PAPER

Temperature-Aware NBTI Modeling Techniques in Digital Circuits*

Hong LUO[†a)], ***Nonmember***, **Yu WANG**[†b)], ***Member***, **Rong LUO**[†], ***Nonmember***, **Huazhong YANG**[†], ***Member***, ***and*** **Yuan XIE**[††], ***Nonmember***

SUMMARY Negative bias temperature instability (NBTI) has become a critical reliability phenomena in advanced CMOS technology. In this paper, we propose an analytical temperature-aware dynamic NBTI model, which can be used in two circuit operation cases: executing tasks with different temperatures, and switching between active and standby mode. A PMOS V_{th} degradation model and a digital circuits' temporal performance degradation estimation method are developed based on our NBTI model. The simulation results show that: 1) the execution of a low temperature task can decrease ΔV_{th} due to NBTI by 24.5%; 2) switching to standby mode can decrease ΔV_{th} by 52.3%; 3) for ISCAS85 benchmark circuits, the delay degradation can decrease significantly if the circuit execute low temperature task or switch to standby mode; 4) we have also observed the execution time's ratio of different tasks and the ratio of active to standby time both have a considerable impact on NBTI effect.

key words: *negative bias temperature instability (NBTI), temperature, reliability*

1. Introduction

As semiconductor manufacturing migrates to more advanced technology nodes, accelerated aging effect for nanoscale devices poses as a key challenge for designers to find countermeasures that effectively mitigate the degradation and prolong system's lifetime. Negative bias temperature instability (NBTI) is emerging as one of the major reliability concerns [1]. NBTI significantly shifts the threshold voltage V_{th} on the order of 20–50 mV for devices operating at 1.2 V or below [2]. Previous research showed that V_{th} shifts due to NBTI is expressed as a power-law time dependence [1], [3], and this aging-induced parameter variation has a negative impact on circuit performance. At the 65 nm technology node, after the aging time of 10^6s arising from NBTI, the variability of circuit delay can reach up to 15% [4].

NBTI phenomena occur when the PMOS transistor is negative biased. Under the negative bias, the Si-H bonds (which are formed during the manufacture procedure) capture the holes tunneling from the inversion layer, which causes the dissociation of the Si-H bonds and yields interface traps. As a result of the interface traps, the threshold voltage of PMOS transistor is shifted, carrier mobility and drain current are reduced [5], and then the performance degradation occurs [6]–[8]. The NBTI phenomena can be classified as *static NBTI* and *dynamic NBTI*. Static NBTI is under the DC stress condition, and the detailed physical mechanism was described in [9]. The impact of electric and environment parameters (such as electric field across the oxide and temperature) on the interface trap generation was studied in [10], [11]. Dynamic NBTI under AC stress condition leads to a less severe parameter's shift over long time because of the recovery phenomenon [6], [11]–[13].

Prior works in analyzing and mitigating performance degradation due to NBTI can be classified into two categories:

- **NBTI modeling and analysis**. Many analytical models for circuit performance degradation due to NBTI have been proposed recently [7], [8], [14]. The impact of NBTI on the worst case performance degradation of digital circuits was analyzed in [15]. An analytical model for multi-cycle dynamic NBTI was proposed in [16], where a recursion process was used to evaluate the NBTI effect. A predictive NBTI model is proposed in [17], [18], the effect of various process and design parameters were described.
- **Circuit optimization to mitigate NBTI effects.** The circuit delay optimization and reliability improvement based on the previous analytical circuit degradation models were also studied. In [19], a gate-sizing optimization technique was used to guarantee the delay of the circuit. An NBTI-aware design for SRAM read stability was proposed in [8]. Various circuit design techniques, such as V_{dd}/V_t tuning, were investigated [17]. A technology mapping technique that incorporates the NBTI effects was presented in [20] based on the model in [16].

Analytical NBTI models were useful in circuit analysis and optimization for NBTI effects. In previous analytical NBTI models [16]–[18], the circuit temperature was considered to be constant (about 400 K) during the circuit operation. However, the temperature of the circuit is variable during the circuit operation, and an analytical model considering arbitrary temperature variation for static NBTI was proposed in [21], but the relaxation phase of NBTI effect

Manuscript received November 7, 2008.

[†]The authors are with TNList and Department of Electronic Engineering, Tsinghua University, Beijing, 100084, P.R. China.

[††]The author is with the Department of CSE, Pennsylvania State University, University Park, PA, USA.

*This work was supported by grants from 863 program of China (No. 2009AA01Z130), and NSFC (No. 60506010, No. 90707002, No. 60870001) and TNList Cross-discipline Foundation. Yuan Xie's work was supported in part by NSF 0454123, NSF CAREER 0643902 and MARCO/DRAPA-GSRC.

a) E-mail: luohong99@mails.thu.edu.cn

b) E-mail: yu-wang@mail.tsinghua.edu.cn

DOI: 10.1587/transele.E92.C.875

Copyright © 2009 The Institute of Electronics, Information and Communication Engineers

was ignored in the derivation, and this model lead to complex transformation in the calculation. Therefore, we consider two simpler but practical cases in our paper (based on our previous simple NBTI modeling work [22], [23]): a) the circuit operates all the time, but executes different tasks with different load, then the temperature varies between the tasks; b) the circuit operates in two modes: active and standby mode, thus the circuit temperature varies periodically between a high temperature and a low temperature. In both cases, the relaxation phase are included in our derivation, and we propose an analytical threshold voltage degradation model for dynamic NBTI effect considering the temperature variation on account of these two cases.

Our contribution in this paper distinguishes itself in the following aspects:

- We propose analytical models to estimate PMOS NBTI effect under AC stress condition considering both constant temperature and variable temperature. Furthermore, our temperature-aware model can be used in two circuit operation cases: a) the temperature varies due to tasks with different workloads; b) the temperature changes between two operation modes: active and standby mode. The threshold voltage degradation can be estimated in both cases.
- Based on our novel temperature-aware NBTI model, we study the impact of NBTI on the temporal performance degradation in combinational circuits. In one case that different tasks lead to diffferent temperature, the relative execution time of these tasks will have impact on the NBTI effect. In another case that the circuit switches between active and standby mode, the time ratio of active to standby mode can affect the performance degradation significantly.

The rest of the paper is organized as follows. In Sect. 2, we first review the physical mechanism of NBTI and previous NBTI models. Then, in Sect. 3, we derive the NBTI-induced interface trap generation model with constant temperature, while in Sect. 4, this model is extended to the case where the temperature varies between a high temperature and a low temperature. The V_{th} degradation model, and the circuit delay estimation method is proposed in Sect. 5. The simulation results of a single PMOS transistor and the ISCAS85 benchmark circuits are shown and analyzed in Sect. 6. Finally, Sect. 7 concludes the paper.

2. Review of NBTI

2.1 Physical Mechanism of NBTI

It is commonly admitted that the interface traps are generated under negative gate bias due to mismatches at the $Si\text{-}SiO_2$ interface, in particular at elevated temperature, and lead to degradation of device performance. During oxidation of Si, some of the Si atoms bond with hydrogen, leading to the formation of weak Si-H bonds. When a PMOS transistor is negative biased, the inversion layer holes tunnel

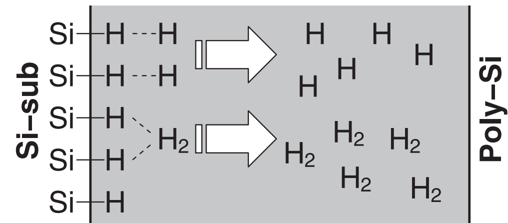


Fig. 1 Mechanism of NBTI.

to and are captured by the Si-H bonds at the interface, which weakens the bond and the subsequent thermally assisted dissociation of a Si-H pair yields one donor-like interface trap (i.e., silicon dangling bond) and one H atom that can diffuse in the oxide.

$$\mathrm{SiH} + h^+ \leftrightarrow \mathrm{Si}^+ + \mathrm{H} \tag{1}$$

The dimerization reaction involves the reaction of H atoms to create molecular hydrogen, which is described by Eq. (2).

$$\mathrm{H} + \mathrm{H} \leftrightarrow \mathrm{H}_2 \tag{2}$$

Then, molecular hydrogen can also diffuse in the oxide. Therefore, the diffusing species can be both atomic hydrogen and molecular hydrogen, as shown in Fig. 1.

Recent researches have shown that molecular hydrogen are the dominant diffusing species [24]. So in this paper, the diffusing species are assumed to be molecular hydrogen H_2. Combining Eq. (1) and (2), the final reaction can be [25]

$$\mathrm{SiH} + h^+ \leftrightarrow \mathrm{Si}^+ + \frac{1}{2}\mathrm{H}_2 \tag{3}$$

Therefore, the rate equation for the quantity of Si^+ is

$$\frac{d[\mathrm{Si}^+]}{dt} = k_{\mathrm{f}}^0[\mathrm{SiH}][h^+] - k_{\mathrm{r}}[\mathrm{Si}^+][\mathrm{H}_2]^{\frac{1}{2}} \tag{4}$$

where k_{f}^0 and k_{r} are forward and reverse reaction rate constants in Eq. (3).

