}$), the V_Ts of NR and PL will also increase due to PBTI and NBTI, respectively. In that case, cell might become symmetric again with similar V_T increase in both the pull-down FETs and pull-up FETs. However, worst condition for WRITE will occur if 'n_{CLK}' is small.

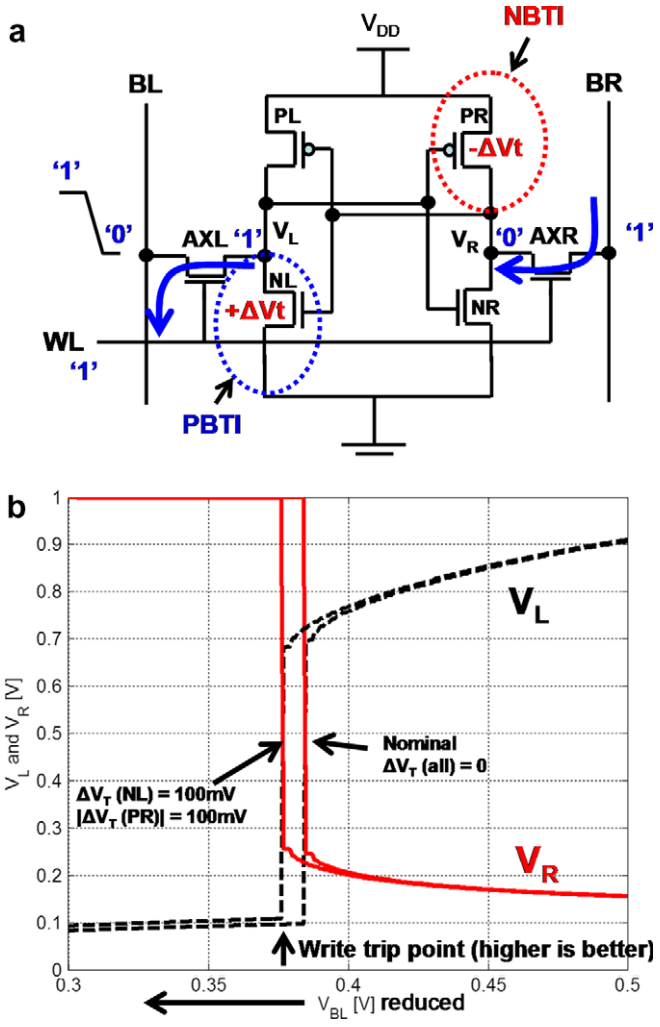

Fig. 7. (a) Worst-case scenario for WRITE ability; (b) calculation of writability as minimum BL voltage necessary to flip the cell.

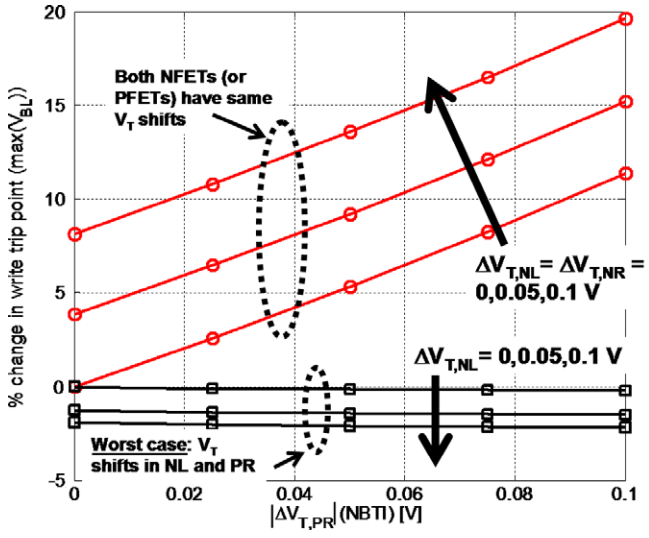

Fig. 8. Writability of a cell in worst-case as well as considering equal effect of NBTI (or PBTI) on both the PFETs (or NFETs).

Writability of a cell is measured by how much the voltage at BL needs to be lowered to flip the cell (Fig. 7b). To measure writability, we start with READ condition, say, $V_{BL} = V_{BR} =$ '1', $V_L =$ '1', $V_R =$ '0' and wordline is activated. Then, BL is pulled down to '0', to write '0' at the left node (as shown in Fig. 7b). V_L is discharged by AXL (while PL is trying to maintain it at '1') and V_R is charged by AXR (while NR is trying to maintain it at '0'). Assuming worst condition, PL and NR are NOT degraded (Fig. 7a). As V_R rises above the threshold of NL to switch it *on* and V_L drops such that PR turns *on*, the cell flips. A higher value of V_{BL} to flip the cell implies better writability. In Fig. 8, we show the impacts of NBTI and PBTI induced V_T shifts on the writability of a cell during (1) worst-case as discussed above, and (2) symmetric degradation in both the PFETs and/or NFETs. In later case, writability improves as also shown in [7]. However, in worst case, writability does not improve; instead it degrades marginally with PBTI while NBTI has negligible effect on it.

