Modeling read SNM considering both soft oxide breakdown and negative bias temperature instability

Behrouz Afzal*, Behzad Ebrahimii, Ali Afzali-Kusha, and Hamid Mahmoodib

*Nanoelectronics Center of Excellence, School of Electrical and Computer Engineering, University of Tehran, College of Engineering, Tehran, Iran

**Department of Electrical and Computer Engineering, San Francisco State University, CA, USA

**Corresponding author. Tel.: +98 21 8208 4920; fax: +98 21 8877 8690. E-mail address: afzali@ut.ac.ir (A. Afzali-Kusha).

1. Introduction

Reducing the operating voltage with each new technology node has been slowed down because of reliability concerns [1]. The gate oxide, however, has become steadily thinner with the scaling [1]. This results in larger electric fields in the gate oxide which eventually forms traps in the oxide. As more traps are formed, they start to overlap and form a conduction path between the gate and channel leading to gate tunneling current. This type of breakdown which is known as soft oxide breakdown (SBD) can deteriorate the functionality of SRAM cells [2,3].

Larger vertical electric fields can cause another reliability problem for PMOS transistors which is called negative bias temperature instability (NBTI) [4]. Under negative gate biases, high energy holes break Si–H bonds at the Si–SiO2 interface leaving interfacial traps. These traps increase the threshold voltage of the device, and hence, affecting the performance of the transistor. This may consequently deteriorate the stability of circuits such as SRAM cells [5].

The effect of NBTI on the stability and performance of SRAM cells have been investigated in the literature [5–7]. It has been shown that the most important parameter of SRAM cells which degrades with NBTI is the read stability (read SNM) [5]. The effect of SBD on the performance and functionality of SRAM cells has been discussed in [3,8–12]. Its effect on the read stability of SRAM cell in a 90 nm technology has been discussed in [8] which shows SBD is the dominant factor in the read stability degradation. To the best of our knowledge, no theoretical framework has been presented to model the read SNM considering the SBD effect.

In this work, first, we model the read SNM considering SBD, and then investigate the combined effect of NBTI and SBD on the read SNM. The reminder of the paper is organized as follows. In Section 2, we present the models used for the SBD as well as NBTI effects. In Section 3, we derive a read SNM model considering the SBD effect. In Section 4, we investigate the accuracy of the model by comparing its results to those of HSPICE and examine the combined NBTI and SBD effects on the read SNM with this model. Finally, Section 5 concludes the paper.

2. SBD and NBTI modeling

The soft oxide breakdown increases the gate leakage. In [8], a model for the gate current, \(I_g\), after the occurrence of the soft breakdown has been proposed as

\[
I_g = I_0 \exp(tGR)
\]

Here, \(I_0\) is the initial oxide tunnel current, \(t\) is the elapsed time, and \(GR\) is the defect current growth rate which has an exponential relation with the stress voltage and oxide thickness as

\[
GR = K_1 \exp(\theta_1 V_g - \theta_2 T_{ox})
\]

where \(V_g\) is the gate stress voltage, \(T_{ox}\) is the gate oxide thickness, and \(\theta_1, \theta_2,\) and \(K_1\) are constants which may be found from experimental data [8].

The breakdown may occur between the gate and diffusion region (source or drain) or between the gate and the channel of the transistors. For a conventional 6T SRAM cell, experimental data shows that the pull down source breakdown causes more severe stability degradation [12]. Thus, to study the SBD effect, we consider only the gate–source breakdown of the pull down transistor.
under stress (NR in Fig. 1). Using Eq. (1), the SBD is modeled as a resistor \( R_{\text{SBD}} \) given by [8]

\[
R_{\text{SBD}} = \frac{V_{\text{dd}}}{I_0} \exp(-tGR)
\]

where \( V_{\text{dd}} \) is the supply voltage.

Next, we include the NBTI effect which increases the threshold voltage of PMOS transistors as a function of time \( t \). The increase, which is denoted by \( \Delta V_{\text{th}} \), may be modeled by the DC reaction–diffusion (RD) framework as [4]

\[
\Delta V_{\text{th}} = K_{\text{DC}} t
\]

where \( K_{\text{DC}} \) is a constant which depends on the gate–source bias \( V_{gs} \), temperature, and other technology parameters. Fig. 2 shows the change in \( V_{th} \) due to NBTI using the reaction diffusion framework which has been calibrated with published data for a 32 nm technology node [13].