Then, we denote the quantity of Si^+ as N_{it}, the initial quantity of Si-H bonds as N_0 and the quantity of H_2 molecules as $N_{\mathrm{i,H_2}}$. Because the quantity of holes is only dependent on $V_{\mathrm{gs}} - V_{\mathrm{th}}$ ($[h^+] = C_{\mathrm{ox}}(V_{\mathrm{gs}} - V_{\mathrm{th}})$), we combine the forward reaction rate constant k_{f}^0 and $[h^+]$ as one parameter k_{f}, which is dissociation rate of Si-H bonds. So Eq. (4) can be rewritten as Eq. (5),

$$\frac{dN_{\mathrm{it}}}{dt} = k_{\mathrm{f}}(N_0 - N_{\mathrm{it}}) - k_{\mathrm{r}} N_{\mathrm{it}} N_{\mathrm{i,H_2}}^{\frac{1}{2}} \tag{5}$$

The hydrogen molecules H_2 can diffuse into the oxide, which satisfies the following

$$\frac{\partial N_{\mathrm{H_2}}}{\partial t} = D\frac{\partial^2 N_{\mathrm{H_2}}}{\partial x^2} \tag{6}$$

$$\frac{dN_{\mathrm{it}}}{dt} = -2 \cdot D \left.\frac{\partial N_{\mathrm{H_2}}}{\partial x}\right|_{x=0} \tag{7}$$

The above three equations Eq. (5) to (7) describe the interface trap generation, which is named as reaction-diffusion

(R-D) model. The generated interface traps can both enlarge the threshold voltage and reduce the mobility. The mobility degradation due to Coulomb scattering can be effected as a shift in the threshold voltage [26]. Therefore, the shift in the threshold voltage ΔV_{th} of the PMOS transistor is proportional to the interface trap generation due to NBTI, which can be expressed [19] as

$$\Delta V_{\mathrm{th}} = (1+m_{\mu})\frac{qN_{\mathrm{it}}(t)}{C_{\mathrm{ox}}} \tag{8}$$

where m_{μ} represents equivalent V_{th} shifts due to mobility degradation described in [27]; q is the electronic charge; C_{ox} is the gate oxide capacitance; and $N_{\mathrm{it}}(t)$ is the interface trap generation, which is the most important factor in evaluating performance degradation due to NBTI.

Interface traps are electrically active physical defects with their energy distributed between the valence and the conduction band in the band diagram. The interface traps shift the threshold voltage of MOSFETs, diminish carrier mobility, and add parasitic capacitances. All these effects reduce the drain current of the devices and thus the temporal performance. An increase of the interface traps not only leads to reduced temporal performance but also may cause reliability issues and potential device failure [5], [6], [8].

2.2 Previous Analytical Models

Kumar et al. proposed an analytical NBTI model [16] to handle multi-cycle AC stress condition; and assume the PMOS transistor is under AC stress with duty cycle of c and period of τ, the interface trap generation after $n+1$ cycles of AC stress can be evaluated by a recursion formula below,

$$N_{\mathrm{it}}[(n+1)\tau] = \frac{\beta N_{\mathrm{it}}(n\tau)}{1+\beta} + \frac{N_{\mathrm{it}}^{0}}{1+\beta}\left[c+\left(\frac{N_{\mathrm{it}}(n\tau)}{N_{\mathrm{it}}^{0}}\right)^{6}\right]^{1/6} \tag{9}$$

and the initial condition is

$$N_{\mathrm{it}}(\tau) = \frac{c^{1/6}}{1+\beta}N_{\mathrm{it}}^{0} \tag{10}$$

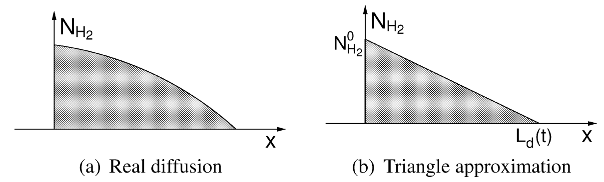
where $N_{\mathrm{it}}^{0} = A\tau^{1/6}$, and $\beta = \sqrt{\frac{1-c}{2}}$.

Predictive NBTI models are proposed in [17] and [18], the effect of various process and design parameters were described. The diffusing species were assumed to be hydrogen atoms in [17], and the default diffusing species in [18] were assumed to be hydrogen molecules. The threshold voltage degradation ΔV_{th} can be calculated as [18]

$$\Delta V_{\mathrm{th}} = \left(\frac{qT_{\mathrm{ox}}}{\varepsilon_{\mathrm{ox}}}\right)\sqrt[3]{K^{2}C_{\mathrm{ox}}(V_{\mathrm{gs}}-V_{\mathrm{th}})\exp\left(\frac{2E_{\mathrm{ox}}}{E_{0}}\right)}\cdot(Ct)^{\frac{1}{6}} \tag{11}$$

3. NBTI Model without Temperature Variation

In this section, an analytical dynamic NBTI model is derived based on the R-D model Eq. (5) to (7). This model will be extended to a temperature-aware model in the next section. As a basic assumption, the diffusing species are assumed to be molecular hydrogen.



(a) Real diffusion (b) Triangle approximation

Fig. 2 Diffusion profile for hydrogen molecules.

3.1 Approximation for Diffusion Profile

The concentration of hydrogen molecules is highest at the interface, where Si-H bonds are broken and interface traps are generated, and gradually decreases as these molecules diffuse into the oxide illustrated in Fig. 2(a). The diffusion profile is described by the diffusion Eq. (6), and can be approximated as a triangle shown in Fig. 2(b), where the concentration is $N_{\mathrm{i,H_2}}(t)$ at the interface and is zero at a point L_{d} known as *diffusion front* or *diffusion length*. Because one hydrogen molecule corresponds to two interface traps, the interface traps can be expressed as

$$N_{\mathrm{it}}(t) = \frac{1}{2}\cdot 2N_{\mathrm{i,H_2}}(t)\cdot L_{\mathrm{d}}(t) = N_{\mathrm{i,H_2}}(t)\cdot L_{\mathrm{d}}(t) \tag{12}$$

3.2 First Stress Phase

The initial density of Si-H bonds is much larger than the number of interface traps generated due to NBTI, so that in the right side of reaction Eq. (5), $N_0 - N_{\mathrm{it}} \approx N_0$. And the reaction process will not deviate far from equilibrium, which is defined as quasi-equilibrium, so that $\frac{dN_{\mathrm{it}}}{dt} \approx 0$. Therefore, Eq. (5) can be derived as

$$\frac{k_{\mathrm{f}}N_0}{k_{\mathrm{r}}} = N_{\mathrm{it}}(t)\sqrt{N_{\mathrm{i,H_2}}(t)} \tag{13}$$

Because of the triangle approximation of diffusion profile, $-\frac{\partial N_{\mathrm{H_2}}(x,t)}{\partial x}\Big|_{x=0}$ can be approximated as $N_{\mathrm{i,H_2}}(t)/L_{\mathrm{d}}(t)$, so substituting it into Eq. (7) and with Eq. (12), we obtain that

$$\frac{dN_{\mathrm{i,H_2}}(t)}{dt}L_d(t) + N_{\mathrm{i,H_2}}(t)\frac{dL_d(t)}{dt} = 2\cdot D_{\mathrm{H_2}}\frac{N_{\mathrm{i,H_2}}(t)}{L_d(t)} \tag{14}$$

From diffusion equation (6), we can derive that

$$\frac{dN_{\mathrm{i,H_2}}(t)}{dt} = \frac{\partial N_{\mathrm{H_2}}(0,t)}{\partial t} = D_{\mathrm{H_2}}(T)\frac{\partial^2 N_{\mathrm{H_2}}(x,t)}{\partial x^2}\Big|_{x=0} \tag{15}$$

Because of the triangle approximation, the right side of the above equation $\frac{\partial^2 N_{\mathrm{H_2}}(x,t)}{\partial x^2}\Big|_{x=0} \approx 0$, then $\frac{dN_{\mathrm{i,H_2}}(t)}{dt} \approx 0$, so that Eq. (14) can be simplified as

$$N_{\mathrm{i,H_2}}(t)\frac{dL_d(t)}{dt} = 2\cdot D_{\mathrm{H_2}}\frac{N_{\mathrm{i,H_2}}(t)}{L_d(t)}$$
$$L_{\mathrm{d}}(t)dL_{\mathrm{d}}(t) = 2\cdot D_{\mathrm{H_2}}dt \tag{16}$$

Integrating the terms of two sides in above equation for time t, we obtain that

$$L_{\mathrm{d}}^2\Big|_{t_0}^{t} = 4\cdot D_{\mathrm{H_2}}(t-t_0) \tag{17}$$

For the first stress phase, $t_0 = 0$, so $L_{\mathrm{d}}(t) = \sqrt{4D_{\mathrm{H_2}}t}$. Therefore, the diffusion length of the triangle approximated profile follows the relation of $\sqrt{pD_{\mathrm{H_2}}t}$, and in fact, according to the results from [25], $p = 9.61$ is a more accurate value.

