Received 10 March 2020; accepted 6 April 2020. Date of publication 22 April 2020; date of current version 7 August 2020.
The review of this article was arranged by Editor L. Lukasiak.

Digital Object Identifier 10.1109/JEDS.2020.2989629

Physical Model of Low-Temperature to Cryogenic Threshold Voltage in MOSFETs

ARNOUT BECKERS [1] (Student Member, IEEE), FARZAN JAZAERI [1] (Member, IEEE), ALEXANDER GRILL [2,3], SUBRAMANIAN NARASIMHAMOORTHY [2], BERTRAND PARVAIS [2,4], AND CHRISTIAN ENZ [1] (Fellow, IEEE)

1 Integrated Circuits Laboratory, Ecole Polytechnique Fédérale de Lausanne, 2000 Neuchâtel, Switzerland
2 imec, 3001 Leuven, Belgium
3 Department of Electrical Engineering (ESAT), MICAS, KU Leuven, 3000 Leuven, Belgium
4 Department of Electronics and Informatics, Vrije Universiteit Brussel, 1050 Brussel, Belgium

CORRESPONDING AUTHOR: A. BECKERS (e-mail: arnout.beckers@epfl.ch)

(Invited Paper)

ABSTRACT This article presents a physical model of the threshold voltage in MOSFETs valid down to 4.2 K. Interface traps close to the band edge modify the saturating temperature behavior of the threshold voltage observed in cryogenic measurements. Dopant freezeout, bandgap widening, and uniformly distributed traps in the bandgap do not change the qualitative behavior of the threshold voltage over temperature. Care should be taken because dopant freezeout results in a different physical definition of the threshold voltage. Using different definitions changes significantly the threshold current level. The proposed model is experimentally validated with measurements in large-area nMOS and pMOS devices of a commercial 28-nm bulk CMOS process down to 4.2 K. Our modeling results suggest that a pMOS-specific phenomenon in the gate stack is responsible for the non-saturating temperature behavior of the threshold voltage in pMOS devices.

INDEX TERMS 28-nm bulk CMOS, cryogenic, cryo-CMOS, freezeout, incomplete ionization, interface traps, MOSFET, threshold voltage.

I. INTRODUCTION

Threshold voltage (V_T) in MOSFETs increases when reducing the temperature [1], [2]. This qualitative trend of V_T is first linear with temperature and then saturates in the cryogenic regime around 50 K. Many works have stated the importance of the temperature dependent Fermi level in the bulk to explain this trend [3]–[6]. Much attention has been paid in the literature to the impact of dopant freezeout on the operation of MOS devices at cryogenic temperatures [7]–[12]. However, the relative impact of the dopant freezeout and the temperature dependence of the bulk Fermi level on V_T remains unclear. It is usually thought that dopant freezeout must be important in predicting V_T because the frozen-out dopants in the channel would need more 'ionizing' voltage to form the depletion region and turn on the MOSFET [13]–[15]. However, dopant freezeout is of minor importance to predict the qualitative behavior of V_T over temperature in enhancement-mode devices [4]. Furthermore, the expected saturation of $V_T(T)$ from the bulk Fermi level behavior is not always measured at cryogenic temperatures. Rather, an almost linear increase is measured in the pMOS devices, showing an additional kink-like feature [13, Fig. 3], [9, Fig. 4]. This was previously attributed to the field-assisted ionization of frozen-out dopants in the channel [13]. However, this ionization process should already be completed before threshold is reached as evidenced by low-temperature C-V plots [16]–[20]. Instead, interface traps close to the band edge are known to be important at cryogenic temperature [21], [22]. Recently, a Gaussian distribution of traps close to the band edge has been introduced to model the inflection of transfer characteristics at cryogenic temperatures [23], [24]. The impact of these traps on the low-temperature V_T has not yet been investigated. Furthermore, care should be taken when choosing an extraction method for the trend of V_T over temperature. The temperature dependence of V_T obtained from the widely-used

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

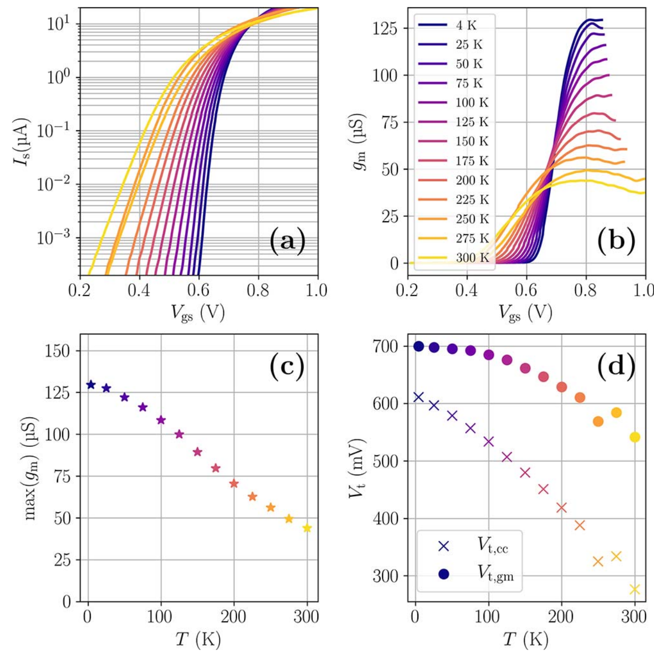

FIGURE 1. **nMOS measurements ($W/L = 10\,\mu m/1\,\mu m$). a) Transfer characteristics ($V_{DS} = 10$ mV, $V_{SB} = 0$ V), b) Transconductance (g_m), c) Maximum transconductance versus temperature, d) Extracted threshold voltage with constant current (cc) and max(g_m) methods.**

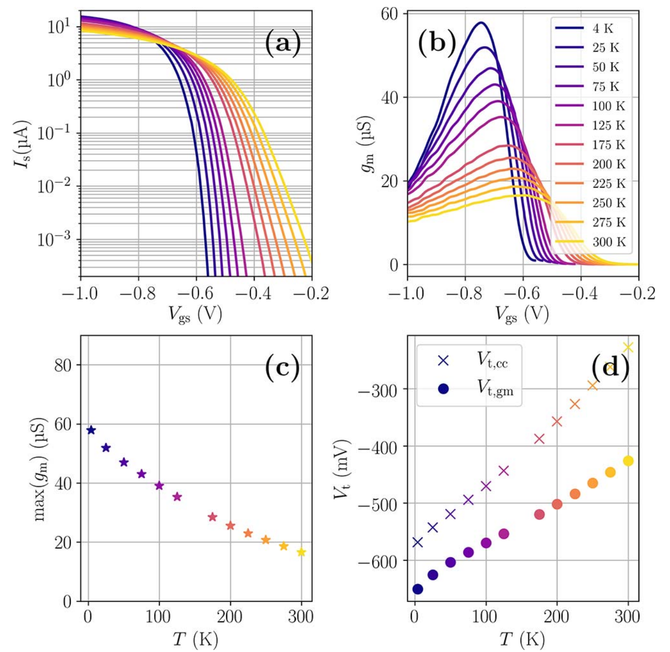

FIGURE 2. **pMOS measurements ($W/L = 10\,\mu m/1\,\mu m$). a) Transfer characteristics ($|V_{DS}| = 10$ mV, $|V_{SB}| = 0$ V), b) Transconductance (g_m), c) Maximum transconductance versus temperature, d) Extracted threshold voltage with constant current (cc) and max(g_m) methods.**

extraction methods is different [25], i.e., (i) the constant current (cc) method, (which defines V_T as the gate voltage at a constant current level), and (ii) the maximum transconductance method, (which finds V_T from a linear extrapolation from the maximum transconductance point on the current in the linear regime) [26], [27]. Furthermore, the physics definition of V_T as used in textbooks, which normally corresponds to the gate voltage at two times the bulk Fermi potential at room temperature, will have to be revised at cryogenic temperatures due to dopant freezeout in the substrate [28], [29]. Finally, the presented results on the temperature behavior of V_T are important for the modeling, reliability studies, and optimization of commercial CMOS processes for cryogenic operation, which is timely for the development of quantum computation systems [30]–[37].