### 3. Read SNM modeling considering SBD

To model the read SNM, consider a 6T SRAM cell during the read operation as is shown in Fig. 1 where static noise sources \( V_n \) and soft oxide breakdown resistance \( R_{\text{SBD}} \) have been added. Gate-to-Source soft breakdown of NR is modeled as a linear resistor between the gate and the source of NR. Such breakdown has been the main source of the read SNM degradation in the SRAM cell [12]. Hence, we are considering only this scenario of breakdown in our analysis. In addition, the chance of multiple breakdown events is quite low, and hence, we ignore multiple breakdown scenarios.

Our goal is to calculate the read SNM considering the SBD effect. To simplify the derivation, we use the simple square law models as

\[
I_D = \begin{cases} 
\frac{1}{2} \beta (V_{gs} - V_{th}^s)^2 & \text{saturation} \\
\beta V_{th}^s (V_{gs} - V_{th}^s - \frac{1}{2} V_{th}^s) & \text{linear}
\end{cases}
\]

where \( V_{th}^s \) and \( V_{th} \) are the drain–source and threshold voltages, respectively, and

\[
\beta = \mu C_{ox} \frac{W}{L}
\]

As will be seen later, the accuracy will be improved by determining the model parameters based on fitting the model results to the simulation results.

Using the KCL equation at the node \( L \) and assuming \( NL \) and \( PL \) to operate in the saturation and linear regions, respectively, (see, e.g., [14]), one may write

\[
(V_{gs-NL} - V_{th-NL})^2 + sV_{gs-NR} - 2mV_{dd-PL}(V_{gs-NL} - V_{th-NL} - \frac{1}{2} V_{th-NL}) = 0
\]

where

\[
s = \frac{2}{\beta_{NL} R_{\text{SBD}}}
\]

\[
m = \frac{\beta_{PL}}{\beta_{NL}}
\]

The parameter \( s \) is related to SBD effect. It has been shown that the transfer characteristics of \( V_{gs-NL}, V_{gs-NR} \) have a fairly constant slope around its operating point where \( NR \) is in the linear region [14]. The linear approximation of this characteristic may be expressed as [14]

\[
V_{gs-NR} = V_0 - k V_{gs-NL}
\]

where

\[
V_0 = kV_{ws} + \frac{1 + z}{1 + z} \frac{V_t}{1 - k/V_t}
\]

\[
k = \left( \frac{z}{z+1} \right) \left( \sqrt{\frac{z+1}{z+1-V_{gr}^2}} - 1 \right)
\]

\[
z = \frac{\beta_{NR}}{\beta_{AR}}
\]

\[
V_s = V_{dd} - V_{th-NL}
\]

\[
V_t = V_s - \left( \frac{z}{z+1} \right) V_{th-AR}
\]

To increase the accuracy, the parameters in Eq. (10) may also be found by fitting the model predictions to the simulation results. Next, we write the KVL equations as

\[
V_{gs-NL} = V_n + V_{gs-NR}
\]

\[
V_{gs-PL} = V_{dd} - V_n - V_{gs-NR}
\]

\[
V_{gs-PL} = V_{dd} - V_n - V_{gs-NL}
\]

\[
V_{gs-NL} = V_n + V_{gs-NR}
\]

Substituting Eqs. (10)–(19) into Eq. (7) yields a quadratic equation with respect to \( V_n \) as

\[
(1 - m) \cdot V_n^2 + 2(1 + m) \cdot V_{th-NL} - V_0 = 0
\]

SNM is equal to \( V_n \) when the condition of coinciding roots for the quadratic equation is satisfied [14]. Therefore,

\[
\text{SNM} = -V_0 + k V_{gs-NL} - \frac{V_{th-NL} + mV_{dd} - mV_{th-NL}}{m - 1}
\]
In this equation, \( V_{gs-NL} \) can be found by setting the delta of the quadratic equation (Eq. (20)) equal to zero. The solution of this equation is approximately given by

\[
V_{gs-NL} \approx \left( \frac{2V_0 + V_{thp}}{k + 1} - \frac{s}{m(k + 1)^2} \right)
\]

\[
\times \left( 1 - \frac{1}{4} \frac{(2V_0 + V_{thp})^2 + (2V_{thp} - 2V_0 - V_{thi})^2}{m(k + 1)^2} \right)
\]

(22)

Taking the derivative of SNM with respect to \( s \), we obtain

\[
\frac{\partial \text{SNM}}{\partial s} = \frac{1 + \frac{1}{4} \frac{(2V_0 + V_{thp})^2 + (2V_{thp} - 2V_0 - V_{thi})^2}{m(k + 1)^2}}{m(k + 1)^2}
\]

(23)

\[ a = \lim_{R_{SBD} \to \infty} \frac{\text{SNM}}{R_{SBD}} \]

(24)