Otherwise, from Eq. (12) and quasi-equilibrium Eq. (13), we get that

$$\left(\frac{k_{\mathrm{f}}N_0}{k_{\mathrm{r}}}\right)^2 = [N_{\mathrm{it}}(t)]^2\cdot\frac{N_{\mathrm{it}}(t)}{L_d(t)} \tag{18}$$

so the generated interface traps can be approximated as

$$N_{\mathrm{it}}(t) = \left(\frac{k_{\mathrm{f}}N_0}{k_{\mathrm{r}}}\right)^{\frac{2}{3}}\cdot[L_d(t)]^{\frac{1}{3}} = \left(\frac{k_{\mathrm{f}}N_0}{k_{\mathrm{r}}}\right)^{\frac{2}{3}}\cdot(pDt)^{\frac{1}{6}} \tag{19}$$

3.3 First Relaxation Phase

We assume that PMOS transistor is being stressed for t_0, after the voltage is removed, the transistor will be in relaxation phase. In the relaxation phase, the generated hydrogen molecules recombine with the interface traps, and become Si-H bonds. This procedure is described by reverse direction of Eq. (3). The transistor has been in relaxation phase for Δt, the number of annealed traps is denoted as $N_{\mathrm{it}}^*(\Delta t)$ [10], then at time $t_0+\Delta t$, the generated interface trap can be described by

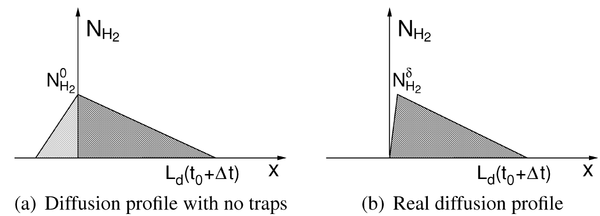
$$N_{\mathrm{it}}(t_0+\Delta t) = N_{\mathrm{it}}(t_0) - N_{\mathrm{it}}^*(\Delta t) \tag{20}$$

In the relaxation phase, because no new traps and hydrogen molecules are generated, the diffusion is two-sided: on one side, the hydrogen molecules diffuse into gate continuously, on the other side, they diffuse back into interface and recombine with the traps. Assuming that no trap exist at the interface, the hydrogen molecules will diffuse into the silicon as illustrated in Fig. 3(a). The number of the annealed interface traps due to backward diffusion is shown as the left triangle area, and can be described as [10]

$$N_{\mathrm{it}}^*(\Delta t) = N_{\mathrm{i,H_2}}(t_0+\Delta t)\sqrt{\xi\cdot pD_{\mathrm{H_2}}\Delta t} \tag{21}$$

where ξ is the two-sided diffusion factor, and theoretically $\xi = 0.5$.

In fact, the concentration of hydrogen molecules is almost zero at the interface. We can still use quasi-equilibrium condition to derive this conclusion. When the transistor is in relaxation, k_{f} in the reaction Eq. (5) is zero, and $\frac{dN_{\mathrm{it}}}{dt}\approx 0$, so



(a) Diffusion profile with no traps (b) Real diffusion profile

Fig. 3 Diffusion profile in relaxation phase.

we have

$$0 \approx -k_{\mathrm{r}}N_{\mathrm{it}}(t_0+\Delta t)\sqrt{N_{\mathrm{i,H_2}}(t_0+\Delta t)} \tag{22}$$

The interface traps $N_{\mathrm{it}}(t_0+\Delta t)$ is much larger than zero, so $N_{\mathrm{i,H_2}}\approx 0$. Therefore, the diffusion profile is shown in Fig. 3(b), and the maximum concentration of hydrogen molecules is at the position $x = \delta$, which is very close to the interface. Since δ is very small ($\delta \approx 0$), the concentration of hydrogen molecules $N_{\mathrm{i,H_2}}(t_0+\Delta t)$ in Fig. 3(a) can be almost the same as $N_{\mathrm{H_2}}^{\delta}(t_0+\Delta t)$. Therefore, we can still use the triangle approximation approach in Section 3.1, the total number of interface traps is given by

$$N_{\mathrm{it}}(t_0+\Delta t) = N_{\mathrm{i,H_2}}(t_0+\Delta t)\sqrt{pD_{\mathrm{H_2}}(t_0+\Delta t)} \tag{23}$$

From (21) and (23), we have

$$N_{\mathrm{it}}^*(\Delta t) = N_{\mathrm{it}}(t_0+\Delta t)\sqrt{\frac{\xi\Delta t}{t_0+\Delta t}} \tag{24}$$

Substituting for $N_{\mathrm{it}}^*(\Delta t)$ in Eq. (20), we have

$$N_{\mathrm{it}}(t_0+\Delta t) = N_{\mathrm{it}}(t_0) - N_{\mathrm{it}}(t_0+\Delta t)\sqrt{\frac{\xi\Delta t}{t_0+\Delta t}} \tag{25}$$

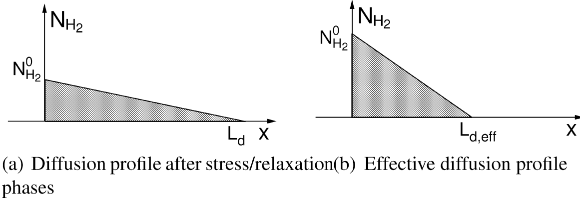
and re-arranging some terms, we get the interface traps after this relaxation phase as the follow

$$N_{\mathrm{it}}(t_0+\Delta t) = \frac{N_{\mathrm{it}}(t_0)}{1+\sqrt{\frac{\xi\Delta t}{t_0+\Delta t}}} \tag{26}$$

3.4 Subsequent Stress/Relaxation Phases

In this section, we use mathematical induction to derive the density of interface traps. We assume that the transistor is stressed by a periodic voltage waveform with the period τ, and the duty cycle of stress phase is c. The interface traps generated after the m stress phases are denoted as $N_{\mathrm{its},m}$, and $N_{\mathrm{itr},m}$ represents the number of traps after m relaxation phases.

After some stress and relaxation phases, the number of the generated interface traps becomes a new boundary condition as illustrated in Fig. 4(a). The relaxation phases reduce the interface traps, so if there are only stress phases, the same interface traps can be generated the effective diffusion profile in Fig. 4(b). Therefore, by this equivalent method, we can derive the generated interface traps in any stress



(a) Diffusion profile after stress/relaxation phases (b) Effective diffusion profile

Fig. 4 Equivalent method for diffusion profile.

phase.

In order to evaluate the interface traps after the $m+1$ stress phase, we substitute $L_{\mathrm{d,eff}}$ for L_{d}, and Eq. (17) becomes

$$L^2_{\mathrm{d,eff}}\Big|_{m\tau}^{m\tau+c\tau} = p\cdot D_{\mathrm{H_2}}(c\tau) \tag{27}$$

We should notice that $L_{\mathrm{d,eff}}(m\tau+c\tau)$ corresponds to the interface traps after $m+1$ stress phases $N_{\mathrm{its},m+1}$, while $L_{\mathrm{d,eff}}(m\tau)$ corresponds to $N_{\mathrm{itr},m}$. From Eq. (18) and (27), we get that

$$\frac{(N_{\mathrm{its},m+1})^6-(N_{\mathrm{itr},m})^6}{\left(\frac{k_{\mathrm{f}}N_0}{k_{\mathrm{r}}}\right)^4} = p\cdot D_{\mathrm{H_2}}(c\tau) \tag{28}$$

so that we can derive the number of interface trap generated after $m+1$ stress phases from the number after m relaxation phases,

$$N_{\mathrm{its},m+1} = \left[(N_{\mathrm{itr},m})^6+\left(\frac{k_{\mathrm{f}}N_0}{k_{\mathrm{r}}}\right)^4\cdot pD_{\mathrm{H_2}}(c\tau)\right]^{\frac{1}{6}} \tag{29}$$

We use the same derivation as in Section 3.3 to get the interface traps generated after m relaxation phases

$$N_{\mathrm{itr},m} = \frac{N_{\mathrm{its},m}}{1+\sqrt{\frac{\xi(1-c)\tau}{m\tau}}} \tag{30}$$

From the above two equations, we get

$$N_{\mathrm{its},m+1} = \left[\alpha_m^6\cdot(N_{\mathrm{its},m})^6+K(c\tau)\right]^{\frac{1}{6}} \tag{31}$$

where

$$\alpha_m = \frac{1}{1+\sqrt{\frac{\xi(1-c)\tau}{m\tau}}}, \qquad K = \left(\frac{k_{\mathrm{f}}N_0}{k_{\mathrm{r}}}\right)^4\cdot pD_{\mathrm{H_2}}$$

