

Analytical Modeling of Read Margin Probability Distribution Function of SRAM Cells in Presence of Process Variations and NBTI Effect

Behrouz Afzal¹, Ali Afzali-Kusha¹, and Massoud Pedram²

¹Nanoelectronics Center of Excellence, School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran

²Department of of Electrical Engineering, University of Southern California, Los Angeles, CA, U.S.A.

Abstract— In this work, we present an analytical model for calculating the Read Margin of SRAM cells as a function of different transistors parameters. Using this model and assuming normal distribution for the threshold voltages of transistors in the presence of process variations, the probability distribution function (PDF) of the Read Margin is analytically derived. In addition, the time variation of the PDF due to the NBTI (negative bias temperature instability) effect is also considered in the model. The accuracy of the model is verified by comparing its results with those of HSPICE simulations in 45 nm and 32 nm technologies. The comparison demonstrates a very high level of accuracy for the proposed model.

Index Terms— Modeling, nano-scale, Read Margin, SRAM, Probability Distribution Function, SRAM Stability, NBTI, Process Variation.

I. INTRODUCTION

SRAM arrays occupy a large portion of System on Chip (SOC) designs and multicore processor chips [1]. To have high levels of integration density, SRAM cells are implemented by using minimum sized transistors [2]. On the other hand, with CMOS technology scaling to nanoscale (critical) feature sizes, random variations in process parameters have emerged as a major challenge in circuit design. For SRAM cells, due to the use of minimum dimensions, the effects of random variations are quite severe [2]. These variations can be divided into global (systematic) and local (totally random) variations [3]. Global variations cause the same change in the parameters of neighboring transistors in the SRAM cell. On the other hand, local variations usually change the parameters of neighboring transistors independently of each other. Only local variations can have a significant effect on the stability of SRAM cells [4]. Among different sources of parameter changes resulting in local variations, random dopant fluctuation (RDF) is the most important [5][6]. The impact of RDF in the channel region, which originates from random variations in the number and location of dopant atoms, is inversely proportional to gate area [7].

In addition to process variations, which are related to the fabricating circuits, the characteristics of transistors change over time due to aging effect which are induced by

phenomena such as NBTI (negative bias temperature instability) and HCI (hot carrier injection). Among the aging effects, NBTI, which is considered to be the most important one [8] since it affects transistors when they are sitting idle doing nothing which is most of the time, causes an increase in the PMOS threshold voltage over time. NBTI results from the depassivation of the Si-H bond at the Si-dielectric interface and diffusion of resultant hydrogen species into gate dielectric and poly Si [9]. This effect becomes important below the 130 nm technology node [10][11]. NBTI degrades the performance, functionality, and stability of SRAM cells (see, e.g., [12]).

As was discussed, both the process variations and NBTI phenomenon affect the SRAM stability parameters such as the Read Margin. These two effects transform the stability parameter to a random variable whose parameters change over time. Having an analytical model that can predict the statistical behavior of the stability parameter (including its time variation) would make the circuit optimization more efficient. There are some works in the literature that investigate the effect of process variations on SRAM parameters and failure probabilities as explained next. In [2], the read stability is investigated using simulations without presenting any analytical analysis. In another work, different sources of failure probabilities have been modeled by assuming Gaussian distributions for all stability metrics [4]. The Gaussian distribution does not necessarily model the probability distribution function (PDF) of the stability parameters. A non-analytical technique for the stability analysis which is faster than the Monte Carlo simulation approach has been proposed in [13]. Also, there are other works which have been focused on the effect of NBTI on the SRAM stability (see, e.g., [12], [14], and [15]). None of these efforts presented an analytical model for the reliability metrics.

In this work, we present an analytical model for the statistical behavior of the Read Margin in the presence of both process variations and NBTI effect. The approach starts with presenting an analytical model for the Read Margin (RM) as a stability parameter. Then, the PDF of this parameter in the presence of process variations is described analytically. Finally, the impact of NBTI on the PDF is also included in the

model. It should be noted that Read SNM is also an important stability parameter. The derivation of an analytical model for this parameter is complicated and the resulting expressions are very long and involved (see, e.g., [16]). Thus, including the effect of process variations and NBTI in its model does not provide much insight to the designer.

The rest of the paper is organized as follows. In Section II, the model for the Read Margin as a function of the threshold voltages and other transistor parameters are derived. The derivation of the PDF for the Read Margin considering both the process variation and NBTI effect is described in Section III. The results of the model predictions and its comparison to HSPICE simulations are discussed in Section IV while Section V concludes the paper.

II. READ MARGIN MODELING

In this section, we present the derivation of the model for the Read Margin as a function of device parameters. For this purpose, we start from the definition of the Read Margin and then using a simple I - V model, we derive the expression for this figure of merit.

Fig. 1 shows the schematic of a conventional 6T SRAM cell. During the read operation of the cell, the voltage at node R increases to a positive value, denoted by V_{read} , due to the voltage division between the right access transistor and the right pull down transistor. If V_{read} is higher than the trip point of the left inverter, denoted by V_{trip} , then the cell flips and a read failure occurs [8].

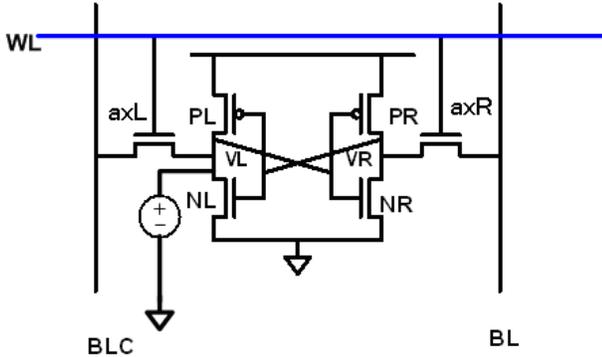


Fig. 1. Schematic of a conventional 6T SRAM cell.

We denote the Read Margin of the right (left) side of the cell by RMR (RML). The margin is calculated as the difference between V_{tripl} and V_{readr} (V_{tripr} and V_{readl})

$$RMR = V_{tripl} - V_{readr} \quad (1)$$

RM is the minimum of RML and RMR . To obtain the Read Margin analytically, first we model V_{read} . For the case of RMR (VL in Fig. 1 is equal to V_{DD}), one may write the KCL at node R as

$$I_{ds,nr} = I_{ds,axr} \quad (2)$$

where $I_{ds,nr}$ is the current of transistor NR in the saturation region and $I_{ds,axr}$ denotes the current of transistor axR in the linear region. To make the mathematics involved manageable, we use the simple square law model for the currents which is not accurate enough for these technologies. In the process of our model derivation, we make some approximations to

compensate for the errors induced by the use of this simple model. Using (2), one may write

$$(V_{DD} - V_{th,nr} - \frac{V_{readr}}{2})V_{readr} = \frac{k_{ax,n}}{2}(V_{DD} - V_{th,axr} - V_{readr})^2 \quad (3)$$

where V_{DD} is the supply voltage, $V_{th,i}$ is the absolute value of the threshold voltage of the transistor i (here, the right access transistor), and for any two transistors i and j , we define $k_{i,j}$ as

$$k_{i,j} = \frac{\mu_i \cdot c_{ox,i} \cdot \frac{W_i}{L_i}}{\mu_j \cdot c_{ox,j} \cdot \frac{W_j}{L_j}} \quad (4)$$

In (4), μ_i is mobility of transistor i , $c_{ox,i}$ denotes the oxide capacitance per unit area, and L_i (W_i) is the length (width) of transistor i .