II. EXPERIMENTAL RESULTS AND DISCUSSION

Cryogenic measurements were performed in large-area ($W/L = 10\,\mu m/1\,\mu m$) n-type and p-type devices from a commercial 28-nm bulk CMOS process with high-k metal gate. The experimental set-up consisted of a Lakeshore CRX-4K cryogenic probe station for cool-down and a dedicated data-acquisition system designed by TU Wien [38]. Similar measurements were performed in the same technology (but a different flavor) in [28], [39], [40]. Figures 1 and 2 show the measurements for the nMOS and pMOS devices, respectively. The temperature was swept from 4.2 K up to 300 K with intermediate steps at multiples of 25 K. The transfer characteristics and transconductance (g_m) of the nMOS and pMOS devices in the linear regime ($|V_{DS}| = 10$ mV and $|V_{SB}| = 0$ V) are shown on the top row of Figs. 1 and 2. The g_m curves were smoothed because of measurement noise. The peak of g_m increases at lower temperatures (due to mobility increase because of reduced phonon scattering) and the overall shape of g_m steepens at lower temperatures (due to exponential scaling of the Fermi-Dirac distribution). The extraction of the maximum of g_m and the threshold voltage is shown in the bottom rows of Figs. 1 and 2. Both max(g_m) and threshold voltage consistently increase, but also show signs of saturation for the lowest temperatures. The threshold voltage was extracted using both the max(g_m) method [26] and the constant current (cc) method [extraction currents for nMOS and pMOS were both 1 nA]. The max(g_m) method allows to remove the temperature variation of the series resistance and the subthreshold swing [26]. The two extraction methods do not give the same temperature dependence of V_T. The absolute increase obtained using the cc method is about 0.3 to 0.4 V, while it is about 0.2 V using the max(g_m) method. Furthermore, the qualitative behavior of V_T over temperature differs between the two methods. It is interesting to note that the saturation of $V_T(T)$ in the cryogenic regime (expected at about 50 K-100 K from the behavior of the bulk Fermi potential over temperature, is only observed for the nMOS. The pMOS shows a rather linear increase in V_T in the cryogenic regime. The measurements from Dao *et al.* also show that the V_T of nMOS saturates more than pMOS [13, Figs. 2 and 3]. Furthermore, a kink-like feature is observed in $V_T(T)$ in pMOS [13, Fig. 3]. Therefore, effects other than the bulk Fermi potential are at play that have an impact on the temperature dependence of V_T.

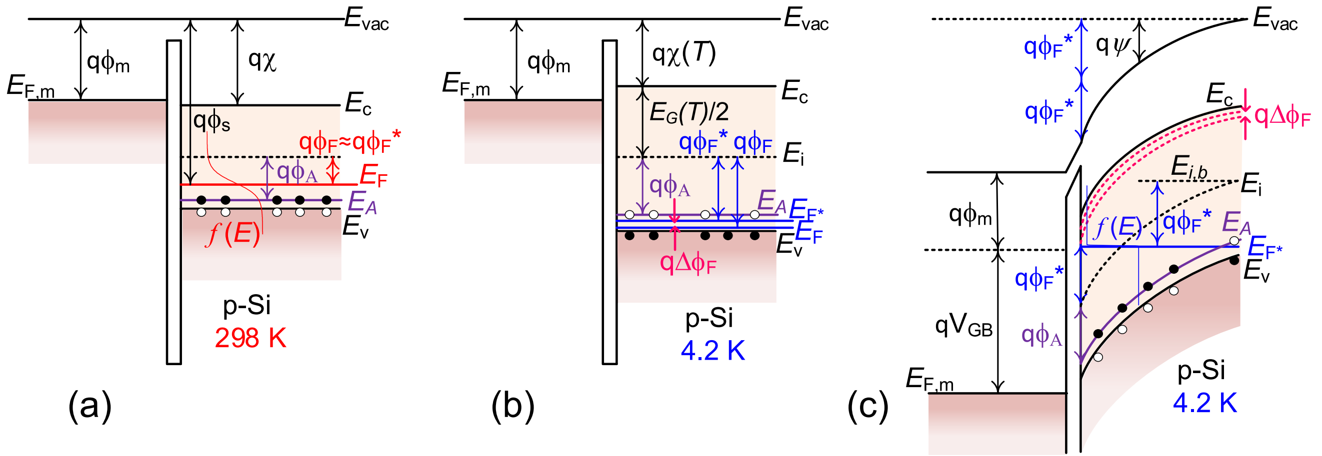

FIGURE 3. MOS band diagrams, a) Flatband at room temperature, b) Flatband at 4.2 K, and c) Band bending $\psi_s' = 2\Phi_F^*$ ($n = N_A^-$ at the interface). An additional band bending of $q\Delta\Phi_F$ is required to increase the inverted carrier density from $n = N_A^-$ to $n = N_A$.

III. ANALYSIS OF DIFFERENT CONTRIBUTIONS TO TEMPERATURE DEPENDENCE OF V_T

In this section, we analyze the different temperature dependencies in the expression of V_T. For this analysis, we assume an n-type, large-area MOSFET without loss of generality for the complementary case of p-type. The goal is to find out which temperature dependencies can be neglected in the overall trend of $V_T(T)$ and which need to be kept in the model. The physics definition of V_T is used, which is based on the inversion condition of the carriers in the channel. To start, a qualitative description of the temperature behavior of V_T is given based on the MOSFET band diagrams in Fig. 3 (interface traps not included). As the temperature is reduced, the energy bands in the substrate of a bulk-silicon FET will shift such that the Fermi level (E_F) takes an energy that satisfies charge neutrality. Because of the Fermi-Dirac distribution scaling, this position of E_F is closer to the valence band edge at lower temperatures [see Figs. 3(a) and 3(b)]. Therefore, at lower temperatures, more band bending is required to generate a sufficient carrier density in the conduction band at the surface, and V_T increases (and saturates as well). However, the imposed charge neutrality condition usually assumes complete ionization of the dopants in the bulk, which is too constraining at cryogenic temperatures. Due to incomplete dopant ionization (freezeout), charge neutrality will already be satisfied when the Fermi level lies closer to the conduction band. Incomplete ionization thus results in a different position of the bulk Fermi level (E_F^*) that is slightly closer to E_c for a given temperature and doping concentration (N_A), as shown in Fig. 3b at 4.2 K. Therefore, less band bending [Fig. 3(c)] is required to reach a given inverted charge density in the channel when incomplete dopant ionization is included, and V_T is therefore slightly lower at all temperatures. However, the overall temperature behavior of V_T (increase and saturation) will not change because of dopant freezeout. As shown in Fig. 3, the Fermi potential assuming complete ionization is defined as $\Phi_F \triangleq (E_i^0 - E_F)/q$. With dopant incomplete ionization, the Fermi potential is defined as $\Phi_F^* \triangleq (E_i^0 - E_F^*)/q$, where E_i^0 or $E_{i,b}$ is the intrinsic energy in the bulk. The widening of the bandgap (typically from 1.12 eV at 300 K to 1.16 eV at 4.2 K (Varshni model [41]) will increase V_T, because more band bending is required to reach inversion. Bandgap widening and dopant freezeout are thus two competing mechanisms.