By mathematical induction, we have the number of interface traps after $m+1$ stress phases,

$$\begin{aligned}N_{\mathrm{its},m+1} &= [K(c\tau)]^{\frac{1}{6}}(1+\alpha_m^6+\alpha_m^6\alpha_{m-1}^6+\cdots\\&\quad+\alpha_m^6\alpha_{m-1}^6\cdots\alpha_2^6\alpha_1^6)^{\frac{1}{6}}\\&= [K(c\tau)]^{\frac{1}{6}}\left[1+\sum_{i=1}^{m}\left(\prod_{j=m-i+1}^{m}\alpha_j^6\right)\right]^{\frac{1}{6}}\end{aligned} \tag{32}$$

We should notice that $\alpha_1<\alpha_2<\cdots<\alpha_m<1$ and $\lim_{m\to\infty}\alpha_m = 1$, so that we can eliminate the style of the series sum and get a bound of the interface traps' number,

$$\begin{aligned}N_{\mathrm{its},m+1} &\le [K(c\tau)]^{\frac{1}{6}}(1+\alpha_m^6+(\alpha_m^6)^2+\cdots+(\alpha_m^6)^m)^{\frac{1}{6}}\\&\le [K(c\tau)]^{\frac{1}{6}}\left(\frac{1}{1-\alpha_m^6}\right)^{\frac{1}{6}} = \left[\frac{K(c\tau)}{1-\alpha_m^6}\right]^{\frac{1}{6}}\end{aligned} \tag{33}$$

When m is large enough, $\sqrt{\frac{\xi(1-c)\tau}{m\tau}} \ll 1$, then

$$1-\alpha_m^6 \approx 1-\left(1-\sqrt{\frac{\xi(1-c)\tau}{m\tau}}\right)^6 \approx 6\sqrt{\frac{\xi(1-c)\tau}{m\tau}} \tag{34}$$

so the number of interface traps after $m+1$ stress cycles

$$N_{\mathrm{its},m+1} \approx \left[\frac{\sqrt{m\tau}\cdot K(c\tau)}{6\sqrt{\xi(1-c)\tau}}\right]^{\frac{1}{6}} \tag{35}$$

If the discrete variable m is transformed to the continuous variable t, the above equation becomes

$$N_{\mathrm{it}}(t) = \left[\frac{Kc\tau\sqrt{t}}{6\sqrt{\xi(1-c)\tau}}\right]^{\frac{1}{6}} \tag{36}$$

From the predictive model Eq. (11) proposed in [18], the threshold voltage degradation ΔV_{th} can be calculated as

$$\Delta V_{\mathrm{th}} = K_{\mathrm{V}}\left[\frac{c\tau\sqrt{t}}{6\sqrt{\xi(1-c)\tau}}\right]^{\frac{1}{6}} \tag{37}$$

where

$$K_{\mathrm{V}} = \left(\frac{qT_{\mathrm{ox}}}{\varepsilon_{\mathrm{ox}}}\right)\sqrt[3]{K^2C_{\mathrm{ox}}(V_{\mathrm{gs}}-V_{\mathrm{th}})\sqrt{C}\exp\left(\frac{2E_{\mathrm{ox}}}{E_0}\right)} \tag{38}$$

All these parameters are defined in [18].

4. NBTI Model with Temperature Variation

In previous section, an dynamic model was derived under the assumption that the temperature remains constant. In fact, the circuit are not operated in a constant temperature. In this section, the NBTI model with temperature variation will be derived. We use a similar approach as in the above section to derive our model. First, we simplify the temperature varies as a periodical waveform, and the waveform of the temperature is described. Second, the recursive equations in stress and relaxation phases are derived. Finally, we can eliminate the recursion and get a closed-form equation.

We should notice that all the coefficients in Eq. (5) to (7) are depend on temperature.

$$D_{\mathrm{H_2}}(T) = D_0\exp\left(-\frac{E_{\mathrm{d}}}{k_{\mathrm{B}}T}\right), \tag{39}$$

$$k_f(T) = k_{f0}\exp\left(-\frac{E_f}{k_B T}4\right), \tag{40}$$

$$k_r(T) = k_{r0}\exp\left(-\frac{E_r}{k_B T}\right) \tag{41}$$

The reaction process can be approximated to be in quasi-equilibrium, because the diffusion is considered to be a much slower process than reaction. Then $dN_{it}(t)/dt \approx 0$ in (5), so that (5) is simplified to

$$\frac{N_{it}(t)\sqrt{N_{H_2}(0,t)}}{N_0 - N_{it}(t)} = \frac{k_f(T)}{k_r(T)} \approx k_{fr} \tag{42}$$

Because $E_f - E_r \approx 0$ [25], $k_f(T)/k_r(T)$ can be considered independent of temperature, so the right-hand side of Eq. (42) k_{fr} is k_{f0}/k_{r0}. Then the interface trap generation is considered to be only affected by temperature through the diffusion coefficient D_{H_2}.

If no relaxation is considered, NBTI-induced PMOS degradation under arbitrary dynamic temperature variation can be derived [21]. The interface trap generation under dynamic temperature variation $N_{it}[t, T(t)]$ can be converted to the equivalent NBTI effect under constant temperature T_{ref} [21]

$$N_{it}[t, T(t)] = N_{it}[\tau(t), T = T_{ref}] \tag{43}$$

where the transformation from time t into equivalent time τ is described in [21] and it is very complex for arbitrary temperature variation. Therefore, we will propose a much simpler but practical enough model in this paper.

4.1 Waveform of Temperature

The digital circuits often switch between active and standby modes, so that the temperature changes relevantly between a high temperature and a low temperature. Before our analysis, we assume that the device operates at a constant temperature T_H in active mode and at another temperature T_L in standby mode as shown in Fig. 5. For convenience, the temperature is considered to change periodically with the period τ_T. In a single period, the duration of high and low temperature is t_h and t_f respectively, and the rise and fall time is t_r and t_f.

The stress voltage on the device changes much faster than the temperature, i.e. $\tau_T \gg \tau$, so we can further assume that the temperature remain constant in a couple of stress and relaxation phases, as illustrated in Fig. 5. This assumption is also reasonable for transition phase if the temperature is simplify to be a step function.

4.2 Stress Phases

Firstly, we consider the impact of temperature variation in stress phases. According to the assumption in the above section, the temperature remains constant in a whole stress/relaxation cycle. So the recursive formula (29) can be still used

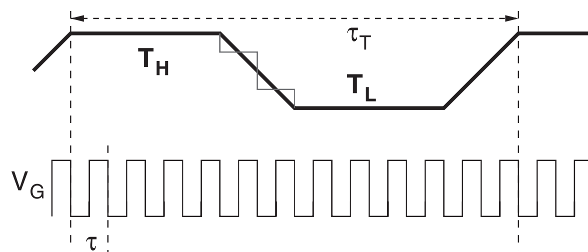


Fig. 5 Waveform of the temperature.

$$N_{its,m+1} = \left[(N_{itr,m})^6 + \left(\frac{k_f N_0}{k_r}\right)^4 \cdot p D_{H_2,m+1}(c\tau)\right]^{\frac{1}{6}} \tag{44}$$

where $D_{H_2,m+1}$ is the diffusion coefficient in the $m+1$ cycle, and as the notation in the above, k_f/k_r is independent on the temperature.

4.3 Relaxation Phases

Here, we will derive the recursive equation in the relaxation phase. The reaction process is in quasi-equilibrium, so that

$$-k_r[T(t)]N_{it}\sqrt{N_{i,H_2}} \approx 0 \tag{45}$$

Though the coefficient $k_r[T(t)]$ is time variant, the generation of interface traps is not relative with this coefficient. Therefore, we can still use the same approach in Sect. 3.3 to get the number of interface traps after m relaxation phases,

$$N_{itr,m} = N_{its,m} \Bigg/ \left(1 + \sqrt{\frac{\xi D_{H_2,m}(1-c)\tau}{\sum_{i=1}^{m}(D_{H_2,m}\cdot\tau)}}\right) \tag{46}$$

We should notice this formula has a similar form with Eq. (30), but all the time-variant coefficient D_{H_2} can not be eliminated.