Solving (3) for V_{readr} , one obtains

$$V_{readr} = \frac{V_{DD} - V_{th,nr} + k_{ax,n}(V_{DD} - V_{th,axr})}{k_{ax,n} + 1} \cdot (1 - \sqrt{1 - \beta}) \quad (5)$$

where

$$\beta = \frac{k_{ax,n}(k_{ax,n} + 1)(V_{DD} - V_{th,axr})^2}{(V_{DD} - V_{th,nr} + k_{ax,n}(V_{DD} - V_{th,axr}))^2} \quad (6)$$

Our simulations for the 65 nm, 45 nm, and 32 nm technologies show that $\beta \ll 1$. Hence, we can approximate V_{readr} as

$$V_{readr} \cong \frac{k_{ax,n}}{2} \cdot \frac{(V_{DD} - V_{th,axr})^2}{(V_{DD} - V_{th,nr}) + k_{ax,n}(V_{DD} - V_{th,axr})} \quad (7)$$

This approximation underestimates the voltage.

Now, note that in the read operation, a stronger pull down NR compared to axR and PR pull ups is desired, and hence, the width of NR is larger than those of axR and PR [17]. Then $k_{ax,n} < 1$, and hence, we can approximate V_{readr} as

$$V_{readr} \cong \frac{k_{ax,n} \cdot (V_{DD} - V_{th,axr})^2}{2(V_{DD} - V_{th,nr})} \quad (8)$$

which overestimates the voltage. Because the first (second) approximation underestimates (overestimates) the voltage, the accuracy of (8) turns out to be acceptable. To increase the accuracy of the approximation further, $k_{ax,n}$ may be treated as a fitting parameter in our model.

To obtain the expression for V_{trip} at the left side of the cell, we set $VL = VR = V_{trip}$ and write the KCL at the node L . This leads to

$$\begin{aligned}
V_{trip} = & \frac{1}{k_{ax,p} - k_{n,p} + 1} \cdot (V_{DD} - V_{th,pl} + k_{ax,p} \cdot V_{DD} \\
& - k_{ax,p} \cdot V_{th,axl} - k_{n,p} \cdot V_{th,nl} + (k_{n,p} \cdot V_{DD}^2 - k_{ax,p} \cdot V_{th,pl}^2 - \\
& k_{ax,p} \cdot V_{th,axl}^2 + k_{n,p} \cdot V_{th,nl}^2 + k_{n,p} \cdot V_{th,pl}^2 + k_{ax,p} \cdot k_{n,p} \cdot V_{DD}^2 \\
& + k_{ax,p} \cdot k_{n,p} \cdot V_{th,nl}^2 + k_{ax,p} \cdot k_{n,p} \cdot V_{th,axl}^2 - 2 \cdot k_{n,p} \cdot V_{DD} \cdot V_{th,nl} \\
& - 2 \cdot k_{n,p} \cdot V_{DD} \cdot V_{th,pl} + 2 \cdot k_{ax,p} \cdot V_{th,p} \cdot V_{th,axl} \\
& - 2 \cdot k_{ax,p} \cdot k_{n,p} \cdot V_{DD} \cdot V_{th,nl} - 2 \cdot k_{ax,p} \cdot k_{n,p} \cdot V_{DD} \cdot V_{th,axl} + \\
& 2 \cdot k_{n,p} \cdot V_{th,nl} \cdot V_{th,pl} + 2 \cdot k_{ax,p} \cdot k_{n,p} \cdot V_{th,nl} \cdot V_{th,axl})^{\frac{1}{2}}
\end{aligned} \quad (9)$$

In [4], the authors considered V_{trip} as a linear function of V_{DD} , $V_{th,p}$, and $V_{th,n}$ by neglecting the effect of the access transistor. As the channel length becomes smaller, V_{trip} will decrease. Thus, the current of the access transistor increases, and hence, it should not be ignored when calculating V_{trip} (especially, as $V_{th,pl}$ becomes more negative due to the NBTI effect). The inclusion of the access transistor in the derivation makes the expression for the trip voltage complex. To reduce the complexity of the model, we approximate the polynomial under the square root of (9) by a second order polynomial (by omitting some terms and adding some other terms such that the acceptable accuracy is achieved). Therefore, one may write

$$V_{trip} = a_1 \cdot V_{DD} - a_2 \cdot V_{th,pl} + a_3 \cdot V_{th,nl} + a_4 \cdot V_{th,axl} \quad (10)$$

where the coefficients a_1 to a_4 are found from

$$a_1 = \frac{1 - \sqrt{k_{n,p} \cdot (k_{ax,p} + 1)} + k_{ax,p}}{k_{ax,p} - k_{n,p} + 1} \quad (11)$$

$$a_2 = \frac{2 \sqrt{k_{n,ax} \cdot \sqrt{k_{ax,n} \cdot (1 - k_{ax,n})}} - 1}{k_{ax,p} - k_{n,p} + 1} \quad (12)$$

$$a_3 = \frac{\sqrt{k_{n,p} \cdot (k_{ax,p} + 1)} - k_{n,p}}{k_{ax,p} - k_{n,p} + 1} \quad (13)$$

$$a_4 = \frac{\sqrt{k_{n,p} \cdot (k_{ax,p} - k_{ax,n})} - k_{ax,p}}{k_{ax,p} - k_{n,p} + 1} \quad (14)$$

If needed, to increase the accuracy of the model for sub-32 nm technologies, the coefficients a_1 to a_4 may be considered as fitting parameters of the model. The Read Margin values for 32 nm, 45 nm, and 65 nm technologies are plotted in Fig. 2 where the maximum error is almost 4%. Note that for these results, the parameters $k_{ax,n}$ and a_1 to a_4 have not been treated as fitting parameters.