V_T is normally defined as the gate-to-source voltage (V_{GS}) at the inversion threshold ψ_s' when $p = N_A \rightarrow n = N_A$ in the channel while assuming that the mobile charge density is still negligible compared to the depletion charge density [42]. Furthermore, the channel is assumed long and in equilibrium (small V_{DS}). Therefore, the equilibrium threshold voltage (V_{T0}) is given by:

$$V_{T0} \triangleq V_{GS}(\psi_s') = \psi_s' + \Phi_{ms} - \frac{Q_{depl}(\psi_s')}{C_{ox}} - \frac{Q_{it}(\psi_s')}{C_{ox}} \quad (1)$$

where the index 0 indicates that the channel is in equilibrium. Φ_{ms} is the metal-semiconductor work function difference, Q_{depl} the depletion charge density, Q_{it} the interface-trap charge density, and C_{ox} the gate-oxide capacitance. The surface potential drop is defined as $\psi \triangleq (E_i^0 - E_i)/q$ [Fig. 3(c)]. C_{ox} can be assumed to be temperature independent in first order. This leaves four contributions in (1) to investigate: (1) ψ_s', (2) Φ_{ms}, (3) $Q_{depl}(\psi_s')$, and (4) $Q_{it}(\psi_s')$, discussed in Sections III.A–III.D. The first step is to find out the correct temperature dependence of the inversion threshold ψ_s'. This is important because ψ_s' appears in three of the four terms in (1).

A. INVERSION THRESHOLD

The channel of an n-type MOSFET is inverted when the hole density in the channel $p = N_A$ is replaced by the same density of electrons as there were of holes before: $n = N_A$. However, this has to be re-examined at cryogenic temperatures because the MOSFET channel starts out at $p = N_A^-$ due to dopant freezeout, where N_A^- is the fraction of ionized dopants. The dopants are initially (in flatband) frozen out in the channel and for bulk silicon FETs remain frozen out deep in the

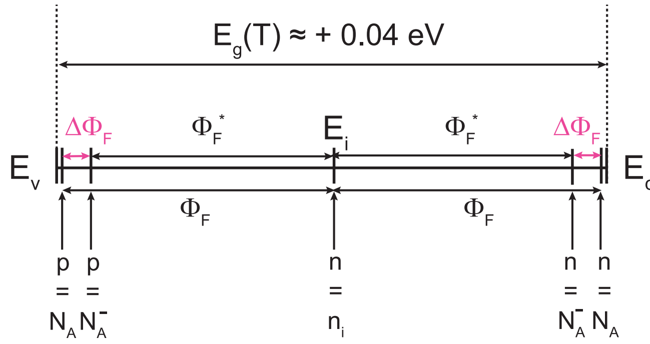

FIGURE 4. Including incomplete ionization of the dopants, the channel in the *p*-type body of an *n*-type FET has to be switched from ($p = N_A^-$) to $n = N_A$ to reach inversion threshold ($\psi_s' = 2\Phi_F^* + \Delta\Phi_F$).

bulk [29]. Therefore, the channel has to be switched from the condition $p = N_A^-$ to $n = N_A^-$ to reach an inversion of the mobile carriers in the channel. However, $n = N_A^-$ is still a very low density of electrons in the channel. The channel cannot be said to be actually inverted under this condition. Three different cases can be distinguished, which correspond to different amounts of band bending (or ψ_s') to reach inversion, as shown in Fig. 4:

- assume complete ionization ($p = N_A$); correct band bending to reach $n = N_A$ is $\psi_s' = 2\Phi_F$ as usual, where Φ_F is the Fermi potential in the bulk (of the holes) defined as $\Phi_F \triangleq (E_i^0 - E_F)/q$.
- include incomplete ionization ($p = N_A^-$) and use $n = N_A^-$ as the threshold; amount of band bending is then $\psi_s' = 2\Phi_F^*$, where Φ_F^* is the Fermi potential calculated from the charge-neutrality condition that includes incomplete ionization ($p = N_A^-$). An analytical expression for $\Phi_F^* = \Phi_F - \Delta\Phi_F$ exists as a function of T and N_A, as shown later in this section. However, the current density is too low when using $n = N_A^-$ as the threshold.
- include incomplete ionization ($p = N_A^-$) and use $n = N_A$ as the threshold; the amount of band bending in this case is $\psi_s' = 2\Phi_F^* + \Delta\Phi_F$.

Only the third case gives the correct inversion threshold when the MOSFET channel is initially frozen out and has to be switched to a decent current density. Therefore, $\psi_s' = 2\Phi_F^* + \Delta\Phi_F = \Phi_F + \Phi_F^* = 2\Phi_F - \Delta\Phi_F$, and not $2\Phi_F$, nor $2\Phi_F^*$. It can already be understood that the impact of bulk freezeout on the inversion threshold ($-\Delta\Phi_F$) will lower V_T, instead of increasing it. The amount of band bending needed to reach inversion is $\Delta\Phi_F$ lower than $2\Phi_F$.

A.1. FERMI POTENTIAL INCLUDING DOPANT FREEZEOUT

The Fermi potential including dopant freezeout (Φ_F^*) can be obtained from the charge-neutrality condition $p = N_A^-$. Since we do not know the relative distance between the Fermi level and the valence band edge (E_v) beforehand, we cannot assess whether the condition of Boltzmann validity ($E_F - E_v > 3k_BT$) is violated or not. Therefore, we have to assume

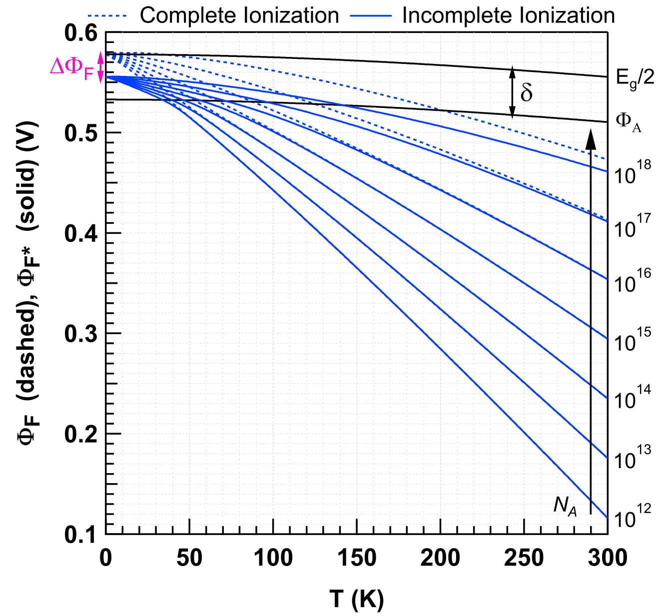

FIGURE 5. Bulk Fermi potential including dopant freezeout (solid lines) and assuming complete ionization (dashed lines). Bandgap widening is also shown using the Varshni model [41]. $\Phi_A \triangleq (E_i - E_A)/q$.