In our assumption, the temperature varies periodically as shown in Fig. 5, then $D_{H_2}[T(t)]$ has a same style of waveform. Therefore, when m is large enough, there is

$$\sum_{i=1}^{m}(D_{H_2,m}\cdot\tau) \approx \frac{m\tau}{\tau_T}\int_{\tau_T} D_{H_2}dt = \tilde{D}_{H_2}\cdot(m\tau) \tag{47}$$

4.4 Full Model

Combining the formulas in Sect. 4.2 and 4.3, we can get complete recursive formula for N_{it}

$$N_{its,m+1} = \left[\alpha_m^6 \cdot (N_{its,m})^6 + K' D_{H_2,m+1}\cdot(c\tau)\right]^{\frac{1}{6}} \tag{48}$$

where

$$\alpha_m = 1 \Bigg/ \left(1 + \sqrt{\frac{\xi D_{H_2,m}\cdot(1-c)\tau}{\tilde{D}_{H_2}\cdot(m\tau)}}\right), \quad K' = p\left(\frac{k_f N_0}{k_r}\right)^4$$

Hence, the close form expression of the generated interface traps after $m+1$ stress phases can be derived as

$$\begin{aligned} N_{\mathrm{its},m+1} &= [K'(c\tau)]^{\frac{1}{6}}(D_{m+1}+\alpha_m^6 D_m+\alpha_m^6\alpha_{m-1}^6 D_{m-1}+\cdots \\ &\quad +\alpha_m^6\alpha_{m-1}^6\cdots\alpha_2^6\alpha_1^6 D_1)^{\frac{1}{6}} \\ &= [K'(c\tau)]^{\frac{1}{6}}\left[D_{m+1}+\sum_{i=1}^{m}\left(D_{m-i+1}\cdot\prod_{j=m-i+1}^{m}\alpha_j^6\right)\right]^{\frac{1}{6}} \\ &\le [K'(c\tau)]^{\frac{1}{6}}\left[\sum_{i=0}^{m}\left(D_{m-i+1}\cdot(\alpha_m^6)^i\right)\right]^{\frac{1}{6}} \end{aligned} \tag{49}$$

Here all the subscripts H_2 of the diffusion coefficients are omitted for brevity.

Because the variation of the temperature is much slower than the variation of the stress voltage applied on the gate of the transistor, in order to simplify the computation, we assume that $\tau_T = n\tau$, while $n \gg 1$, and $t_{\mathrm{h}} = n_{\mathrm{h}}\tau$, $t_{\mathrm{l}} = n_{\mathrm{l}}\tau$, $t_{\mathrm{r}} = n_{\mathrm{r}}\tau$ and $t_{\mathrm{f}} = n_{\mathrm{f}}\tau$, obviously, there is $n = n_{\mathrm{h}}+n_{\mathrm{l}}+n_{\mathrm{r}}+n_{\mathrm{f}}$. If we consider the number of the interface trap generation after a long time, i.e. $m \gg 1$, we can further assume that $m+1 = qn$ and the temperature starts at the high temperature. According to these assumptions above and the waveform of the temperature as shown in Fig. 5, there is $D_{m+qn} = D_m$, $D_{p+qn} = D_{\mathrm{h}}\,(1 \le p \le n_{\mathrm{h}})$, and $D_{p+qn} = D_{\mathrm{l}}\,(n_{\mathrm{h}}+n_{\mathrm{r}}+1 \le p \le n_{\mathrm{h}}+n_{\mathrm{r}}+n_{\mathrm{l}})$. Therefore, we can combine some terms in Eq. (49) and get the interface traps

$$\begin{aligned} N_{\mathrm{it},qn} &= (K'c\tau)^{\frac{1}{6}}\cdot\Big[(\alpha_m^6)^{(q-1)n}\beta_{\mathrm{h}}+(\alpha_m^6)^{(q-1)n}\beta_{\mathrm{r}}+\cdots \\ &\quad +(\alpha_m^6)^n\beta_{\mathrm{h}}+(\alpha_m^6)^n\beta_{\mathrm{r}}+\cdots+\beta_{\mathrm{l}}+\beta_{\mathrm{f}}\Big]^{\frac{1}{6}} \\ &= (K'c\tau)^{\frac{1}{6}}\cdot\left[\frac{1-\alpha_m^{6qn}}{1-\alpha_m^{6n}}(\beta_{\mathrm{h}}+\beta_{\mathrm{r}}+\beta_{\mathrm{l}}+\beta_{\mathrm{f}})\right]^{\frac{1}{6}} \\ &\le (K'c\tau)^{\frac{1}{6}}\cdot\left[\frac{1}{1-\alpha_m^{6n}}(\beta_{\mathrm{h}}+\beta_{\mathrm{r}}+\beta_{\mathrm{l}}+\beta_{\mathrm{f}})\right]^{\frac{1}{6}} \end{aligned} \tag{50}$$

where

$$\beta_{\mathrm{h}} = \left[(\alpha_m^6)^{n-1}+\cdots+(\alpha_m^6)^{n-n_{\mathrm{h}}}\right]D_{\mathrm{h}} \tag{51}$$

$$\beta_{\mathrm{l}} = \left[(\alpha_m^6)^{n-n_{\mathrm{h}}-n_{\mathrm{r}}-1}+\cdots+(\alpha_m^6)^{n-n_{\mathrm{h}}-n_{\mathrm{r}}-n_{\mathrm{l}}}\right]D_{\mathrm{l}} \tag{52}$$

$$\beta_{\mathrm{r}} = (\alpha_m^6)^{n-n_{\mathrm{h}}-1}D_{n_{\mathrm{h}}+1}+\cdots+(\alpha_m^6)^{n-n_{\mathrm{h}}-n_{\mathrm{r}}}D_{n_{\mathrm{h}}+n_{\mathrm{r}}} \tag{53}$$

$$\beta_{\mathrm{f}} = (\alpha_m^6)^{n-n_{\mathrm{h}}-n_{\mathrm{r}}-n_{\mathrm{l}}-1}D_{n_{\mathrm{h}}+n_{\mathrm{r}}+n_{\mathrm{l}}+1}+\cdots+(\alpha_m^6)^0 D_n \tag{54}$$

Because $m \gg 1$, the difference between $D_{\mathrm{H}_2,m}$ and $\tilde{D}_{\mathrm{H}_2}$ can be ignored. From Eq. (48), we define α_{qn} as α,

$$\alpha \approx \alpha_m = 1\Big/\left(1+\sqrt{\frac{\xi\cdot(1-c)\tau}{qn\tau}}\right) \tag{55}$$

Firstly, we consider that the temperature changes rapidly between the high and low temperature, which can significantly simplify the solution. In this case, β_{r} and β_{f} are both zero. Notice that $n = n_{\mathrm{h}}+n_{\mathrm{l}}$, then

$$\beta_{\mathrm{h}} = \alpha^{6n_{\mathrm{l}}}\frac{1-\alpha^{6n_{\mathrm{h}}}}{1-\alpha^6}D_{\mathrm{h}} \tag{56}$$

$$\beta_{\mathrm{l}} = \frac{1-\alpha^{6n_{\mathrm{l}}}}{1-\alpha^6}D_{\mathrm{l}} \tag{57}$$

and using Taylor approximation $\alpha^k \approx 1-k\sqrt{\frac{\xi\cdot(1-c)\tau}{qn\tau}}$, we have

$$\beta_{\mathrm{h}}+\beta_{\mathrm{l}} = \frac{6n_{\mathrm{h}}D_{\mathrm{h}}+6n_{\mathrm{l}}D_{\mathrm{l}}}{1-\alpha^6}\sqrt{\frac{\xi\cdot(1-c)\tau}{qn\tau}} \tag{58}$$

Hence

$$N_{\mathrm{it,qn}} = (K'c\tau)^{\frac{1}{6}}\cdot\left[\frac{n_{\mathrm{h}}D_{\mathrm{h}}+n_{\mathrm{l}}D_{\mathrm{l}}}{(1-\alpha^6)n}\right]^{\frac{1}{6}} \tag{59}$$

Secondly, the temperature transition between the high and low temperature is often slow, so β_{r} and β_{f} can not be ignored. If the average value of the diffusion coefficient in the transition phase is used, using a similar technique as above, we have

$$N_{\mathrm{it,qn}} = (K'c\tau)^{\frac{1}{6}}\cdot\left[\frac{n_{\mathrm{h}}D_{\mathrm{h}}+n_{\mathrm{l}}D_{\mathrm{l}}+n_{\mathrm{r}}D_{\mathrm{r}}+n_{\mathrm{f}}D_{\mathrm{f}}}{(1-\alpha^6)n}\right]^{\frac{1}{6}} \tag{60}$$

Compared the above equation with Eq. (33), the interface trap generation due to NBTI with temperature variation can be estimated by the same equation as the constant temperature case, if the diffusion coefficient is replaced by an effective one

$$D_{\mathrm{eff}} = \frac{t_{\mathrm{h}}D_{\mathrm{h}}+t_{\mathrm{l}}D_{\mathrm{l}}+t_{\mathrm{r}}D_{\mathrm{r}}+t_{\mathrm{f}}D_{\mathrm{f}}}{\tau_{\mathrm{T}}} \tag{61}$$

According to Eq. (39), we can use the diffusion coefficient under a reference temperature, such as D_{h}, to denote D_{eff}. First of all, because the variation between high and low temperature is linear, we can use three point trapezium approximation to compute D_{f} and D_{r}

$$D_{\mathrm{f}} = D_{\mathrm{r}} \approx \frac{D_{\mathrm{h}}+D_{\mathrm{l}}}{4}+\frac{D\left(\frac{T_{\mathrm{h}}+T_{\mathrm{l}}}{2}\right)}{2} \tag{62}$$