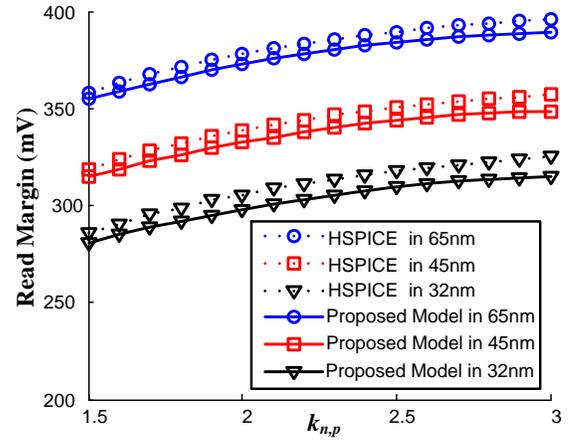


Fig. 2. Read Margin versus $k_{n,p}$ for 32 nm, 45 nm, and 65 nm technologies.

By substituting (8) and (10) into (1), we obtain RMR . RML may be found similarly.

III. PDF MODELING OF READ MARGIN

In this section, we model the variation of the Read Margin in the presence of process variations. Here, we only focus on the threshold voltage variations which are modeled by Gaussian distributions [2]. The read voltage obtained in the previous section was modeled as a nonlinear function of $V_{th,ax}$ and $V_{th,n}$. To obtain the Read Margin PDF, we should approximate V_{read} in (8) as a linear function of the threshold voltages as we will see later in this section. This enables us to find the Read Margin PDF easily (for the case of Gaussian distribution). It should be noted that both PDF's of V_{trip} and V_{read} on each side are modeled by Gaussian distributions. Therefore, RMR and RML will also follow the same distribution. The RM which is the minimum of RMR and RML , however, will not have a Gaussian distribution. In [4], the PDF of RM was assumed to be Gaussian.

Now, we proceed with linearizing the read voltage in terms of the threshold voltages. Rewriting V_{read} as

$$V_{read} = K \cdot T \quad (15)$$

where K is a constant and

$$T = \frac{x^2}{y} \quad (16)$$

Here,

$$x = V_{DD} - V_{th,axr} \quad (17)$$

$$y = V_{DD} - V_{th,nr} \quad (18)$$

which shows that statistical behavior of these two parameters are related to the statistical behavior of $V_{th,axr}$ and $V_{th,nr}$. If μ_n and σ_n are the mean and variance for the NMOS threshold voltage, x and y are Gaussian variables with σ_n and $\mu_d = V_{DD} - \mu_n$. By differentiating (16), we have

$$dT = \frac{2x \cdot dx}{y} - \frac{x^2 \cdot dy}{y^2} \quad (19)$$

Because x and y are random variables with the same mean and variance and because $\mu_d \gg \sigma_n$, one may assume $y \approx x$ and write

$$dT \approx 2dx - dy \quad (20)$$

Integrating (20) yields

$$T \approx 2x - y + c \quad (21)$$

where c is the integration constant. This relation expresses the Read Margin as a linear function of two Gaussian variables, and hence, T should also be a Gaussian variable. To increase the accuracy of this approximation, we obtain c such that statistical behaviors of T in (16) and (21) become very close to each other. For this purpose, first, we calculate the PDF of T from the PDFs of x and y using

$$P_T(T)dT = \int_{y=-\infty}^{y=\infty} P_y(y) \left(P_x(x = \sqrt{T \cdot y}) + P_x(x = -\sqrt{T \cdot y}) \right) dx dy \quad (22)$$

where P_x , P_y , and P_T are the PDFs of x , y , and T , respectively. Next, for a given y , we should express x and dx in terms of T and dT . First, using (16) we can obtain

$$dT = \frac{2x dx}{y} \quad (23)$$

Obtaining x from (16) and substituting it into (23) yields

$$dx = \frac{1}{2} \sqrt{\frac{y}{T}} dT \quad (24)$$

One can use (24) to express the right hand side of (22) in terms of variable y only. Now, to obtain $P_T(T)$ analytically, we need to use one approximation. For the region of interest where $(y - \mu_d) \ll \mu_d$, one may write

$$\sqrt{y} \approx \sqrt{\mu_d + (y - \mu_d)} \approx \sqrt{\mu_d} \cdot \left(1 + \frac{y - \mu_d}{2\mu_d}\right) \quad (25)$$

By substituting (25) into (22) and performing the integration, one obtains

$$P_T(T) = \frac{1}{2\sigma_z \cdot \sqrt{T}} \cdot \left(\phi\left(\frac{\sqrt{T} - \mu_z}{\sigma_z}\right) + \phi\left(\frac{\sqrt{T} + \mu_z}{\sigma_z}\right) \right) \quad (26)$$

where ϕ is a standard normal probability distribution function,

$$\mu_z = \sqrt{\mu_d}, \quad (27)$$

and

$$\sigma_z = \frac{\sigma_n}{\sqrt{\mu_d}} \cdot \sqrt{\frac{5}{4}} \quad (28)$$

The expression for the PDF of T given by (26) is a Non-Central-Chi-Square distribution function [18]. Here, $T = z^2$ and z is a normal variable with the mean and standard deviation of μ_z and σ_z , respectively. The mean (μ_T) and standard deviation (σ_T) of T are given by [18].

$$\mu_T = \mu_d + \frac{\sigma_n^2}{\mu_d} \cdot \frac{5}{4} \quad (29)$$

$$\sigma_T = \sqrt{5} \cdot \sigma_n \quad (30)$$

Since when $\mu_z \gg 0$ (as is our case), the Non-Central-Chi-Square function becomes similar to the Gaussian function, we use a Gaussian PDF for T (V_{read}) with the mean and standard deviation of μ_T and σ_T , respectively.

By equating the mean of the PDF of T obtained from (21) and (29), one obtains c as

$$c = \frac{\sigma_n^2}{\mu_d} \cdot \frac{5}{4} \quad (31)$$

Also, note that the standard deviation obtained from (21) is equal to the one given in (30).

To determine the accuracy of the approximate model for the read voltage PDF, we compared the PDF obtained from (21) and Monte Carlo simulations (100,000 runs) for the 32 nm and 45 nm technologies in Fig. 3. The maximum errors for the mean of V_{read} , 1%, 5%, 95%, and 99% percentiles were 1.8% (1.6%), 9.3% (3.9%), 3.6% (1.6%), 2.4% (1.2%) and 5.1% (2.8%) for the 32 nm (45 nm) technology, respectively. The comparison reveals a very good accuracy for the model.