Fermi-Dirac statistics for p from the start: $p = \int g_v(E)[1 - f(E)]dE$, where $g_v(E)$ is the density-of-states in the valence band and $f(E)$ is the Fermi-Dirac distribution function. This exercise was done previously in [34, Sec. 3]. The Fermi level position was calculated numerically using Fermi-Dirac statistics. It was found that in the 0-K limit $E_F \to E_v + (E_A - E_v)/2$ when dopant freezeout is included (E_A is the acceptor dopant energy), while $E_F \to E_v$, when dopant freezeout is not included ($p = N_A$). Typically, $(E_A - E_v)/2 \approx 22.5$ meV is much larger than three times the thermal energy for instance at 4.2 K, $3k_BT \approx 1$ meV. Therefore, the Boltzmann statistics can be assumed for p when dopant freezeout is included, in order to derive an analytical expression for Φ_F^* from $p = N_A^-$:

$$n_i \exp\left(\frac{\Phi_F^*}{U_T}\right) = \frac{N_A}{1 + g_A \exp\left(\frac{\Phi_F^* - \Phi_A}{U_T}\right)} \tag{2}$$

where $\Phi_A \triangleq (E_i - E_A)/q$ (Fig. 3), $U_T \triangleq k_BT/q$ is the thermal voltage, n_i is the intrinsic carrier concentration, and $g_A = 4$ is a degeneracy factor. This gives a quadratic expression for $\exp(-\Phi_F^*/U_T)$ which has the following solution:

$$\Phi_F^* = \underbrace{U_T \ln\frac{N_A}{n_i}}_{\Phi_F} - \underbrace{U_T \ln\frac{1 + \sqrt{1 + (4\alpha N_A)/n_i}}{2}}_{\Delta\Phi_F}, \tag{3}$$

where $\alpha = g_A \exp(-\Phi_A/U_T)$. Φ_A and Φ_F^* are shown in Fig. 5 with solid lines. Φ_F is shown with dotted lines. The function $\Delta\Phi_F = \Phi_F - \Phi_F^*$ measures the difference in Fermi levels at a given T and N_A due to dopant incomplete ionization (freezeout). $\Delta\Phi_F(T, N_A)$ is shown in Fig. 6 versus T for different N_A. As shown in this figure, the maximum of $\Delta\Phi_F$ is around 25 mV. It can be checked that $\lim_{T\to 0}\Delta\Phi_F = \delta/2$

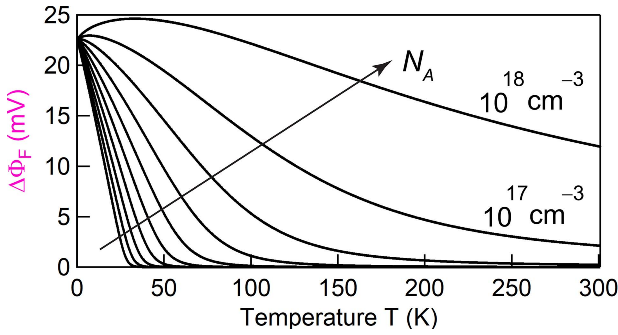

FIGURE 6. $\Delta\Phi_F = \Phi_F - \Phi_F^*$, **i.e., the difference between the dashed and the solid lines in Fig. 5. Purple color corresponds to** $\Delta\Phi_F$ **in Figs. 3, 4, and 5.**

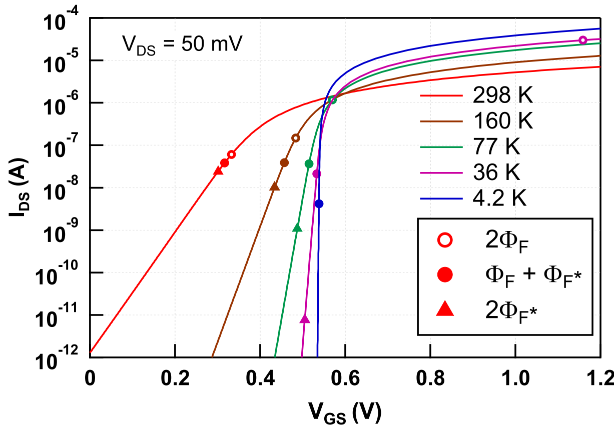

FIGURE 7. Discrepancy in threshold current levels when using different threshold definitions (starting from bulk freezeout $p = N_A^-$**).**

with $\delta = E_A - E_v = 0.045\,\text{eV}$, for typical, hydrogen-like Si:B doping and thus $\lim_{T\to 0} \Phi_F^* = \Phi_A + \delta/2$ as indicated in Fig. 5. Even though $\Delta\Phi_F$ is in millivolt range, the discrepancy of the predicted current level at threshold can be quite large at cryogenic temperatures, as shown in Fig. 7.

It should be emphasized that in (3), we have obtained $\Phi_F = U_T \ln(N_A/n_i)$ as part of the expression of Φ_F^*. $\Phi_F = U_T \ln(N_A/n_i)$ would be obtained when assuming Boltzmann statistics for p and complete ionization for N_A, which we knew was not valid because $E_F \to E_v$ in that case. However, Φ_F can be used here, because Φ_F is just a part of the derived expression for Φ_F^*, which has nothing to do with the validity of the Boltzmann statistics anymore.

B. METAL-SEMICONDUCTOR WORK FUNCTION DIFFERENCE

The difference in the metal and semiconductor work functions (including dopant freezeout) is given by

$$\Phi_{ms} = \Phi_m - \Phi_s = \Phi_m - \chi - \Phi_F^* - E_g/(2q) \quad (4)$$

where Φ_m is the metal work function, Φ_s the semiconductor work function, χ the electron affinity, and E_g the bandgap (as shown in Fig. 3). The increase in the width of the band gap, $E_g(T) \approx +0.04\,\text{eV}$ (Varshni model [41]), and the temperature dependence of the electron affinity (χ) are assumed to be

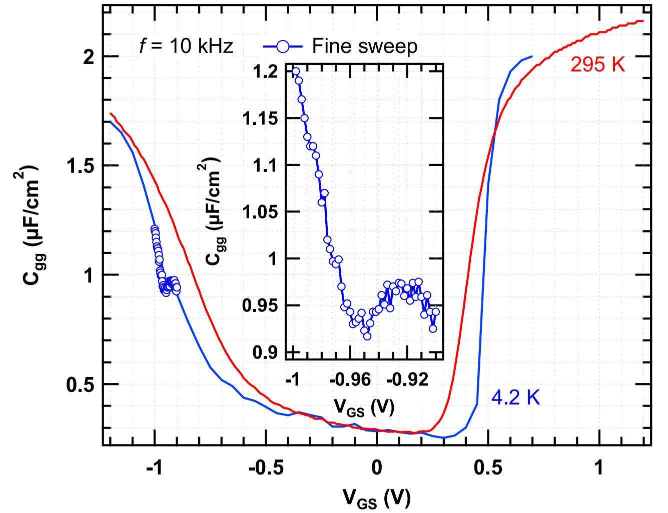

FIGURE 8. Capacitance-voltage measurements in a 1µm by 1µm n-type device of a 28-nm bulk CMOS process at 298 K and 4.2 K. Inset shows the bump close to flatband at 4.2 K due to field-assisted dopant ionization. This process is completed near inversion and can thus be neglected in the calculation of V_T**.**

balanced in this expression because of the lack of measurement data of $\chi(T)$ at cryogenic temperatures in the literature. Hence, the bandgap widening is assumed to be compensated by the reduction in $\chi(T)$ [Fig. 3(b)]. Therefore, the standard room-temperature values for silicon can be used in (4), i.e., $E_g = 1.12\,\text{eV}$ and $\chi = 4.05\,\text{eV}$. The metal work function Φ_m can be assumed temperature independent. Therefore, the only temperature dependence in (4) is inside Φ_F^* as given by (3).

C. DEPLETION CHARGES

The contribution of the depletion charges is given by

$$Q_{depl}(\psi_s') = -\Gamma_b C_{ox}\sqrt{\psi_s'} \quad (5)$$

where $\Gamma_b = \sqrt{2qN_A\varepsilon_{si}}/C_{ox}$ is the body factor. The silicon permittivity ε_{si} has no significant temperature dependence. To derive (5), we have neglected the voltage-dependent dopant ionization process close to flatband [34]. The validity of this assumption is explained and experimentally validated in the next subsection.