In Fig. 3, we have shown the SBD induced variations in the read SNM multiplied by SBD resistance squared for 45, 32 and 22 nm technologies [14]. The results, which were obtained from HSPICE simulations, show that this quantity closely follows a linear relationship with \( R_{SBD} \) for values which SNM is positive (see Fig. 4 for these values of \( R_{SBD} \)). The dependence is valid for different threshold voltages of the PMOS transistor affected by NBTI (PL in Fig. 1). Therefore, we can use the following linear relationship for modeling this dependence:

\[
\text{SNM}(R_{SBD}) = \gamma + \lambda \cdot R_{SBD}
\]

(25)

The expression may be used for simplifying the relation of SNM and \( R_{SBD} \). Next, we derive the expression for the coefficients \( \lambda \) and \( \gamma \) of Eq. (25).

Taking the derivative of Eq. (25) with respect to \( R_{SBD} \), we obtain

\[
\frac{\partial \text{SNM}}{\partial R_{SBD}} = \frac{2\gamma - \lambda \cdot R_{SBD}}{R_{SBD}^2} = \frac{1}{R_{SBD}^2} \left( a + \frac{b}{R_{SBD}} \right)
\]

(26)

We obtain the parameters \( a \) and \( b \) as

\[
a = -\lambda = \lim_{R_{SBD} \to \infty} \frac{\text{SNM}}{R_{SBD}^2}
\]

(27)

and

\[
b = -2\gamma = \frac{\partial \left( \frac{\text{SNM}}{R_{SBD}} \right)}{\partial \left( \frac{1}{R_{SBD}} \right)}
\]

(28)

One may also find \( a \) by using the more exact equation of Eq. (24) (and Eq. (27)) as

\[
a = \frac{2k \left( 1 + \frac{1}{4} \frac{(2V_0 + V_{thp})^2 + (2V_{thp} - 2V_0 - V_{thi})^2}{m(k + 1)^2} \right)}{\beta_{PL}(k + 1)^2}
\]

(29)

Similarly, \( b \) can be found as

\[
b = 8k \left( \frac{1}{4} \frac{(2V_0 + V_{thp})^2 + (2V_{thp} - 2V_0 - V_{thi})^2}{m(k + 1)^2} \right) \frac{1}{\beta_{PL}(k + 1)^2} \left( 2V_0 + V_{thp} - \frac{1}{\beta_{PL}(k + 1)^2} \right)
\]

(30)

As \( R_{SBD} \) normally has a relatively large value, \( b \) can be approximated as

\[
b \approx 8k \left( \frac{1}{4} \frac{(2V_0 + V_{thp})^2 + (2V_{thp} - 2V_0 - V_{thi})^2}{m(k + 1)^2} \right) \frac{1}{\beta_{PL}(k + 1)^2} \left( 2V_0 + V_{thp} \right)
\]

(31)

Rewriting Eq. (25) (with parameters \( a \) and \( b \)) yields

\[
\text{SNM} = \text{SNM}(R_{SBD} = \infty) - \frac{a}{R_{SBD}} - \frac{b}{2 \cdot R_{SBD}^2}
\]

(32)

For the first term in Eq. (32), we may make use of accurate models for SNM which are functions of the threshold voltages of transistors (see, e.g., [16–18]). Note that these models do not consider SBD. To include NBTI, we add \( AV_{th} \) due to the NBTI effect using the models reported in the literature (e.g., [13]). The other terms in Eq. (32), which are attributed to SBD, can be found from our expressions for \( a \), \( b \), and \( R_{SBD} \). There are other models for \( R_{SBD} \) which also may be used in our model (see, e.g., [8]).
It should be noted that we could directly use Eq. (21) for the calculation of SNM. However, since its derivation was based on the simple square law model for the \(I-V\) characteristic, it would not have a very high accuracy.

4. Results and discussion

To investigate the accuracy of the proposed model, we compare the model predictions with those of HSPICE simulations for 22, 32, and 45 nm technologies [15]. Fig. 4 which presents the comparison of the read SNM versus \(R_{\text{RSBD}}\) reveals a very good accuracy for the model. Among the technologies considered here, the accuracy for the 22 nm technology is slightly lower. This may be attributed to the degradation of the accuracy of the approximation of Eq. (10) where we assumed a linear relation between the drain-to-source and gate-to-source voltages of \(NR\). The accuracy degrades as the technology scales more. Even for this technology, the error is very small (e.g., the mean error for the case of “10% PMOS \(V_{\text{th}}\) increase” is \(\sim 4.5\%\)). The model enables a fast estimation of SNM as compared to the simulation method.