As mentioned before, the threshold voltage of transistors are Gaussian variables. To simplify the equations, which have been utilized to calculate RM, we use

$$V_{DD} - V_{th,nl} = \mu_d + \sigma_n \cdot M_L \quad (32)$$

$$V_{DD} - V_{th,nr} = \mu_d + \sigma_n \cdot M_R \quad (33)$$

$$V_{DD} - V_{th,axl} = \mu_d + \sigma_n \cdot N_L \quad (34)$$

$$V_{DD} - V_{th,axr} = \mu_d + \sigma_n \cdot N_R \quad (35)$$

$$V_{th,pl} = \mu_p + V_{nbtic} + V_{nbtii} + \sigma_p \cdot P_L \quad (36)$$

$$V_{th,pr} = \mu_p + V_{nbtic} + \sigma_p \cdot P_R \quad (37)$$

where M_L , M_R , N_L , N_R , P_R , and P_L are standard normal variables. Also, V_{nbtic} is the value of threshold voltage increase of one of the PMOS transistors (PR in this case) due to the NBTI effect, V_{nbtii} is the excess threshold voltage increase of the other PMOS transistor due to non-symmetry in the NBTI degradation (PL in this case). The non-symmetric behavior was resulted due to assuming more stress on the left PMOS transistor. This assumption of non-symmetric behavior generalizes the formulation.

Having found the relations for V_{read} and V_{trip} , one may obtain the left and right Read Margins, respectively as

$$RML = \frac{1}{2} FL \cdot k_{ax,n} \cdot \sigma_n + \mu_{RMC} \quad (38)$$

$$RMR = \frac{1}{2} FR \cdot k_{ax,n} \cdot \sigma_n + \mu_{RMC} \quad (39)$$

where

$$\mu_{RMC} = a_1 \cdot V_{DD} + a_2 \cdot V_{nbtic} + a_2 \cdot \mu_p + (a_3 + a_4) \cdot \mu_n - \frac{1}{2} k_{ax,n} \left(V_{DD} - \mu_n + \frac{5}{4} \cdot \frac{\sigma_n^2}{V_{DD} - \mu_n} \right) \quad (40)$$

and FL and FR are two auxiliary functions defined by

$$FL = h + k_1 \cdot pL + k_2 \cdot M_L + k_3 \cdot N_L - 2N_R + M_R \quad (41)$$

$$FR = k_1 \cdot P_R + k_2 \cdot M_R + k_3 \cdot N_R - 2N_L + M_L \quad (42)$$

Here,

$$h = \frac{2a_2 \cdot V_{nbtii}}{\sigma_n \cdot k_{ax,n}} \quad (43)$$

$$k_1 = \frac{2a_2 \cdot \sigma_p}{\sigma_n \cdot k_{ax,n}} \quad (44)$$

$$k_2 = \frac{-2a_2}{k_{ax,n}} \quad (45)$$

$$k_3 = \frac{-2a_3}{k_{ax,n}} \quad (46)$$

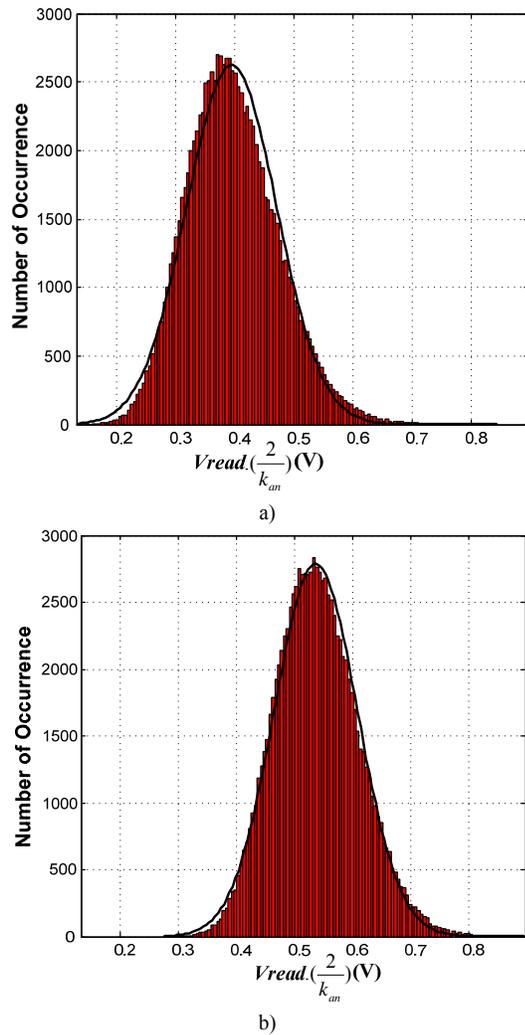


Fig. 3. The PDF results of $V_{read}(2/k_{an})$ obtained from HSPICE and proposed model for (a) 32 nm and (b) 45 nm technologies. We used 100,000 Mont Carlo simulations.

Note that since μ_{RMC} is constant, by finding the minimum values of the auxiliary functions, one can find the Read Margin. Let us define f as

$$f = \min(FL, FR) \quad (47)$$

To obtain the probability distribution function of f , we use

$$P_f(f) = P_1 + P_2 \quad (48)$$

where

$$P_1 = P[(FL = f) \text{ and } (FR > f)] \quad (49)$$

$$P_2 = P[(FR = f) \text{ and } (FL > f)] \quad (50)$$

After some manipulation, probability P_1 is obtained from (see Appendix 1)

$$P_1(f).df = \int_{P_R=-\infty}^{P_R=\infty} \int_{P_L=-\infty}^{P_L=\infty} \int_{N_L=-\infty}^{N_L=\infty} \int_{M_L=-\infty}^{M_L=\infty} P_{P_R}(P_R).P_{P_L}(P_L). \quad (51)$$

$$P_{M_L}(M_L).P_{N_L}(N_L).P_t(t = t_0) \cdot \frac{(1 + \operatorname{erf}(\frac{N_{R0}}{\sqrt{2}}))}{2}.$$

$$dP_R.dP_L.dM_L.dN_L.df$$

In order to obtain (51), we have defined a variable, denoted by t as follows:

$$t = -2.N_R + M_R \quad (52)$$

which is a normal variable with $\mu_t = 0$ and $\sigma_t = 5^{1/2}$. Also, we defined another variable t_0 which is t when $FL = f$. Using, (41) one obtains t_0 as

$$t_0 = f - h - k_1.P_L - k_2.M_L - k_3.N_L \quad (53)$$

Also, the variable N_{R0} is given by

$$N_{R0} = \frac{1}{2k_2 + k_3} (f + 2N_L - M_L - k_2.(f - h - k_1.P_L - k_2.M_L - k_3.N_L) - k_1.P_R) \quad (54)$$

For the values around zero, one only needs to keep the first few terms of the erf series expansion as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} (x - \frac{x^3}{3} + \frac{x^5}{10} + \dots) \quad (55)$$