C.1. FIELD-ASSISTED DOPANT IONIZATION

The dopants in the deeper part of the bulk remain thermally de-ionized (frozen-out) during operation [29]. The dopants near the surface can gradually be ionized by increasing the gate or drain voltage [16]. The thickness of the ionized layer can be computed analytically, as done in earlier work in [40, Figs. 4 and 5]. Furthermore, the ionization of the dopants due to the applied voltage is observable in $C-V$ plots as a bump close to flatband [18]. In Fig. 8, this bump is measured for the first time at 4.2 K. As shown in Fig. 8, the bump happens near flatband, which means that the process of ionizing the dopants is completed before inversion threshold

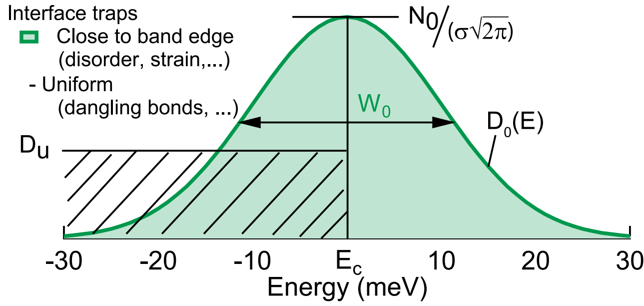

FIGURE 9. Uniform and Gaussian distributions of interface traps (standard deviation $\sigma \triangleq W_0/2$).

(ψ_s') is reached. In principle, the bump could be modeled by including N_A^- in Poisson's equation, which includes a dependence on the field [18]. However, this process can be neglected in the calculation of V_T. Therefore, the anomalous kink feature in $V_T(T)$ of pMOS (non-saturation) cannot be explained by field-assisted ionization of dopants (fitting parameters were introduced in [13]). This feature might be explained by interface traps, which is investigated in the next section.

D. INTERFACE TRAPS

The charge density per unit area captured in interface states Q_{it} is:

$$Q_{it} = -q \int_{E_i}^{+\infty} D_{it}(E) f(E) dE, \tag{6}$$

where D_{it} is the total distribution of interface traps over energy. Q_{it} can be split in two contributions: $Q_{it} = Q_u + Q_0$, which have different origins. Q_u is the standard uniform distribution of traps in the bandgap (due to dangling bonds) and Q_0 models the exponential increase of the interface-trap density near the band edge (due to disorder, strain, defects, etc.) which is measured in FETs at cryogenic temperatures [21], [22]. This exponential increase of the trap density near the band edge was also necessary to correctly model the inflection of the transfer characteristics of MOSFETs on top of the average degradation of the subthreshold slope due to conduction-band or valence-band tails [23]. A Gaussian distribution of traps (which has an exponentially decaying tail in the bandgap) was therefore introduced around the band edge [23]. As shown in Fig. 9, $D_{it}(E) = D_u(E) + D_0(E)$ and

$$D_0(E) = \frac{N_0}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(E - E_c)^2}{2\sigma^2}\right). \tag{7}$$

The Gaussian has a width W_0 of twice the standard deviation (σ) and a maximum density of $N_0/(\sigma\sqrt{2\pi})$ at E_c. Furthermore, only the negatively charged acceptor traps in the top half of the bandgap have to be considered in Q_u for n-type FET. Therefore, $Q_u = -qD_u(\psi_s - \Phi_F^*)$ (assuming a step function for $f(E)$ of the uniform distribution of traps). D_u is typically very low in advanced processes ($\approx 10^{10}\,\mathrm{cm^{-2}\,eV^{-1}}$ [43]) and can be assumed to be temperature independent. This can be justified by the fact that the density of dangling bonds at the interface will not change significantly down to cryogenic temperatures. The apparent temperature-dependent increase of the density of interface traps (as observed for instance in charge-pumping measurements [21], [22]) is due to the fact that at cryogenic temperatures the Fermi-level scans a substantial portion of the bandgap including the Gaussian tail as well; on top of the uniform distribution of traps. Because the Gauss-Fermi integral for Q_0 cannot be solved analytically [23], [44], we assume a step function for $f(E)$. This gives:

$$Q_0 = -q \int_{E_i}^{E_F} D_0(E) dE, \tag{8}$$

which has the following solution:

$$Q_0 = -q\frac{N_0}{2}\left[\mathrm{erf}\left(\frac{E_F - E_c}{(W_0/2)\sqrt{2}}\right) + 1\right]. \tag{9}$$

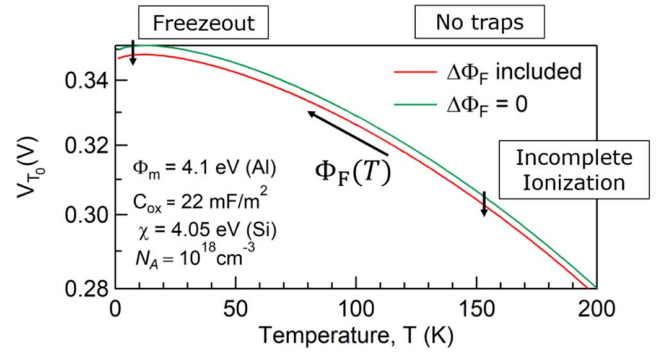

FIGURE 10. Expression (11) without traps ($D_u = D_0 = 0$). Bulk dopant freezeout ($\Delta\Phi_F$) lowers V_{T0} and does not change the qualitative behavior of V_{T0} over T. The trend is already predicted correctly by only considering the temperature dependence of the Fermi level in the bulk [bulk Fermi potential $\Phi_F(T)$].

IV. THRESHOLD VOLTAGE MODEL

The equation of V_{T0} in (1), including the depletion charges of (5) and the two distributions of interface traps in Fig. 9, is thus given by:

$$V_{T0} \triangleq V_{GS}(\psi_s') = \psi_s' + \Phi_{ms} + \Gamma_b\sqrt{\psi_s'} + \underbrace{\frac{qD_u(\psi_s' - \Phi_F^*)}{C_{ox}} + \frac{qN_0}{2C_{ox}}\left[\mathrm{erf}\left(\frac{q(\psi_s' - \psi_s^*)}{(W_0/2)\sqrt{2}}\right) + 1\right]}_{\text{traps}} \tag{10}$$

where $\psi_s^* \triangleq \Phi_F^* + E_g/(2q)$.

Combining (10), Φ_{ms} from (4) and the correct threshold $\psi_s' = 2\Phi_F^* + \Delta\Phi_F = \Phi_F + \Phi_F^* = 2\Phi_F - \Delta\Phi_F$ due to bulk freezeout, gives

$$V_{T0} = \underbrace{\Phi_F + \Phi_m'}_{\text{Sec. IV.A}} + \Gamma_b\sqrt{\underbrace{2\Phi_F}_{\text{bulk}} \underbrace{- \Delta\Phi_F}_{\text{freezeout, Sec. IV.B}}} + \underbrace{\underbrace{\frac{qD_u\Phi_F}{C_{ox}}}_{\text{uniform}} + \underbrace{\frac{qN_0}{2C_{ox}}\left[\mathrm{erf}\left(\frac{q\Phi_F - E_g(T)/2}{(W_0/2)\sqrt{2}}\right) + 1\right]}_{\text{Gaussian}}}_{\text{traps, Sec. IV.C}} \tag{11}$$