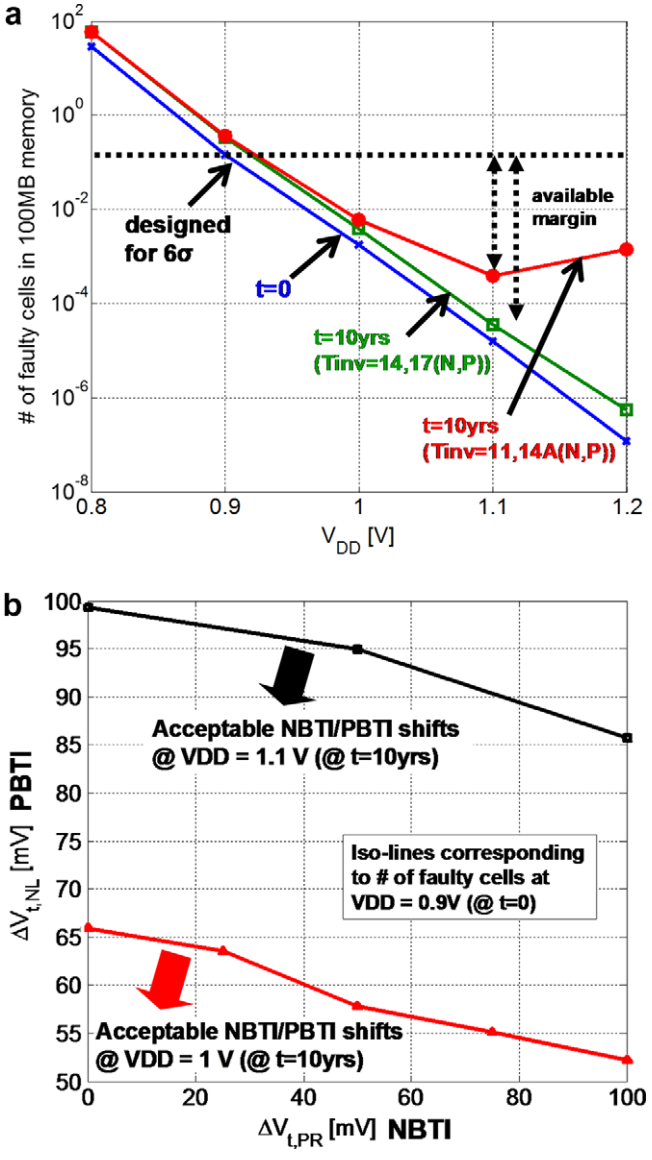

Fig. 9. (a) Number of faulty cells (under worst-case static stress) for different V_{DD} at $t = 0$ and $t = 10$ years assuming the degradation as shown in Fig. 3(a) and (b) for different T_{inv}. $V_{DD} = 0.9$ V is chosen to meet six-sigma requirement at $t = 0$. Using higher V_{DD} lowers the failure at $t = 0$, giving us acceptable margins for NBTI and PBTI shifts; (b) For un-scaled T_{inv} (14,17Å(NFET,PFET)): Acceptable combinations of NBTI/PBTI for $V_{DD} = 1$, 1.1V @ $t = 10$ years to achieve iso-failure as $V_{DD} = 0.9$ V @ $t = 0$.

This is because NFETs (NL or NR) are designed to be stronger than PFETs (PL and PR) to improve READ stability. Further, in worst case, for V_T increase of even 100 mV, the degradation in writable bit-line voltage is negligibly small. Hence, we do not consider writability degradation as a major concern.

4. Discussion

4.1. Long-term vs. short-term stability

A higher supply voltage has been used for SRAM than for logic to improve cell stability. However, a higher supply voltage results in larger stress oxide field, resulting in a larger ΔV_{MAX} in (1). This may result in a larger long-term V_T shift. Also, the degradation rate, τ, increases with the stress oxide field, causing a faster shift in V_T. As a result, the larger V_T shift may result in poorer cell stability in the long run than using lower supply voltage. In Fig. 9a, we show the impact of increasing V_{DD} on number of faulty cells at time $t = 0$ and $t = 10$ yrs. The NBTI (Fig. 3a) and PBTI (Fig. 3b) degradations in SiO_2/HfO_2 devices with TiN and Re as gates are used for computing failure at $t = 10$ yrs. The PBTI effect in these devices is greatly reduced. However, if one were to use a technology with potentially higher PBTI impact (say FUSI NiSi gated FETs with HfSiO as gate dielectric [3]), we would need to choose the supply voltage in order to meet six-sigma requirement at $t = 10$ yrs. Further, T_{inv} will be scaled with technology scaling resulting in increased V_T degradation (Fig. 3a and b). That V_T degradation, along with assuming the same sensitivities of NBTI and PBTI to failures as shown in Fig. 6a, we obtain the increase in failures with scaled T_{inv} (shown in Fig. 9a). Please note that this exercise is solely for understanding the trends. Numerical values may change for different technologies. It can be observed that with scaled T_{inv}, increasing V_{DD} may not always improve cell stability due to enhanced time-dependent V_T degradation.