By replacing (55) in (51) and performing the integration, the probability distribution function P_1 is obtained as

$$P_1(f) = g_1 \left[1 + \frac{2}{\sqrt{\pi}} \left(\frac{r_1.(1 + q + \frac{3}{2}q^2 + \frac{5}{2}q^3 + \dots) - r_1^3.(1 + 3q + \frac{15}{2}q^2) + \dots}{3} \right) \right] \quad (56)$$

where

$$g_1 = \frac{1}{2\sqrt{2\pi}} \cdot \frac{e^{-\frac{1}{2} \frac{(f-h)^2}{k_1^2 + 5 + k_3^2 + k_2^2}}}{\sqrt{k_1^2 + 5 + k_3^2 + k_2^2}} \quad (57)$$

$$r_1 = \frac{f.(k_1^2 + k_2^2 + k_3^2 - 6k_2 + 2.k_3 + 5) + 2h.(3k_2 - k_3)}{\sqrt{2}(2k_2 + k_3)(k_1^2 + 5 + k_3^2 + k_2^2)} \quad (58)$$

$$q = -\frac{1}{2} \frac{5k_2^4 + 5k_2^2k_3^2 - 6k_2^2 + 24k_2.k_3 + k_3^2 + 25}{(2k_2 + k_3)^2(5 + k_3^2 + k_2^2)} \quad (59)$$

Since the simulations show that $q \ll 1$, one can approximate it by zero and obtain (56) as

$$P_1(f) \approx g_1.(1 + \operatorname{erf}(r_1)) \quad (60)$$

Similarly, the probability function of “ $(FL > f)$ and $(FR = f)$ ” is obtained as

$$P_2(f) \approx g_2.(1 + \operatorname{erf}(r_2)) \quad (61)$$

where

$$g_2 = \frac{1}{2\sqrt{2\pi}} \cdot \frac{e^{-\frac{1}{2} \frac{f^2}{k_1^2 + 5 + k_3^2 + k_2^2}}}{\sqrt{k_1^2 + 5 + k_3^2 + k_2^2}} \quad (62)$$

$$r_2 = \frac{f.(k_1^2 + k_2^2 + k_3^2 - 6k_2 + 2k_3 + 5) + 2h.(3k_2 - k_3)}{\sqrt{2}(2k_2 + k_3)(k_1^2 + 5 + k_3^2 + k_2^2)} - \frac{h}{\sqrt{2}(2k_2 + k_3)} \quad (63)$$

Note that if h (which shows the non-symmetric behavior of the left and right sides of the cell) is 0, both P_1 and P_2 will become equal. Having found P_1 and P_2 , one may obtain the PDF of f in (47). Then, using (38) (or (39)), the PDF of the Read Margin can be obtained analytically. Now, we summarize the equations were used to obtain the read margin and the PDF of the read margin as was our objective.

Obtain analytical Read Margin

1-Right Read Margin (RMR) = $V_{tripl} - V_{readr}$

2- V_{tripl} obtain from (10) $\rightarrow a_1, a_2, a_3,$ and a_4 from (11 - (14)

3- V_{readr} from (8), approximate V_{readr} from (15)

$$\downarrow$$

 K is constant, T from (21)

$$\downarrow$$

 $x, y,$ and c from (17), (18), and (31)

Obtain analytical PDF of Read Margin

1-Read Margin = $\min(RML, RMR)$

$$\downarrow \quad \downarrow$$

(38) (39)

$$\downarrow \quad \downarrow$$

 μ_{RMC} from (40). FL and FR from (41) and (42)

$N_L, P_L, M_L, N_R, P_R,$ and M_R from (32 - (37)

h, k_1, k_2, k_3 from (43 - (46)

2- P_f = PDF of $\min(FL, FR) \Rightarrow$ PDF of $\min(RMR, RML)$

3- $P_f = P_1 + P_2$

$$\downarrow \quad \downarrow$$

(49) (50)

4- P_1 and P_2 from (60) and (61)

$$\downarrow \quad \downarrow$$

 r_1 and g_1 from (58) and (57)
 r_2 and g_2 from (62) and (63)

In the next section, we study the accuracy of the proposed model by comparing its results with those of Monte Carlo simulations.

IV. RESULTS AND DISCUSSION

In this section, first, we study the accuracy of the analytical model for predicting the PDF of the Read Margin. Then, we investigate on the effects of NBTI on the Read Margin PDF. Finally, we present results for the SRAM yield.

The Read Margin PDF obtained from model has been compared to that of HSPICE simulations for 45 nm and 32 nm technologies in Fig. 4. The results of HSPICE were obtained using the Monte Carlo simulations for 100,000 data points. In Fig. 5, we investigate the NBTI effect on the PDF of the Read Margin. To consider this effect, we apply the threshold voltage shift due to this effect after one and ten years on a PMOS transistor and compare the simulation results with those predicted by the proposed model. The maximum errors of these two figures for 1%, 5%, 95%, and 99% percentiles for 32 nm (45 nm) are 1.6% (1%), 2.9% (1.1%), 1.8% (1.4%), and 1.2% (1.6%), respectively. This shows a very good accuracy for the model. Note that the results for HSPICE simulations (model) were obtained in a couple of days (seconds). Also, it should be mentioned that with the elapse of time, the mean of the Read Margin decreases while its variance increases (see Fig. 5). Finally, Fig. 6 demonstrates the results of the model for the evolution of the PDF after manufacturing, one year, and ten years for both the 32 nm and 45 nm technologies.

Now, we compare the NBTI effect on the Read Margin PDF when both PMOS transistors are equally affected and when only one transistor is affected. As stated in the previous section, the threshold voltage increase of one transistor (PR) is V_{nbtic} while the value for the other PMOS transistors (PL) is $(V_{nbtic} + V_{nbt})$. We can use the Read Margin PDF obtained in the previous section to estimate the shift of the mean value of the Read Margin due to the NBTI effect for the cases where the threshold voltages have large variances. The approximation is given by (see Appendix 2)