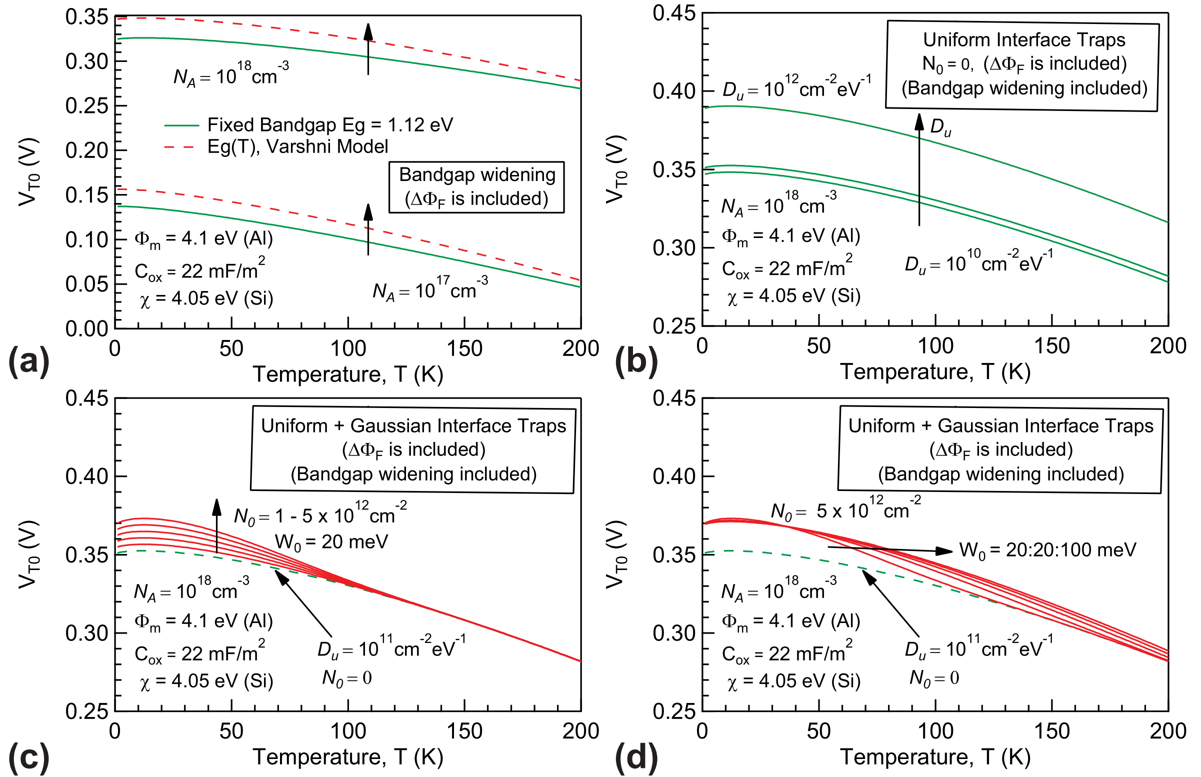

FIGURE 11. V_{T0} **plotted using (11) a) Impact of bandgap widening inside Φ_F (impact of bulk freezeout $\Delta\Phi_F$ is also included as in Fig. 10 and no traps: $D_u = 0$ and $N_0 = 0$). Varshni model is used for $E_g(T)$ [41]. b) Adding the impact of uniform interface traps (D_u) does not alter the qualitative behavior of V_{T0} over temperature. c) Adding a Gaussian distribution of additional traps close to the band edge can change the qualitative behavior of V_{T0} over temperature. The saturation of V_{T0} is delayed to lower temperatures and the overall trend is more linear around 50 to 100 K if N_0 increases. d) Impact of the Gaussian is seen at higher temperatures if W_0 increases.**

In (11), Φ_F, $\Delta\Phi_F$, and E_g are temperature dependent. It is interesting to note that the impact of the bulk freezeout (Φ_F^*) inside Φ_{ms} and ψ_s' have compensated each other in the first two terms and also in the trap terms. The only left-over impact of the dopant freezeout on V_{T0} is inside the term of the depletion charges ($-\Delta\Phi_F$ under the square root). Furthermore, we had already assumed that the temperature dependences in χ and E_g cancel out (see Section III.B). Therefore, $\Phi_m' = \Phi_m - \chi - E_g/(2q)$ is temperature independent. [Note that the temperature dependence of $E_g(T)$ is still present in the error function and in Φ_F (through n_i)].

In the following subsections the remaining temperature dependences in (11) are analyzed.

A. BULK FERMI POTENTIAL

The dominant temperature dependence in V_{T0} is the first term of (11), Φ_F, which predicts a saturation of V_{T0} at low temperatures, as shown in Fig. 10. Even though we included incomplete dopant ionization in our derivation, the first term has the usual temperature dependence of the bulk Fermi potential when assuming complete ionization, i.e., $\Phi_F = U_T \ln(N_A/n_i)$, where n_i is the intrinsic carrier concentration given by $\sqrt{N_c N_v}\exp[-E_g(T)/(2U_T)]$ (N_c and N_v are the effective density-of-states in the conduction and valence bands, respectively). Furthermore, despite Φ_F seems linearly dependent on temperature (due to the U_T dependence in front), and therefore appears to decrease when reducing T, the exponential temperature dependence in n_i is stronger, which makes Φ_F increase and saturate at low temperatures.

n_i reaches extremely low values at deep-cryogenic temperatures due to its exponential drop-off when reducing the T [34, Fig. 3]. Therefore, a variable precision arithmetic was used to compute Φ_F in Fig. 10. n_i cannot be calculated below about 10 K using standard tools that use IEEE double precision arithmetic, which is not practical. However, in the expression of Φ_F, we can benefit from the logarithm and expand Φ_F as follows:

$$\Phi_F = U_T \ln(N_A) - \frac{U_T}{2}\ln(N_c N_v) + \frac{E_g(T)}{2} \tag{12}$$

We can further expand N_c and N_v to reveal all temperature dependencies that contribute to Φ_F:

$$N_c = 2\left[\frac{2k_B T m_e^* \pi}{h^2}\right]^{3/2} \tag{13}$$

$$N_v = 2\left[\frac{2k_B T m_h^* \pi}{h^2}\right]^{3/2} \tag{14}$$

where the effective masses m_e^* and m_h^* can be assumed temperature independent. Therefore, $N_c = \alpha U_T^{3/2}$ and $N_v = \beta U_T^{3/2}$, where α and β are temperature independent constants. This gives:

$$\Phi_F = U_T(A - 1.5\ln U_T) + E_g(T)/2, \tag{15}$$

where $A = \ln(N_A) - \ln(\sqrt{\alpha\beta})$.

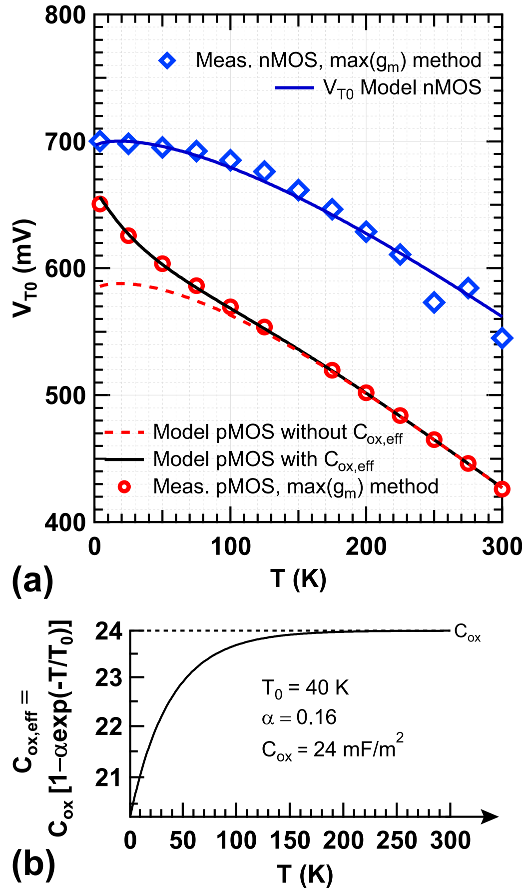

FIGURE 12. a) Experimental validation of V_{T0} model. Physical parameters given in Table 1. b) Effective gate-oxide capacitance to model the rising trend of V_{T0} in pMOS below ≈ 100 K, which cannot be explained by dopant freezeout or traps (see also [13, Fig. 3]).