We can get the effective diffusion coefficient if D_{h} is taken as a reference,

$$\begin{aligned} D_{\mathrm{eff}} &= \frac{t_{\mathrm{h}}D_{\mathrm{h}}+t_{\mathrm{l}}\cdot\mu_{\mathrm{l}}D_{\mathrm{h}}+(t_{\mathrm{f}}+t_{\mathrm{r}})\left(\frac{D_{\mathrm{h}}+\mu_{\mathrm{l}}D_{\mathrm{h}}}{4}+\frac{\mu_{\mathrm{m}}D_{\mathrm{h}}}{2}\right)}{\tau_T} \\ &= \frac{t_{\mathrm{h}}+\mu_{\mathrm{l}}t_{\mathrm{l}}+(t_{\mathrm{f}}+t_{\mathrm{r}})\left(\frac{1+\mu_{\mathrm{l}}}{4}+\frac{\mu_{\mathrm{m}}}{2}\right)}{\tau_T}D_{\mathrm{h}} \end{aligned} \tag{63}$$

From Eq. (39), we can derive the ratios μ_{l} and μ_{m} in Eq. (63)

$$\mu_{\mathrm{l}} = \frac{D_{\mathrm{l}}}{D_{\mathrm{h}}} = \exp\left[-\frac{E_{\mathrm{d}}}{k_{\mathrm{B}}}\left(\frac{1}{T_{\mathrm{L}}}-\frac{1}{T_{\mathrm{H}}}\right)\right]$$

$$\mu_{\mathrm{m}} = \frac{D\left(\frac{T_{\mathrm{h}}+T_{\mathrm{l}}}{2}\right)}{D_{\mathrm{h}}} = \exp\left[-\frac{E_{\mathrm{d}}}{k_{\mathrm{B}}}\left(\frac{2}{T_{\mathrm{H}}+T_{\mathrm{L}}} - \frac{1}{T_{\mathrm{H}}}\right)\right] \quad (64)$$

The above derivation shows that we can still use the same dynamic model to estimate the number of the generated interface traps due to NBTI effect with variable temperature, if the diffusion coefficient of a constant temperature is transformed to the effective diffusion coefficient by Eq. (63). In fact, the above derivation can be extended to the multiple temperature case, and the effective coefficient can be calculated.

5. V_{th} and Circuit Delay Degradation Model

Previous section describes the analytical model of interface trap generation due to NBTI considering the temperature variation. In order to evaluate the temporal performance degradation due to NBTI, the threshold voltage degradation model and circuit delay model are described in the following sections.

5.1 PMOS V_{th} Degradation Model

In this section, we will derive the threshold voltage degradation model due to NBTI. Two circuit operation cases are considered.

(1) Case A

The circuit runs all the time, but execute different tasks. The tasks have different work load, so the temperature corresponding different executing task is different. We assume that there are two tasks, $\mathrm{TASK_H}$ and $\mathrm{TASK_L}$. When the circuit execute $\mathrm{TASK_H}$, the temperature is T_{H}, and $\mathrm{TASK_L}$ corresponds to T_{L}. If the circuit switch between the two tasks, the temperature will change as the waveform in Fig. 5. Therefore, the threshold voltage degradation is that

$$\Delta V_{\mathrm{th}} = \Delta V_{\mathrm{th,H}}\left(\frac{D_{\mathrm{eff}}}{D_{\mathrm{H}}}\right)^{\frac{1}{6}} \quad (65)$$

where $\Delta V_{\mathrm{th,H}}$ is the V_{th} degradation at a constant temperature T_{H}, and the ratio $D_{\mathrm{eff}}/D_{\mathrm{H}}$ can be calculated by Eq. (63).

(2) Case B

The circuit has two modes, active and standby modes. In active mode, the circuit runs as normal and the temperature will be high; while in standby mode, the circuit stops, and remains in low power mode, so the circuit is considered to be in room temperature (about 300 K). This case is describe in Fig. 6. In active mode, the temperature is T_{H} and the period of the input voltage is τ with the duty cycle of c. In standby mode, the temperature is T_{L}, and the input voltage remains in logic 1 or logic 0. We assume the circuit operates in active and standby mode alternately, and in one cycle, the active time is t_{A}, while the standby time is t_{S}. We also denote the ratio of active to standby time as R_{AS}. As shown in Fig. 6, when the circuit changes from standby mode to active mode,

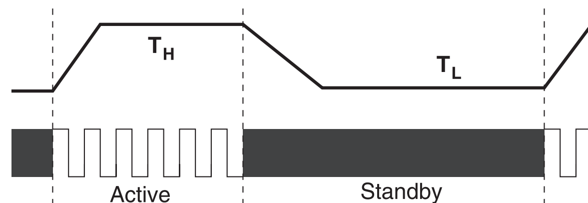


Fig. 6 Temperature variation in active and standby mode.

the temperature begins to increase from T_{L} to T_{H}, and the rise time is t_{R}. The fall time of the temperature is defined as t_{F}, similarly.

In order to calculate the threshold voltage degradation, firstly, the effective diffusion coefficient should be used just as in *case A*. In active mode, the input voltage is a periodic square wave, but in standby mode, the input voltage is set to logic 1 or logic 0 constantly. From [16], the NBTI effect by a random signal can be equivalent as a square wave signal, if these two waves have the same signal probability. Therefore, we should transform the overall input signal to a square wave signal, and use an effective duty cycle c to calculate the threshold voltage degradation.

If the input signal in standby mode remains logic 0, the overall signal probability of logic 0 (effective duty cycle of stress phase) is

$$c_0 = \frac{ct_{\mathrm{A}} + t_{\mathrm{S}}}{t_{\mathrm{A}} + t_{\mathrm{S}}} = \frac{cR_{\mathrm{AS}} + 1}{R_{\mathrm{AS}} + 1} \quad (66)$$

If the input signal remains logic 1 in standby mode, the effective duty cycle is

$$c_1 = \frac{cR_{\mathrm{AS}}}{R_{\mathrm{AS}} + 1} \quad (67)$$

Therefore, if we know the input of the given PMOS transistor in standby mode, the effective duty cycle c_0 or c_1 can be calculated. If the information is unknow, we assume that the chance of logic 1 and logic 0 is the same, so the effective duty cycle c_{M} is

$$c_{\mathrm{M}} = \frac{cR_{\mathrm{AS}} + 0.5}{R_{\mathrm{AS}} + 1} \quad (68)$$

In both *case A* and *case B*, we can estimate the threshold voltage degradation due to NBTI with temperature variation.

5.2 Gate and Circuit Delay Degradation Analysis

In this paper, the delay of a gate v can be approximately expressed as [19]

$$d(v) = \frac{C_{\mathrm{L}}V_{\mathrm{dd}}}{I_{\mathrm{d}}} = \frac{K}{(V_{\mathrm{gs}} - V_{\mathrm{th}})^{\alpha}}$$
$$K = \frac{C_{\mathrm{L}}V_{\mathrm{dd}}}{\mu C_{\mathrm{ox}}W_{\mathrm{eff}}/L_{\mathrm{eff}}}. \quad (69)$$

where α is the velocity saturation index, whose value ranges from 1 to 2, and depends on the type of logic gate.

Depending on the circuit operation states, the shift in the transistor threshold voltage ΔV_{th} can be derived using the approach in the above section. Hence, the delay degradation $\Delta d(v)$ for gate v can be derived as

$$\begin{aligned}\Delta d(v) &= \frac{K}{(V_{gs}-V_{th}-\Delta V_{th})^{\alpha}} - \frac{K}{(V_{gs}-V_{th})^{\alpha}} \\ &= \left[\left(1-\frac{\Delta V_{th}}{V_g - V_{th}}\right)^{-\alpha} - 1\right] d(v)\,.\end{aligned} \quad (70)$$

We use Taylor series expansion on the right side of Eq. (70), neglecting the higher order terms, giving that

$$\Delta d(v) = \frac{\alpha \Delta V_{th}}{V_{gs} - V_{th0}} \times d(v) \quad (71)$$

where V_{th0} is the original transistor threshold voltage and $d(v)$ is the original delay of gate v.

6. Simulation Results

In this section, we present the simulation results in order to validate the NBTI models. Firstly, a single PMOS transistor is used to evaluate the NBTI effect considering temperature variation. Secondly, ISCAS85 circuits are used as the benchmark. In this paper, the standard cell library is constructed using the PTM 65 nm bulk CMOS model [28]. V_{dd} = 1.2V, $|V_{th}|$ = 220 mV are set for all the transistors in the circuits. Our NBTI aware Static Timing Analysis is based on our previous STA tool [29].

6.1 Single PMOS Device NBTI Analysis

As in Sect. 5.1, we consider two circuit operation cases. The default duty cycle of the PMOS transistor c is set to 0.5, and the overall operation time is 10 yr (about 3×10^8s).