$$\Delta\mu_{RM} \cong a_2 \cdot (V_{nbtic} + \frac{V_{nbt}}{2}) \quad (64)$$

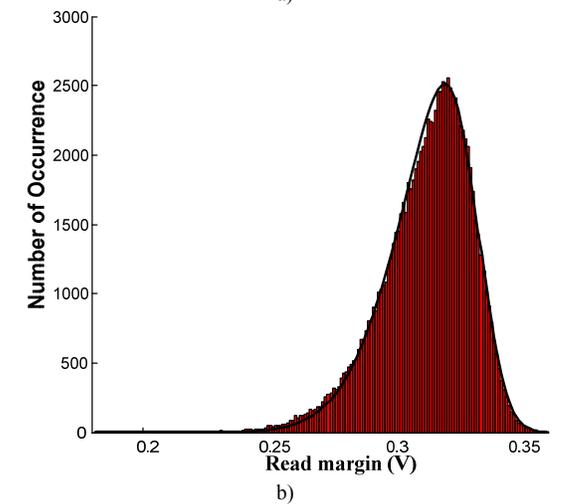
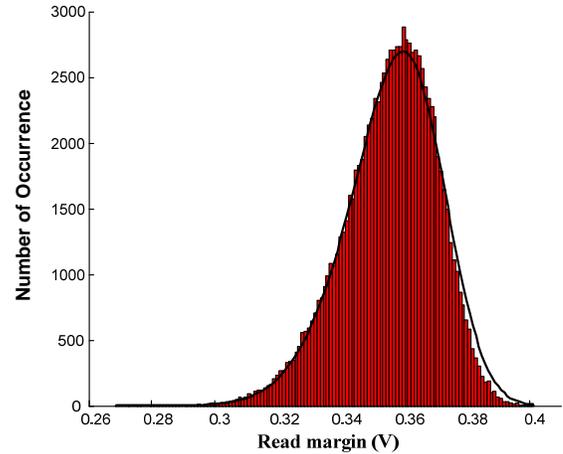
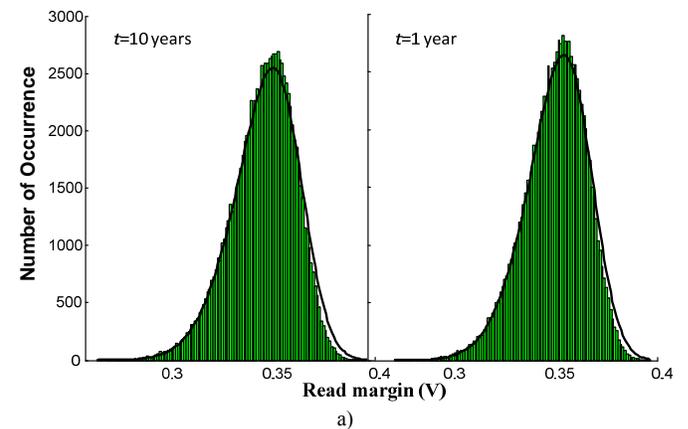


Fig. 4. The PDF of the Read Margin obtained by both the model and HSPICE simulations for a) 45 nm and b) 32 nm technologies.



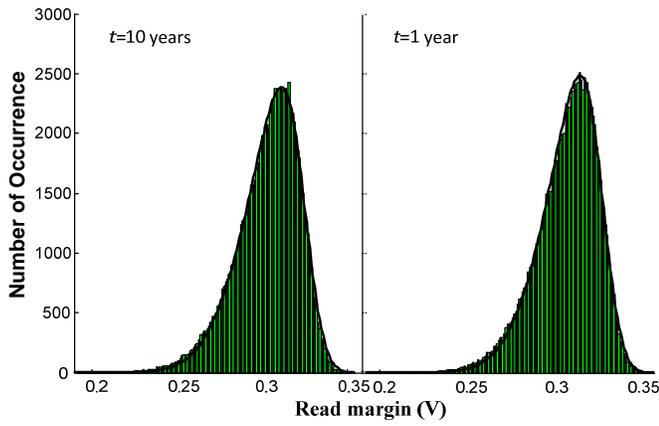


Fig. 5. The PDF of the Read Margin obtained by both the model and HSPICE simulations for a) 45 nm and b) 32 nm technologies with NBTI effect on a PMOS transistor after one and ten years.

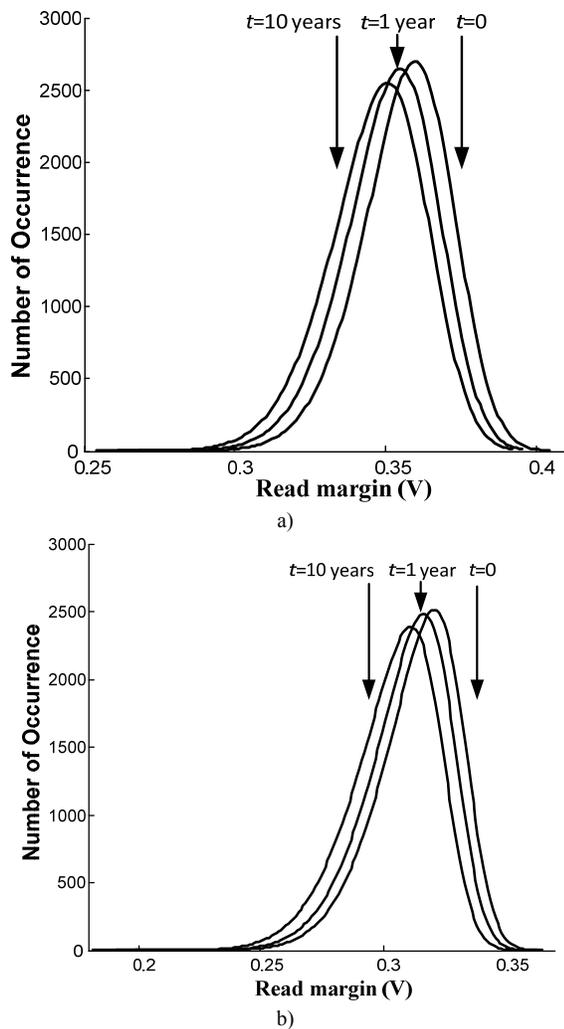


Fig. 6. The evolution of the predicted PDF of the Read Margin obtained by the model for a) 45 nm and b) 32 nm technologies with NBTI effect on a PMOS transistor after one and ten years.

Using this simple model obtained from the approximation provides us with some insights on the effect of the NBTI components on the Read Margin change over the time. Note

that (64) suggests that the change is proportional to the coefficient a_2 for the threshold voltage variation in both transistors while the coefficient is $a_2/2$ for the variation in one transistor. The threshold voltage variation due to the NBTI effect is a function of time which is given by [12]

$$\Delta V_{thp} = K_{DC} \cdot \alpha(s) \cdot t^n \quad (65)$$

where s is the probability of the stress of the transistor and α is constant prefactor. The stress condition corresponds to the case where the source-to-gate voltage of the PMOS transistor is equal to the supply voltage. When almost all the times a transistor is under stress, then $\alpha(s)$ will be approximately one and $V_{nbti} = 0$. In this case, we may write

$$\Delta \mu_{RM-single} = a_2 \frac{K_{DC} t^n}{2} \quad (66)$$

If the two transistors are under stress with equal probability, we have

$$\Delta \mu_{RM-Dual} = a_2 \cdot \alpha(s = 0.5) \cdot K_{DC} \cdot t^n \quad (67)$$