To evaluate the importance of considering both NBTI and SBD, we have plotted the difference of the read SNM changes for the case of considering both NBTI and SBD with those obtained from the addition of the read SNM changes when considering the NBTI and SBD effects separately with threshold voltages of \(PL\) as the running parameter for (a) 22, (b) 32, and (c) 45 nm technologies.
as is modeled by Eq. (32). The dependence, expressed by this equation, is not a simple addition of the independent NBTI and SBD terms. The coefficients for the SBD terms \(a\) and \(b\) are themselves functions of the PMOS threshold voltages, and hence, for a better accuracy both the NBTI and SBD effects should be considered together. This can be justified by noting that the slope in Eq. (25) \(\lambda\) becomes more negative when the NBTI effect increases, and consequently, \(a (-\lambda)\) becomes higher.

Thus, \(\Delta\text{SNM}\) due to SBD becomes more which is also apparent from Eq. (32) (in which the second term becomes larger). It is also apparent from Eq. (32) that \(\Delta\text{SNM}\) due to SBD is more sensitive to \(a\) (PMOS threshold voltage) as \(R_{\text{SBD}}\) decreases. Fig. 6 shows \(\lambda\) as a function of the change in the threshold voltage of PL for the three technologies. This graph demonstrates that increasing the NBTI effect makes the SBD effect more detrimental on SNM verifying our previous observation in Fig. 5.

Next, we plot the read SNM as a function of the stress time for the 32 nm technology. For this graph, we used the data in Fig. 2 for the \(V_{th}\) drift due to NBTI, Eq. (3) for the SBD resistance calculation, and three arbitrary values of \(GR\) as 1.6, 3.2, and \(6.4 \times 10^{-8}/s\) (we did not have access to industrial data which is process dependent). As shown in Fig. 7a, different growth rates result in different times for the onset of substantial change in the read SNM. Fig. 7b shows the difference of the read SNM decrease values when considering both NBTI and SBD simultaneously with those obtained from the addition of the results when considering the NBTI and SBD effects separately versus stress time for the 32 nm technology. The parameter \(GR\) is the running parameter. As the results suggest, the difference in the \(GR\) values does not have a major effect on the maximum of the difference. This is justified by noting that the change of the threshold voltage of PL almost saturates when the stress time increases (see Fig. 2).

We also study the effect of the supply voltage on the read SNM degradation. For this purpose, the read SNM values for different supply voltages versus time have been shown in Fig. 8. For these results, the changes of \(V_{th}\) due to NBTI for the supply voltages of 0.8 and 0.9 V have been obtained from the analytical expressions.
given in [19] and the SBD resistance from Eq. (3) by assuming that GR increases 5 dec/V with \( V_g \) [8]. The results show that while for smaller supply voltages the initial SNM value is lower, the rate of the SNM degradation due to NBTI and SBD is smaller too. Hence, for instance, after 8 (8.5) years, the SNM value for \( V_{dd} \) equal to 0.9 V (0.8 V) becomes more than that for \( V_{dd} \) equal to 1 V. After 9 years, the SNM values for \( V_{dd} \) equal to 0.8 V are larger than those for the other two supply voltages.

Finally, Fig. 9 compares the Cumulative Distribution Function (CDF) of the read SNM versus minimum read SNM (RSNM0) obtained using 15,000 HSPICE Monte Carlo simulations and our proposed model under process variations and aging effects (NBTI and SBD) for the 32 nm technology. We consider the threshold voltage of transistors due to process variations as Gaussian random variables [20]. The 3\( \sigma \) of the threshold voltages were set to 20% of their nominal values [18]. The CDF value at each RSNM0 shows the percentage of the cells whose read SNM values are smaller than RSNM0. The figure shows, for example, if the target read SNM is assumed to be 30 mV, the percentage of the cells with smaller read SNM values than 30 mV are about 2% and 45% after 8 and 9 years, respectively. The comparison reveals a very good accuracy for the model which is evaluated in a very short period of time due to its analytical nature. Therefore, when we study the impact of process variations and aging on SRAM cell, using the proposed model is a very efficient method of calculating the read SNM compared to HSPICE simulations.

5. Conclusion

In this work, we proposed a read SNM model which considered the soft oxide breakdown effect. The model which used a resistance for modeling this effect, added two terms to the expression for the read SNM model. The NBTI effect was considered in the model by including the change of the threshold voltages in the SNM model which was a function of the threshold voltages. The accuracy of the model was verified by comparing its prediction with those of simulation for 45, 32, and 22 nm technologies. The comparison revealed a very good accuracy for the model. In addition, the results showed that NBTI aggravates the SBD effect on the read SNM. This suggested that the effect of NBTI and SBD should be included in the model concurrently as has been performed in our model.

Acknowledgements

The first three authors acknowledge the financial support by the Iranian National Science Foundation (INSF).

References