For comparison, we look at acceptable NBTI and PBTI effect if higher V_{DD} (>0.9 V) is used. Note that higher V_{DD} will give lower than acceptable failure rate at $t = 0$ (as seen from Fig. 9a). Higher V_{DD} will also result in higher leakage which we have not addressed in this work. With higher V_{DD}, we increase the threshold-voltages of NFET and PFET such that after 10 yrs, we still meet the six-sigma requirement we set at $t = 0$. As seen from Fig. 9b, using $V_{DD} = 1$ V allows us to have a technology which results in, say, up to 100 mV V_T shift due to NBTI and up to 52 mV V_T shift due to PBTI after 10 yrs. Other combinations of NBTI/PBTI (e.g., $V_{T,\,NBTI} = 50$ mV and $V_{T,\,PBTI} = 58$ mV) are also possible to achieve same number of fails for $V_{DD} = 1$ V @ $t = 10$ yrs. Hence, depending upon the technology, supply voltage should be chosen to meet desired fail requirements during the life of an IC.

5. Conclusions

SRAMs are driven towards six-sigma immunity to errors. Hence, time-dependent threshold-voltage increase due to NBTI and PBTI can result in reduced stability of an SRAM cell resulting in faulty cells as the age of an SRAM array increases. A faulty cell might result in false data resulting in whole memory failure. Hence, worst-cases of threshold-voltage shift resulting in a cell failure need to be analyzed and accounted. In this work, we provide circuit insights into worst-case conditions resulting in functional failure of an SRAM cell. We show that technique of analyzing cell stability under process induced variations, such as static noise margin, is not sufficient while analyzing the combined effects of NBTI and PBTI. Hence, proper failure analysis (such as Monte-Carlo simulations) needs to be carried out. In worst case static stress, NBTI and PBTI degrade the stability during READ (significantly) and WRITE (mar-

ginally) operations. Further, we analyzed the trade-off between short-term stability (due to process variations) and long-term stability (due to NBTI/PBTI) to achieve desired six-sigma confidence in functionality. We showed that higher V_{DD} might be required, depending on the technology, to achieve desired stability during the life of an IC.

Acknowledgements

This work was performed by the Research Alliance Teams at various IBM Research and Development Facilities. Jae-Joon Kim, Rahul Rao, and Ching-Te Chuang were partially supported by DARPA Contract HR0011-07-9-0002.

References

[1] Alam M et al. A comprehensive model of PMOS NBTI degradation. Microelectron Reliab 2005;45(1):71–81.
[2] Zafar S et al. Charge trapping related threshold voltage instabilities in high permittivity gate dielectric stacks. In: JAP 2003.
[3] Zafar S et al. A comparative study of NBTI and PBTI (charge trapping) in SiO_2/HfO_2 stacks with FUSI, TiN, re gates. In: VLSI tech. sym; 2006.
[4] Mukhopadhyay S et al. Modeling of failure probability and statistical design of SRAM array for yield enhancement in nanoscaled CMOS. In: TCAD 2005, December. p. 1859–80.
[5] Bernstein K et al. High-performance CMOS variability in the 65-nm regime and beyond. IBM J Res Dev 2006;50:433–49.
[6] Lin JC et al. Prediction and control of NBTI-induced SRAM vccmin drift. In: IEDM 2006.
[7] Lin JC et al. Time dependent vccmin degradation of SRAM fabricated with high-*k* gate dielectris. In: IRPS 2007. p. 439–44.
[8] Kang K et al. Impact of negative-bias temperature instability in nanoscale SRAM array: modeling and analysis. In: TCAD 2007. p. 1770–81.
[9] Zafar S. Statistical mechanics based model for negative bias temperature instability induced degradation. In: JAP 2005.
[10] Kanj R et al. Mixture importance sampling and its application to the analysis of SRAM designs in the presence of rare failure events. In: DAC 2006. p. 69–72.
[11] Denais M et al. Interface traps and oxide traps under NBTI and PBTI in advanced CMOS technology with a 2 nm gate-oxide. Integrated reliability workshop final report, October 2003. IEEE International. p. 1–6.
[12] Grasser T et al. Simultaneous extraction of recoverable and permanent components contributing to bias-temperature instability. In: IEDM 2007. p. 801–4.
[13] Taur Y et al. Fundamentals of modern VLSI devices, Cambridge, MA.