B. BULK DOPANT FREEZEOUT AND BANDGAP WIDENING

As shown in Fig. 10, the impact of incomplete ionization (or freezeout) of the dopants in the bulk ($\Delta\Phi_F$) can be included for accurate extraction of physical parameters, but is not the dominant effect to explain the trend of V_{T0} over T, which is already predicted correctly by $\Phi_F(T)$. Dopant freezeout in the substrate thus lowers V_T and does not change the qualitative trend of V_T over temperature.

Figure 11(a) show the impact of bandgap widening on V_{T0}. Similarly to bulk freezeout, band gap widening does not change the trend of V_{T0}, only shifts the trend upwards by a few tens of millivolts.

C. INTERFACE TRAPS

As shown in Fig. 11(b), adding a uniform distribution of interface traps does not change the qualitative behavior of V_T from $\Phi_F(T)$ either. This can be understood from the model in (11), which keeps the same temperature dependence as $\Phi_F(T)$ since D_u is assumed to be temperature independent.

On the other hand, the Gaussian distribution of traps close to the band edge can change the trend of V_{T0} over temperature, as shown in Fig. 11(c). Increasing N_0 increases the

TABLE 1. Model parameters used in Fig. 12 for 28 nm bulk CMOS process.

$(W/L = 10\,\mu\mathrm{m}/1\,\mu\mathrm{m})$	nMOS	pMOS
Φ_m (V) (Al, TiN)	4.34	4.2
C_{ox} (mF m^{-2})	24	$C_{ox,\mathrm{eff}}$ Fig. 12(b)
χ (Si) (V)	4.05	4.05
$N_{\mathrm{A,D}}$ (cm^{-3})	2×10^{18}	2×10^{18}
D_u (cm^{-2} eV^{-1})	10^{11}	10^{11}
N_0 (cm^{-2})	7×10^{11}	1.5×10^{12}
W_0 (meV)	100	100

absolute saturation value of V_{T0} and the saturating behavior sets in at lower temperatures. This behavior resembles the pMOS measurements from Hanamura et al. [9, Fig. 4]].

This behavior can be understood from the temperature dependence of Φ_F inside the error function in (11). This dependence on T appears because we plugged ψ_s' in the error function. The exponential dependence of $f(E)$ on T was neglected in Q_0 to solve the Gauss-Fermi integral and derive the error function model. Therefore, the impact on the V_{T0} seen in Fig. 11(c) is only due to the Fermi level scanning a larger portion of the bandgap at cryogenic temperatures including the Gaussian tail, and not due to the probability of the thermal occupation of the traps. As T is reduced, Φ_F inside the error function increases. When Φ_F comes closer to $E_g/(2q)$ (Fermi level closer to band edge), the influence of the error function increases. Because W_0 is typically extending in meV-range below the band edge, the contribution of the error function is only important at cryogenic temperatures as shown in Fig. 11(c). For larger W_0 (larger exponential decay of the Gaussian in the band gap), the impact is already seen at higher temperatures [Fig. 11(d)].

However, V_{T0} still saturates when including the traps close to the band edge. Therefore, the pMOS measurements in Fig. 12(a) and by Dhao et al. [13], which show a rising trend even at cryogenic temperatures, cannot be explained by these traps alone (neither with dopant freezeout, nor field-assisted dopant ionization as proposed in [13]). Thus, another phenomenon that is specific to pMOS must be responsible for this (strain, different gate-stack, a quantum-mechanical or band-structure-related effect of the holes, etc.). The model in (11) requires that an empirical temperature dependence for C_{ox} is used, as shown in Figs. 12(a) and (b). The effective, temperature-dependent C_{ox} drops off exponentially below a critical temperature (T_0) from 24 mF m^{-2} at 295 K to 20 mF m^{-2} at 4.2 K.

V. CONCLUSION

We present a systematic study of the impacts of the bulk Fermi potential, dopant freezeout, field-assisted dopant ionization, bandgap widening, and interface traps on the temperature dependence of the equilibrium threshold voltage (V_{T0}) in bulk silicon MOSFETs down to 4.2 K. The temperature dependence of the bulk Fermi potential results in an increase and saturation of V_{T0} down to cryogenic temperatures. On top of this saturation, interface traps close to the band edge (modeled by a Gaussian distribution) add an error

function model at cryogenic temperatures, which increases the saturation value of V_{T0}. The non-saturating behavior of V_{T0} that is observed in some pMOS measurements, can be modeled by an effective gate-oxide capacitance. Dopant freezeout, field-assisted dopant ionization, bandgap widening, and interface traps uniformly distributed over energy, do not change the qualitative behavior of V_{T0} over temperature.

REFERENCES

[1] E. A. Gutiérrez, M. J. Deen, and C. L. Claeys, *Low Temperature Electronics : Physics, Devices, Circuits, and Applications*. San Diego, CA, USA: Academic, 2001.

[2] F. Balestra and G. Ghibaudo, Eds., *Device and Circuit Cryogenic Operation for Low Temperature Electronics*. Boston, MA, USA: Springer, 2001.

[3] F. H. Gaensslen, V. L. Rideout, E. J. Walker, and J. J. Walker, "Very small MOSFET's for low-temperature operation," *IEEE Trans. Electron Devices*, vol. 24, no. 3, pp. 218–229, Mar. 1977.

[4] F. H. Gaensslen, R. C. Jaeger, and J. J. Walker, "Low temperature threshold behavior of depletion mode devices—Characterization and simulation," in *Proc. IEEE Int. Electron Devices Meeting*, Washington, DC, USA, 1977, pp. 520–524.

[5] R. C. Jaeger and F. H. Gaensslen, "Simple analytical models for the temperature dependent threshold behavior of depletion-mode devices," *IEEE J. Solid-State Circuits*, vol. 14, no. 2, pp. 423–429, Apr. 1979.

[6] R. M. Fox and R. C. Jaeger, "MOSFET behavior and circuit considerations for analog applications at 77 K," *IEEE Trans. Electron Devices*, vol. 34, no. 1, pp. 114–123, Jan. 1987.

[7] A. Akturk, J. Allnutt, Z. Dilli, N. Goldsman, and M. Peckerar, "Device modeling at cryogenic temperatures: Effects of incomplete ionization," *IEEE Trans. Electron Devices*, vol. 54, no. 11, pp. 2984–2990, Nov. 2007.

[8] E. Simoen, B. Dierickx, L. Warmerdam, J. Vermeiren, and C. Claeys, "Freeze-out effects on NMOS transistor characteristics at 4.2 K," *IEEE Trans. Electron Devices*, vol. 36, no. 6, pp. 1155–1161, Jun. 1989.

[9] H. Hanamura, M. Aoki, T. Masuhara, O. Minato, Y. Sakai, and T. Hayashida, "Operation of bulk CMOS devices at very low temperatures," *IEEE J. Solid-State Circuits*, vol. 21, no. 3, pp. 484–490, Jun. 1986. [Online]. Available: http://ieeexplore.ieee.org/document/1052555/

[10] S. Selberherr, "MOS device modeling at 77 K," *IEEE Trans. Electron Devices*, vol. 36, no. 8, pp. 1464–1474, Aug. 1989.

[11] P. P. Altermatt, A. Schenk, and G. Heiser, "A simulation model for the density of states and for incomplete ionization in crystalline silicon. I. Establishing the model in Si:P," *J. Appl. Phys.*, vol. 100, no. 11, 2006, Art. no. 113714.

[12] G. Xiao, J. Lee, J. Liou, and A. Ortiz-Conde, "Incomplete ionization in a semiconductor and its implications to device modeling," *Microelectron. Rel.*, vol. 39, no. 8, pp. 1299–1303, Aug. 1999.

[13] N. C. Dao, A. E. Kass, M. R. Azghadi, C. T. Jin, J. Scott, and P. H. Leong, "An enhanced MOSFET threshold voltage model for the 6–300 K temperature range," *Microelectron. Rel.*, vol. 69, pp. 36–39, Feb. 2017.