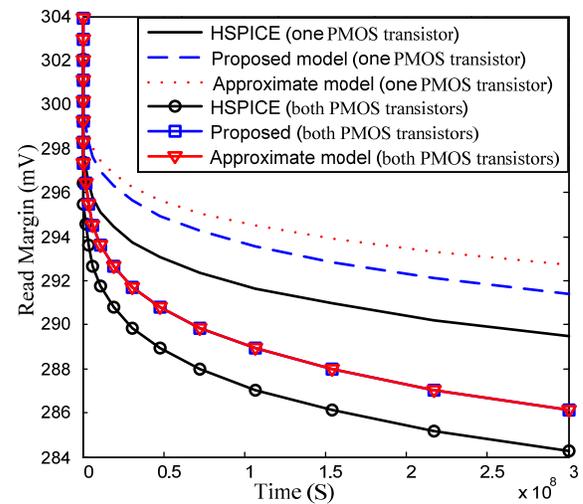
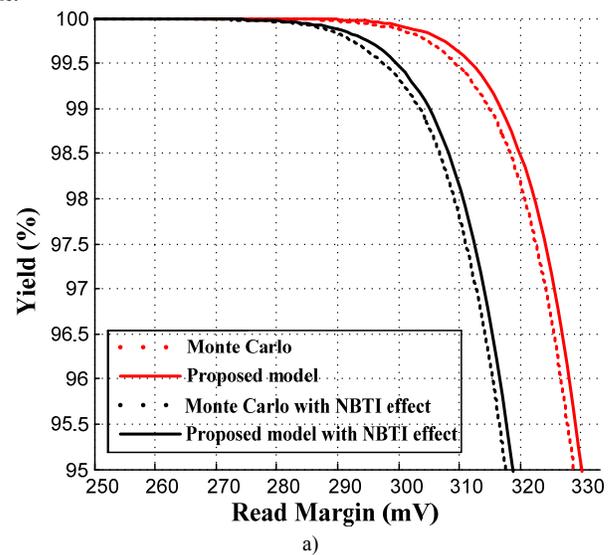


Fig. 7. Time variation of Read Margin mean obtained from HSPICE simulations, proposed model of (A2. 3), and approximate model of (64) for two cases with NBTI effect on one and both PMOS transistors in terms of time.



a)

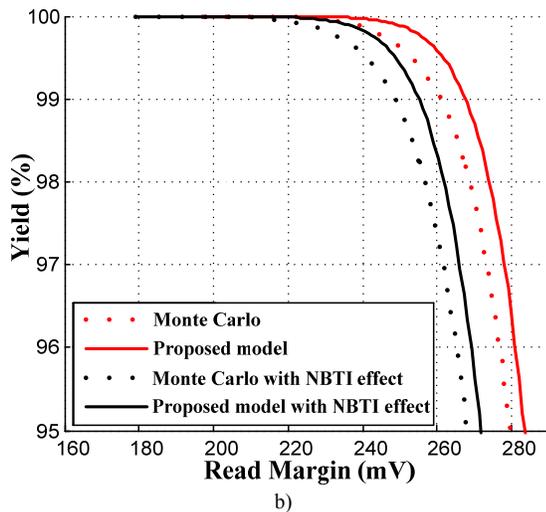


Fig. 8. Yield versus target read margin obtained from HSPICE simulations and proposed model for a) 45 nm and b) 32 nm technologies without considering NBTI effect and with considering NBTI effect after 10 years.

When the two transistors are under equal stress, $\alpha(s = 0.5) = 0.796$ [12]. In Fig. 7, we have plotted the time variations of the Read Margin average obtained from the HSPICE simulations, obtained Read Margin PDF model, and the approximate expression of (64). The variation of the threshold voltages was assumed to be $3\sigma_i = 0.2V_{th,i}$ where $V_{th,i}$ is the nominal value of the threshold voltage for the transistor I [19]. The results indicate that the time variation of the Read Margin mean is more compared to the case where one transistor is under stress.

In Fig. 8, we have plotted yield versus target Read Margin obtained from the model and HSPICE simulations for 45 nm and 32 nm technologies. Also, the ratio of the working SRAM cells to the total cells after ten years (still denoted by yield) is given in the figure. Again, the comparison reveals a very good accuracy for the model. For the case of HSPICE simulations, we considered the shift in the threshold voltage based on the expression given by (65). As the results reveal, for a given Read Margin, the NBTI effect gives rise to the reduction in the number of working cells, and hence, this should be taken into account by the designer.

V. CONCLUSION

In this work, we obtained an analytical expression for the distribution function (PDF) of the Read Margin of the SRAM cell. Also, using the model, we studied the effect of the NBTI on the PDF of the read margin to determine the number of cells which have acceptable Read Margin value as a function of time. First, we obtained analytical expressions for the read and trip voltages and then used these expressions to model the Read Margin analytically. The model then was used to obtain the PDF of Read Margin. Since the PDF was an analytical function of the threshold voltages, we managed to include the impact of NBTI. The model accuracy was verified by comparing its predictions with those of HSPICE Monte Carlo analysis for 45 nm and 32 nm technologies. The comparison showed a very good accuracy for the model. We also compared the two cases of both PMOS transistors were equally affected and only one transistor was affected by NBTI.