(1) Case A

The circuit works all the time, and executes two tasks, $TASK_H$ and $TASK_L$. We assume the circuit switches between these two tasks periodically, and in one cycle, the executing time of $TASK_H$ is t_H, while time of $TASK_L$ is t_L. The transition time of the temperature is denoted as t_{TR}. The temperature executing $TASK_H$ is 373 K (100°C), and $TASK_L$ corresponds to 323 K (50°C). The V_{th} degradation of PMOS transistor in the following five cases is shown in Table 1.

A1) the circuit executes only one task $TASK_H$, and $t_L = 0$, so the temperature keep at 373 K;
A2) t_H = 1 h, t_L = 1 h, and t_{TR} = 0;
A3) t_H = 1 h, t_L = 1 h, and t_{TR} = 10 min;
A4) t_H = 1 h, t_L = 1 h, and t_{TR} = 20 min;
A5) t_H = 1 h, t_L = 1 h, and t_{TR} = 30 min.

Table 1 shows that the execution of the task $TASK_L$ with low temperature compensates 10.7% of V_{th} degradation for the PMOS transistor. We should also notice that the impact of the temperature transition time can be ignored.

Table 1 V_{th} degradation in case A1 to A5.

Case	A1	A2	A3	A4	A5
ΔV_{th}	52.6 mV	47.5 mV	47.3 mV	47.0 mV	46.7 mV

Table 2 V_{th} degradation with diffrent R_{HL} and T_L.

R_{HL}	9:1	3:1	1:1	1:3	1:9
T_L = 353 K	52.1 mV	51.2 mV	49.7 mV	47.8 mV	46.5 mV
T_L = 323 K	51.7 mV	50.4 mV	47.5 mV	43.5 mV	39.7 mV

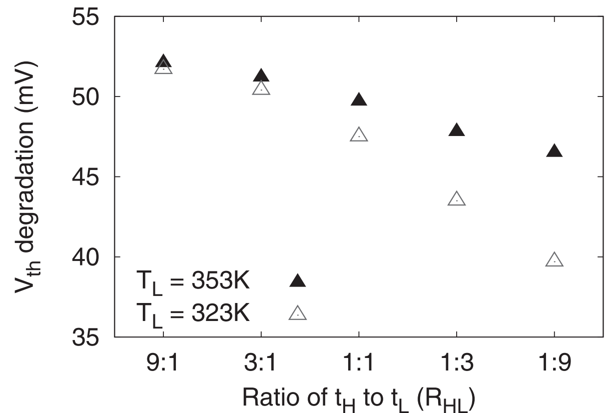


Fig. 7 ΔV_{th} with different R_{HL} and T_L.

We denote the ratio of time t_H to t_L as R_{HL}, and V_{th} degradation with different R_{HL} is shown in Table 2. The simulation is also run at another $TASK_L$ temperature T_L = 353 K. The comparison is shown in Fig. 7.

Figure 7 shows that longer execution time of $TASK_L$ decreases V_{th} degradation, and lower $TASK_L$ temperature enlarges this effect. If the temperature of $TASK_L$ is 323 K, and the ratio of these two tasks' execution time is 1 : 9, the threshold voltage degradation can be decreased by 24.5%.

(2) Case B

The circuit periodically switches between active and standby mode. We still assume that the temperature transition time both from active to standby and from standby to active is the same, which is denoted as t_{TR}. The temperature of active mode is T_A = 373 K, and temperature of standby mode T_S = 303 K. The active and standby time is t_A and t_S respectively. The V_{th} degradation of PMOS transistor in the following five cases is shown in Table 3.

B1) the circuit is always active, thus $t_S = 0$, so the temperature keep at 373 K;
B2) t_A = 1 h, t_S = 1 h, t_{TR} = 0, and gate input V_g = 0;
B3) t_A = 1 h, t_S = 1 h, t_{TR} = 30 min, and V_g = 0;
B4) t_A = 1 h, t_S = 1 h, t_{TR} = 0 min, and V_g = 1;
B5) t_A = 1 h, t_S = 1 h, t_{TR} = 30 min, and V_g = 1.

Table 3 also shows that impact of temperature transition time can be ignored, and logic 1 of the gate input can significantly decrease V_{th} degradation by 23.0%. We denote the ratio of active time t_A to standby time t_S as R_{As}, and V_{th} degradation with different R_{AS} and different gate input

Table 3 V_{th} degradation in case B1 to B5.

Case	B1	B2	B3	B4	B5
ΔV_{th}	52.6 mV	50.9 mV	49.5 mV	41.6 mV	40.5 mV

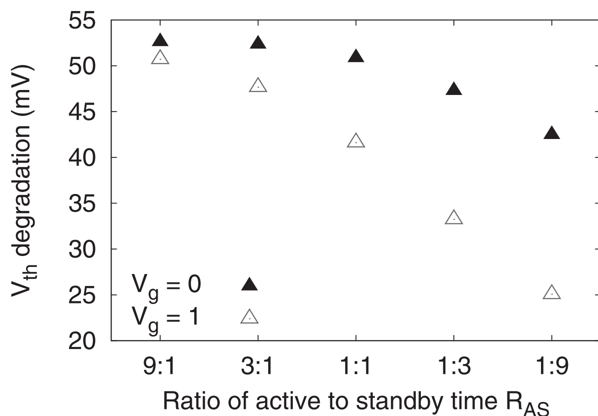


Fig. 8 ΔV_{th} with different R_{AS} and gate input.

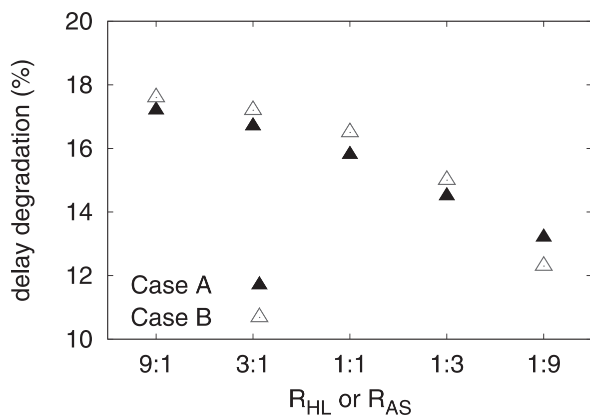


Fig. 9 ΔV_{th} with different R_{HL} in Case A and R_{AS} in Case B.

is shown in Fig 8.

From Fig. 8, we can conclude that 1) longer standby time leads to a smaller NBTI effect on the V_{th} degradation; 2) logic 1 gate input significantly increase the above effect, and the maximum decrease of ΔV_{th} can reach up to 52.3%.

6.2 Performance Degradation Analysis in Digital Circuits

Figure 9 shows the performance degradation of ISCAS85 c432 benchmark with different R_{HL} in Case A and R_{AS} in Case B. In circuit simulation, we do not set all the states of internal node at standby mode, but use a logic simulator to get the logic states. From Fig. 9, we can conclude that longer time of a low temperature state ($TASK_L$ execution time in Case A and standby mode in Case B decreases the circuit performance degradation. Therefore, we can use this to mitigate the NBTI effect. Table 4 shows the result of the circuit delay degradation analysis. In Case A, R_{HL} is set to 1:1, and in Case B, the ratio of active to standby time R_{AS} is set to 1 : 9. The Δdelay is around 18.0% of the original circuit delay when the circuit runs at a constant temperature 373 K. If the temperature variation is considered, the Δdelay in Case A is around 15.9%, which decreases by 13.2% compared to the constant temperature case, and the Δdelay in Case B decreases by 45.2%.

Table 4 Delay degradation of ISCAS85 benchmark circuits under NBTI.

ISCAS85 Circuits	Nominal delay (ns)	373 K Δdelay (%)	Case A Δdelay	Case B Δdelay
c432	4.993	17.5	15.6	12.3
c499	1.978	17.0	15.2	12.2
c880	3.386	18.5	16.5	12.6
c1355	3.818	17.1	15.3	12.7
c1908	4.576	17.4	15.6	11.6
c2670	4.706	17.2	15.4	13.5
c3540	6.035	19.7	17.7	11.8
c5315	6.262	19.5	17.5	13.9
c6288	21.06	18.3	16.3	12.3
c7552	6.487	17.4	14.0	11.5
Average	N/A	18.0	15.9	12.4

7. Conclusion

In this paper, we propose an analytical model of temporal performance degradation due to PMOS NBTI effect under AC stress condition. We also demonstrate that the temperature variation due to the execution of different tasks and the change of circuit operation mode (active and standby mode) can have significant impact on PMOS NBTI effect: executing a low temperature task or remaining in standby mode can mitigating the effect of NBTI. Therefore, we propose an analytical model that considers the temperature variation, such that a more accurate performance degradation estimation can be performed. Our results of the circuit delay degradation analysis show that the existence of the standby mode can significantly decrease the delay degradation by 45.2%.