[14] S. R. Ekanayake, T. Lehmann, A. S. Dzurak, R. G. Clark, and A. Brawley, "Characterization of SOS-CMOS FETs at low temperatures for the design of integrated circuits for quantum bit control and readout," *IEEE Trans. Electron Devices*, vol. 57, no. 2, pp. 539–547, Feb. 2010.

[15] R. M. Incandela, L. Song, H. Homulle, E. Charbon, A. Vladimirescu, and F. Sebastiano, "Characterization and compact modeling of nanometer CMOS transistors at deep-cryogenic temperatures," *IEEE J. Electron Devices Soc.*, vol. 6, pp. 996–1006, 2018.

[16] D. P. Foty, "Impurity ionization in MOSFETs at very low temperatures," *Cryogenics*, vol. 30, no. 12, pp. 1056–1063, 1990.

[17] R. G. Pires, R. M. Dickstein, S. L. Titcomb, and R. L. Anderson, "Carrier freezeout in silicon," *Cryogenics*, vol. 30, no. 12, pp. 1064–1068, Dec. 1990.

[18] P. Bouillon and T. Skotnicki, "Theoretical analysis of kink effect in C-V characteristics of indium-implanted NMOS capacitors," *IEEE Electron Device Lett.*, vol. 19, no. 1, pp. 19–22, Jan. 1998.

[19] P. Bouillon, R. Gwoziecki, T. Skotnicki, J. Alieu, and P. Gentil, "Universal impurity ionization parameters in MIS C-V freeze-out characteristics and direct extraction of surface doping concentration," *IEEE Trans. Electron Devices*, vol. 47, no. 4, pp. 871–877, Apr. 2000.

[20] A. Pirovano, A. L. Lacaita, A. Pacelli, and A. Benvenuti, "Novel low-temperature C-V technique for MOS doping profile determination near the Si/SiO_2 interface," *IEEE Trans. Electron Devices*, vol. 48, no. 4, pp. 750–757, Apr. 2001.

[21] I. M. Hafez, G. Ghibaudo, and F. Balestra, "Assessment of interface state density in silicon metal-oxide-semiconductor transistors at room, liquid-nitrogen, and liquid-helium temperatures," *J. Appl. Phys.*, vol. 67, no. 4, pp. 1950–1952, Feb. 1990.

[22] M. Cassé, K. Tachi, S. Thiele, and T. Ernst, "Spectroscopic charge pumping in Si nanowire transistors with a high-*k*/metal gate," *Appl. Phys. Lett.*, vol. 96, no. 12, 2010, Art. no. 123506.

[23] A. Beckers, F. Jazaeri, and C. Enz, "Inflection phenomenon in cryogenic MOSFET behavior," *IEEE Trans. Electron Devices*, vol. 67, no. 3, pp. 1357–1360, Mar. 2020.

[24] A. Kamgar, "Subthreshold behavior of silicon MOSFETs at 4.2 K," *Solid-State Electron.*, vol. 25, no. 7, pp. 537–539, Jul. 1982.

[25] J. Michl *et al.*, "Quantum mechanical extension of defect model to explain BTI at cryogenic temperatures," in *Proc. Int. Rel. Phys. Symp. (IRPS)*, 2020.

[26] A. Ortiz-Conde, F. J. García-Sánchez, J. Muci, A. T. Barrios, J. J. Liou, and C.-S. Ho, "Revisiting MOSFET threshold voltage extraction methods," *Microelectron. Rel.*, vol. 53, no. 1, pp. 90–104, 2013.

[27] D. K. Schroder, *Semiconductor Material and Device Characterization*, Wiley, Hoboken, NJ, USA, 2015.

[28] A. Beckers, F. Jazaeri, and C. Enz, "Cryogenic MOSFET threshold voltage model," in *Proc. 49th Eur. Solid-State Device Res. Conf. (ESSDERC)*, Cracow, Poland, Sep. 2019, pp. 94–97.

[29] F. Balestra, L. Audaire, and C. Lucas, "Influence of substrate freeze-out on the characteristics of MOS transistors at very low temperatures," *Solid-State Electron.*, vol. 30, no. 3, pp. 321–327, 1987.

[30] R. M. Glidden, S. C. Lizotte, J. S. Cable, L. W. Mason, and C. Cao, "Optimization of cryogenic CMOS processes for sub-10 K applications," in *Proc. Infrared Readout Electron.*, Orlando, FL, USA, Jul. 1992, p. 2.

[31] K. J. Dwyer, J. M. Pomeroy, D. S. Simons, K. L. Steffens, and J. W. Lau, "Enriching ^{28}Si beyond 99.9998% for semiconductor quantum computing," *J. Phys. D, Appl. Phys.*, vol. 47, no. 34, Aug. 2014, Art. no. 345105.

[32] L. Vandersypen and A. V. Leeuwenhoek, "1.4 Quantum computing—The next challenge in circuit and system design," in *Proc. IEEE Int Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, USA, Feb. 2017, pp. 24–29.

[33] D. Sabbagh *et al.*, "Quantum transport properties of industrial ^{28}Si / ^{28}Si O_2," *Phys. Rev. Appl.*, vol. 12, no. 1, Jul. 2019, Art. no. 014013.

[34] A. Beckers, F. Jazaeri, and C. Enz, "Cryogenic MOS transistor model," *IEEE Trans. Electron Devices*, vol. 65, no. 9, pp. 3617–3625, Sep. 2018.

[35] P. A. 't Hart, M. Babaie, E. Charbon, A. Vladimirescu, and F. Sebastiano, "Characterization and modeling of mismatch in cryo-CMOS," *IEEE J. Electron Devices Soc.*, vol. 8, pp. 263–273, Feb. 2020.

[36] A. Grill *et al.*, "Reliability and variability of advanced CMOS devices at cryogenic temperatures," in *Proc. Int. Rel. Phys. Symp. (IRPS)*, 2020.

[37] F. Jazaeri, A. Beckers, A. Tajalli, and J. Sallese, "A review on quantum computing: From qubits to front-end electronics and cryogenic MOSFET physics," in *Proc. 26th Int. Conf. Mixed Design Integr. Circuits Syst. (MIXDES)*, Rzeszow, Poland, Jun. 2019, pp. 15–25.

[38] M. Waltl, *et al.*, "Nanoscale evidence for the superior reliability of SiGe high-k pMOSFETs," in *Proc. Int. Rel. Phys. Symp. (IRPS)*, Pasadena, CA, USA, Apr. 2016, pp. 1–6.

[39] A. Beckers, F. Jazaeri, A. Ruffino, C. Bruschini, A. Baschirotto, and C. Enz, "Cryogenic characterization of 28 nm bulk CMOS technology for quantum computing," in *Proc. 47th Eur. Solid-State Device Res. Conf. (ESSDERC)*, Leuven, Belgium, Sep. 2017, pp. 62–65.

[40] A. Beckers, F. Jazaeri, and C. Enz, "Characterization and modeling of 28-nm bulk CMOS technology down to 4.2 K," *IEEE J. Electron Devices Soc.*, vol. 6, pp. 1007–1018, Mar. 2018.

[41] Y. P. Varshni, "Temperature dependence of the energy gap in semiconductors," *Physica*, vol. 34, no. 1, pp. 149–154, Jan. 1967.

[42] N. D. Arora, *MOSFET Models for VLSI Circuit Simulation: Theory and Practice*. Vienna, Austria: Springer, 2012.

[43] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*. Hoboken, NJ, USA: Wiley, 2006.

[44] G. Paasch and S. Scheinert, "Charge carrier density of organics with Gaussian density of states: Analytical approximation for the Gauss–Fermi integral," *J. Appl. Phys.*, vol. 107, no. 10, 2010, Art. no. 104501.