References

- [1] "International technology roadmap for semiconductors," 2009. [Online]. Available: <http://www.itrs.net/Links/2009ITRS/Home2009.htm>
- [2] K. Agarwal and S. Nassif, "The impact of random device variation on SRAM cell stability in Sub-90-nm CMOS technologies," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 16, no. 1, pp. 86–97, Jan. 2008.
- [3] C.-H. Lin, M. V. Dunga, D. D. Lu, A. M. Niknejad, and C. Hu, "Performance-Aware Corner Model for Design for Manufacturing," *IEEE Tran. on Electron Devices*, vol. 56, no. 4, pp. 595–600, 2009.
- [4] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, "Modeling of failure probability and statistical design of SRAM array for yield enhancement in nanoscaled CMOS," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 24, no. 12, pp. 1859–1880, Dec. 2005.
- [5] Y. Li, C.-H. Hwang, and T.-Y. Li, M.-H. Han, "Process-Variation Effect, Metal-Gate Work-Function Fluctuation, and Random-Dopant Fluctuation in Emerging CMOS Technologies," *IEEE Trans. Electron Devices*, vol. 57, no. 2, pp. 437–447, Dec 2009.
- [6] Y. Li, C.-H. Hwang, and T.-Y. Li, "Random-Dopant-Induced Variability in Nano-CMOS Devices and Digital Circuits," *IEEE Trans. Electron Devices*, vol. 56, no. 8, pp. 1588 – 1597,
- [7] Y. Li, C.-H. Hwang, and T.-Y. Li, "Discrete-dopant-induced timing fluctuation and suppression in nanoscale CMOS circuit," *IEEE Trans. Circuits Syst. II: Express Briefs*, vol. 56, no. 5, pp. 379–383, May 2009.
- [8] M.A. Alam, K. Roy, and C. Augustine, "Reliability- and process-variation aware design of integrated circuits," *Microelectronics Reliability*, vol. 48, no. 8, pp. 1114–1122, 2008.
- [9] A.E. Islam, H. Kuflluoglu, D. Varghese, S. Mahapatra, and M.A. Alam, "Recent Issues in Negative-Bias Temperature Instability: Initial Degradation, Field Dependence of Interface Trap Generation, Hole Trapping Effects, and Relaxation," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2143 – 2154, Aug. 2007.
- [10] M. A. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectron. Reliab.*, vol. 45, no. 1, pp. 71–81, Jan. 2005.
- [11] M. A. Alam, H. Kuflluoglu, D. Varghese, and S. Mahapatra, "A comprehensive model for PMOS NBTI degradation: Recent progress," *Microelectron. Reliab.*, vol. 47, no. 6, pp. 853–862, Jun. 2007.
- [12] K. Kang, H. Kuflluoglu, K. Roy, and M. A. Alam, "Impact of Negative-Bias Temperature Instability in Nanoscale SRAM Array: Modeling and Analysis," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 26, no. 10, pp. 1770–1781, Oct. 2007.
- [13] S. Srivastava, and J. Roychowdhury, "Rapid Estimation of the Probability of SRAM Failure due to MOS Threshold Variations," in *Proc. Custom Integrated Circuits*, 2007, pp. 229 – 232.
- [14] R.A. Fonseca, L. Dilillo, A. Bosio, and P. Girard, "Detecting NBTI Induced Failures in SRAM Core-Cells," in *Proc. VLSI Test Symposium (VTS)*, 2010 28th, pp. 75 – 80.
- [15] H. Singh and H. Mahmoodi, "Analysis of SRAM Reliability under Combined Effect of NBTI, Process and Temperature Variations in Nano-Scale CMOS," in 5th *Proc. Future Information Technology*, 2010, pp. 1-4.
- [16] B. Afzal, B. Ebrahimi, and A. Afzali-Kusha, "An accurate analytical I - V model for sub-90-nm MOSFETs and its application to read static noise margin modeling," *Journal of Zhejiang University-SCIENCE C (Computers & Electronics)*, vol. 13, no. 1, pp. 58–70, Jan. 2012.
- [17] I. Carlson, "Design and evaluation of high density 5T SRAM cache for advanced microprocessors," M.S. thesis, Dept. Elec. Eng., Linkoping Univ., Linkoping, Sweden, 2004.
- [18] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, 6th ed. Oxford, U.K.: Oxford Univ. Press, 1999, vol. 2A.
- [19] L. Cheng, P. Gupta, and L. He, "On Confidence in Characterization and Application of Variation Models," in *Proc. ASP-DAC*, 2010, pp. 751 – 756.

APPENDIX 1

To obtain P_1 , ($FL = f$ and $FR > f$), we substitute t defined by (52) into (41). Also, noting that $FL = f$, one may use (53) and vary N_L , M_L , and P_L from $-\infty$ and $+\infty$. Therefore,

$$P_1(f).df = \int_{P_L=-\infty}^{P_L=\infty} \int_{N_L=-\infty}^{N_L=\infty} \int_{M_L=-\infty}^{M_L=\infty} P_{P_L}(P_L)P_{M_L}(M_L).P_{N_L}(N_L).P_t(t=t_0).P(F_R > f).dP_L.dM_L.dN_L.dt \quad (A1. 1)$$

Also, since $dt = df$ (see (53)), we may write

$$P_1(f).df = \int_{P_L=-\infty}^{P_L=\infty} \int_{N_L=-\infty}^{N_L=\infty} \int_{M_L=-\infty}^{M_L=\infty} P_{P_L}(P_L)P_{M_L}(M_L).P_{N_L}(N_L).P_t(t=t_0).P(F_R > f).dP_L.dM_L.dN_L.df \quad (A1. 2)$$

The only unknown is $P(F_R > f)$ which may be obtained as is explained here. Using (52) and (53), we obtain M_R in terms of N_R, P_L, M_L, N_L , and f . This relation for M_R is then plugged into (42). Note that, as mentioned above, N_L, P_L , and M_L are assumed to vary from $-\infty$ and $+\infty$ in the integral. The two remaining variables are N_R and P_R . One of these two parameters (P_R in our case) also may vary from $-\infty$ and $+\infty$. Therefore, N_R is selected such that the condition $FR > f$ is satisfied. Thus, for the condition of $FR > f$ corresponds to $N_R < N_{R0}$. Since N_R is a normal standard variable, the probability for this condition is obtained from

$$P_{N_R}(N_R < N_{R0}) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{N_{R0}}{\sqrt{2}} \right) \right) \quad (A1. 3)$$

Therefore,

$$P(F_R > f) = \int_{P_R=-\infty}^{P_R=\infty} P_{P_R}(P_R) \frac{(1 + \operatorname{erf}(\frac{N_{R0}}{\sqrt{2}}))}{2} dP_R \quad (A1. 4)$$

By combining (A1. 4) and (A1. 2), we obtain (51).

APPENDIX 2

To obtain the average of the Read Margin, first, we should obtain the average of $P_f(f)$, which is denoted by μ_{P_f} , using

$$\mu_{P_f} = \int_{f=-\infty}^{f=\infty} f.P_f(f).df \quad (A2. 1)$$

This integral may be calculated using (48), (60), and (61) with the help of

$$\int_{x=-\infty}^{x=\infty} x.e^{-x^2}.\operatorname{erf}(a.x+b).dx = \frac{a}{e^{a^2+1}\sqrt{a^2+1}} \quad (A2. 2)$$

Using the result of (A2. 1) as well as (38) and (39), the average of the Read Margin (μ_{RM}) may be obtained as

$$\begin{aligned} \mu_{RM} = & \mu_{RMC} + \frac{a_2.V_{nbti}}{2} + \frac{a_2.V_{nbti}}{2\sqrt{\pi}.\sqrt{k_1^2+k_2^2+k_3^2+5}} \\ & + \frac{(k_1^2+k_2^2+k_3^2+5+2k_3-6k_2).k_{ax,n}.\sigma_n}{\sqrt{\pi}.(2k_2+k_3).z.e^{2.(\frac{a_2.V_{nbti}}{k_{ax,n}.\sigma_n.z.(2k_2+k_3)})^2}} \end{aligned} \quad (A2. 3)$$

where

$$z = \sqrt{1 + \frac{(k_1^2+k_2^2+k_3^2+5+2k_3-6k_2)^2}{(2k_2+k_3)^2.(k_1^2+k_2^2+k_3^2+5)}} \quad (A2. 4)$$

The simulation results show that the last two terms in (A2. 3) may be ignored (especially, when σ_i 's are large). Note that since V_{nbti} and V_{nbic} vary with time, we obtain the approximate change of μ_{RM} as is given by (64).