References

[1] V. Huard, M. Denais, and C. Parthasarathy, “NBTI degradation: From physical mechanisms to modelling,” Microelectron. Reliab., vol.46, no.1, pp.1–23, 2006.

[2] L. Peters, “NBTI: A growing threat to device reliability,” Semiconductor International, vol.27, no.3, March 2004, www.reed-electronics.com/semiconductor/article/CA386329

[3] J. Stathis and S. Zafar, “The negative bias temperature instability in MOS devices: A review,” Microelectron. Reliab., vol.46, no.2-4, pp.270–286, 2006.

[4] K. Bernstein, D.J. Frank, A.E. Gattiker, W. Haensch, B.L. Ji, S.R. Nassif, E.J. Nowak, D.J. Pearson, and N.J. Rohrer, “High-performance CMOS variability in the 65-nm regime and beyond,” IBM J. Res. Dev., vol.50, no.4/5, pp.433–449, 2006.

[5] H. Kufluoglu and M. Alam, “Theory of interface-trap-induced NBTI degradation for reduced cross section MOSFETs,” IEEE Trans. Electron Devices, vol.53, no.5, pp.1120–1130, 2006.

[6] R. Wittmann, H. Puchner, L. Hinh, H. Ceric, A. Gehring, and S. Selberherr, “Impact of NBTI-driven parameter degradation on lifetime of a 90 nm p-MOSFET,” Proc. Intl. Integrated Reliab. Workshop Final Report, pp.99–102, 2005.

[7] V. Reddy, A.T. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, and S. Krishnan, “Impact of negative bias temperature instability on digital circuit reliability,” Microelectron. Reliab., vol.45,

no.1, pp.31–38, 2005.

[8] S. Kumar, C. Kim, and S. Sapatnekar, "Impact of NBTI on SRAM read stability and design for reliability," Proc. ISQED, pp.210–218, 2006.

[9] S. Ogawa and N. Shiono, "Generalized diffusion-reaction model for the low-field charge-buildup instability at the Si-SiO_2 interface," Phys. Rev. B, vol.51, no.7, pp.4218–4230, 1995.

[10] M. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," Microelectron. Reliab., vol.45, no.1, pp.71–81, 2005.

[11] S. Mahapatra, D. Saha, D. Varghese, and P. Kumar, "On the generation and recovery of interface traps in MOSFETs subjected to NBTI, FN, and HCI stress," IEEE Trans. Electron Devices, vol.53, no.7, pp.1583–1592, 2006.

[12] G. Chen, M. Li, C. Ang, J. Zheng, and D. Kwong, "Dynamic NBTI of p-MOS transistors and its impact on MOSFET scaling," IEEE Electron Device Lett., vol.23, no.12, pp.734–736, 2002.

[13] S. Mahapatra, P.B. Kumar, T. Dalei, D. Sana, and M. Alam, "Mechanism of negative bias temperature instability in CMOS devices: Degradation, recovery and impact of nitrogen," IEDM Tech. Dig., pp.105–108, 2004.

[14] V. Reddy, A. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, and S. Krishnan, "Impact of negative bias temperature instability on digital circuit reliability," Proc. IEEE IRPS, pp.248–254, 2002.

[15] B. Paul, K. Kang, H. Kufluoglu, M. Alam, and K. Roy, "Impact of NBTI on the temporal performance degradation of digital circuits," IEEE Electron Device Lett., vol.26, no.8, pp.560–562, 2005.

[16] S. Kumar, C. Kim, and S. Sapatnekar, "An analytical model for negative bias temperature instability," Proc. IEEE/ACM ICCAD, pp.493–496, 2006.

[17] R. Vattikonda, W. Wang, and Y. Cao, "Modeling and minimization of PMOS NBTI effect for robust nanometer design," Proc. DAC, pp.1047–1052, 2006.

[18] S. Bhardwaj, W. Wang, R. Vattikonda, Y. Cao, and S. Vrudhula, "Predictive modeling of the NBTI effect for reliable design," Proc. CICC, pp.189–192, 2006.

[19] B. Paul, K. Kang, H. Kufluoglu, M. Alam, and K. Roy, "Temporal performance degradation under NBTI: Estimation and design for improved reliability of nanoscale circuits," DATE, pp.780–785, 2006.

[20] S.V. Kumar, C.H. Kim, and S.S. Sapatnekar, "NBTI-aware synthesis of digital circuits," Design Automation Conference, 2007, DAC'07, 44th ACM/IEEE, pp.370–375, 2007.

[21] B. Zhang and M. Orshansky, "Modeling of NBTI-induced PMOS degradation under arbitrary dynamic temperature variation," 9th International Symposium on Quality Electronic Design, pp.774–779, 2008.

[22] H. Luo, Y. Wang, K. He, R. Luo, H. Yang, and Y. Xie, "Modeling of PMOS NBTI effect considering temperature variation," International Symposium on Quality Electronic Design (ISQED), pp.139–144, 2007.

[23] Y. Wang, H. Luo, K. He, R. Luo, H. Yang, and Y. Xie, "Temperature-aware NBTI modeling and the impact of input vector control on performance degradation," Design, Automation & Test in Europe Conference & Exhibition, 2007, DATE'07, pp.1–6, 2007.

[24] A.T. Krishnan, C. Chancellor, S. Chakravarthi, P.E. Nicollian, V. Reddy, and A. Varghese, "Material dependence of hydrogen diffusion: Implication for NBTI degradation," IEDM, 2005.

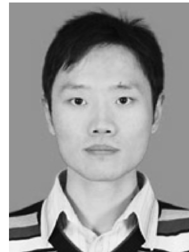
[25] A.T. Krishnan, S. Chakravarthi, P. Nicollian, V. Reddy, and S. Krishnan, "Negative bias temperature instability mechanism: The role of molecular hydrogen," Appl. Phys. Lett., vol.88, no.15, p.153518, 2006.

[26] A. Krishnan, V. Reddy, S. Chakravarthi, J. Rodriguez, S. John, and S. Krishnan, "NBTI impact on transistor and circuit: Models, mechanisms and scaling effects [MOSFETs]," IEDM, pp.14.5.1–14.5.4, 2003.

[27] R. Wittmann, Miniaturization problems in CMOS technology: Investigation of doping profiles and reliability, Ph.D. Thesis, Vienna University of Technolog, 2007.

[28] Nanoscale Integration and Modeling (NIMO) Group, ASU, "Predictive technology model (PTM)," 2007.

[29] Y. Wang, H. Yang, and H. Wang, "Signal-path level dual-Vt assignment for leakage power reduction," J. Circuits Syst. Comput., vol.15, no.2, pp.179–216, 2006.



Hong Luo received his B.S. and Ph.D. degrees in Circuits and Systems Division, E.E. Dept, Tsinghua University. His research interests includes leakage current modeling and optimization, reliability aware modeling.

Yu Wang received his B.S. degree in Tsinghua University, China in 2002, and then Ph.D. degree with honor in NICS Group, Electronics Engineering Department, Tsinghua University in 2007, supervised by Prof. Huazhong Yang (Tsinghua University) and Prof. Yuan Xie (Penn. State University). He is now a faculty member in E.E. Dept., Tsinghua University. Dr. Wang's research mainly focuses on fast circuit analysis, low power circuit design methodology, reliability-aware circuit design methodology, application specific FPGA design, and on-chip communication strategies for MPSOC.

Rong Luo received her double B.S. in Engineering Physical and Electronic Engineering from Tsinghua University in 1992 and her Ph.D. from Tsinghua University in 1997. Currently, she is an associate Professor in the Department of Electronic Engineering, Tsinghua University, Beijing. Now, her research work is mainly on SoC Design Technology, VLSI Design and Embedded System Design Technology.

Huazhong Yang was born in Ziyang, Sichuan Province, P.R. China, on Aug.18, 1967. He received B.S. degree in microelectronics in 1989, M.S. and Ph.D. degree in electronic engineering in 1993 and 1998, respectively, all from Tsinghua University, Beijing. In 1993, he joined the Department of Electronic Engineering, Tsinghua University, Beijing, where he has been a Full Professor since 1998. Dr. Yang was recognized as '2000 National Palmary Young Researcher by NSFC. His research interests include chip design for communication and multimedia applications, synthesis of analog integrated circuits (IC), power estimation and synthesis of digital ICs, noise and delay estimation of deep submicron ICs, yield enhancement, optimization and modeling. Dr. Yang has authored and co-authored over 80 technical papers, 6 books, and 23 patents.

Yuan Xie is Associate Professor in Computer Science and Engineering department at the Pennsylvania State University. He received the B.S. degree in electronic engineering from Tsinghua University in Beijing, the M.S. and Ph.D. degrees in electrical engineering from Princeton University. He was a recipient of the SRC Inventor Recognition Award in 2002, NSF CAREER award in 2006, and IBM Faculty Award in 2008. He also received Best Paper Award in ASP-DAC 2008 and ASICON 2001. He is currently Associate Editor for IEEE Transaction on VLSI and IET Computers and Digital Techniques.