Abstract—This paper presents a physics-based analytical model for the MOS transistor operating continuously from room temperature down to liquid-helium temperature (4.2 K) from depletion to strong inversion and in the linear and saturation regimes. The model is developed relying on the 1D Poisson equation and the drift-diffusion transport mechanism. The validity of the Maxwell-Boltzmann approximation is demonstrated in the limit to zero Kelvin as a result of dopant freeze-out in cryogenic equilibrium. Explicit MOS transistor expressions are then derived including incomplete dopant-ionization, bandgap widening, mobility reduction, and interface charge traps. The temperature-dependency of the interface trapping process explains the discrepancy between the measured value of the subthreshold swing and the thermal limit at deep-cryogenic temperatures. The accuracy of the developed model is validated by experimental results on a commercially available 28 nm bulk CMOS process. The proposed model provides the core expressions for the development of physically-accurate compact models dedicated to low-temperature CMOS circuit simulation.

Index Terms—cryogenic MOSFET, cryo-CMOS, freeze-out, incomplete ionization, interface traps, low temperature, MOS transistor, physical modeling

I. INTRODUCTION

ADVANCED CMOS processes perform increasingly well from room temperature down to deep-cryogenic temperatures (< 10 K) [1]–[4]. At these temperatures the ideal switch with a step-like subthreshold slope comes within reach [5]. Furthermore, cryo-electronics [6]–[8] can provide an interface with superconducting devices on the quest for exascale supercomputing [9]. Ultimately, quantum-engineered devices controlled by cryo-CMOS circuits can bring new functionality to existing computing technologies [10], [11].

Large-scale integration of silicon spin qubits [12], [13] and cryo-CMOS control circuits is envisioned to take solid-state quantum computing to the next level [14]. Digital, analog, and RF CMOS circuits [15]–[17] are then required to operate at millikelvin temperatures for initialization, manipulation, and read-out of the qubits, as well as error correction [18], [19]. Since the cooling power at millikelvin temperatures is reduced, the system could feature a cryogenic temperature gradient, where the control circuits operate at a higher cryogenic temperature than the qubits, e.g., 4.2 K [15]. However, the optimal design of power-hungry and thermal-noise dissipating circuits operating in close proximity to the qubits is yet to be explored. In this context, the main hurdle to overcome is the lack of compact MOS transistor models in circuit simulators, remaining physically accurate below 10 K [15], [17].

II. CRYO-MOS TRANSISTOR MODELING

The low-temperature circuits developed for spacecraft [20], [21], scientific equipment [22], ultra-low-noise detectors [23], cryobiology [24], and others, have been custom-designed relying on a semi-empirical approach. This approach requires laborious and expensive low-temperature measurements to extract model parameters for tuning room-temperature compact models to the target low-temperature [23], [25], [26]. Empirical temperature-scaling laws have been added to the room temperature physics-based MOS transistor model [27], [28] to capture cryogenic operation down to 4.2 K [29]–[31]. However, the discrepancy between the measured value of the subthreshold swing for a long device at 4.2 K (∼ 10 mV/decade) [3], [4], [32], and the theoretical thermal limit, $U_T \ln 10 \approx 0.8$ mV/decade) reveals that something more fundamental is missing. As we will demonstrate along this paper, important physical phenomena at low temperatures such as interface trapping [28], [33] and incomplete ionization [34], [35] have not been properly included to date. Furthermore, the intrinsic carrier concentration, $n_i$, takes on extremely small values below 10 K, causing arithmetic underflow in implemented analytical expressions or convergence problems in computer-aided-design simulations [36]–[38]. Therefore, standard references on semiconductor devices treat only the cryogenic equilibrium condition in bulk semiconductors above 10 K [27], [28], [39]. Analytical device-physics models, starting from the Poisson equation at low temperature, leave a gap unfilled between the zero-Kelvin approximation and 77 K [40]–[43].

In this work, we develop a MOS transistor model valid from room temperature (RT) down to deep-cryogenic temperatures, entirely based on physics principles and validated with experimental results. We start by verifying the continued validity of the Boltzmann statistics down to the deep-cryogenic regime.

III. MOS ELECTROSTATICS FROM RT TO 4.2 K

We model a long, planar $n$-channel MOS field-effect transistor in silicon, depicted in Fig. 1. Uniform operation across the width of the transistor is assumed and the gradual channel approximation is adopted. The electrostatics can then be described by the 1D Poisson equation [27], [28].
A. Poisson-Fermi equation

Merging the 1D Poisson equation with the mobile carrier concentrations, \( n \) and \( p \), given by Fermi-Dirac statistics, gives
\[
\frac{\partial^2 \psi(y)}{\partial y^2} = -\frac{q}{\varepsilon_{si}} \left( -n + p - N_A^+ \right),
\]

where \( q \) is the elementary charge, \( \varepsilon_{si} \) the silicon permittivity, and \( \psi \) the potential, with \( E_F \) the Fermi-level and \( E_i \) the intrinsic energy level. The first term on the RHS of equation (1) represents the electron contribution, \( n \), the second term the hole contribution, \( p \), and the third term the ionized dopant contribution, \( N_A^- \).

1) Incompletely-ionized dopants: Under thermal equilibrium, both at room and cryogenic temperatures, the majority carrier concentration can defer from the implanted doping value, \( N_A^- \), due to incomplete ionization of the dopants. In cryogenic equilibrium, incomplete ionization is strong and known as freeze-out, since thermal dopant-ionization is very low [35]. However, during MOS operation, also field-assisted ionization comes into play. Fermi-Dirac statistics provides a low [35]. During MOS operation, also field-assisted ionization [35] and temperature (thermal ionization).

\[
N_A^- = \frac{N_A}{1 + g_A e^{-E_F,n/V_{ch}}} = \frac{N_A}{1 + g_A e^{-(E_A - E_i)/q}},
\]

where the electron quasi-Fermi-level is given by \( E_F,n = E_F - qV_{ch} \). The RHS of (2) is obtained by replacing \( E_A - E_F,n \) with \( E_A - E_i + E_i - E_F,n \) in the exponential term, and by defining an acceptor potential, \( \psi_A \equiv (E_A - E_i)/q \), as indicated in Fig. 1. The channel voltage, \( V_{ch} \), denotes the shift of the quasi-Fermi-potential due to the drain-to-source voltage, \( V_{DS} \). The second expression in (2) highlights the two dopant-ionization contributions, i.e. the potential (field-assisted ionization [35]) and temperature (thermal ionization).

2) Mobile carrier concentrations: Since \( n \) and \( p \) given by Fermi-Dirac statistics in (1) require numerical integration over energy, this inhibits explicit solutions for the charge densities and current in the MOS transistor. Expressing \( n \) and \( p \) using Boltzmann statistics allows to obtain such relations. However, the validity of the Maxwell-Boltzmann approximation down to deep-cryogenic temperatures is questionable. It has been reported [38], [42] that semiconductors become strongly degenerate at deep-cryogenic temperatures, preventing its use. This is however inconsistent with the zero-Kelvin limits of the Fermi-level position in the bandgap derived by Pierret [39]. Therefore, in the next subsection we aim to verify the Maxwell-Boltzmann approximation down to deep-cryogenic temperatures.

3) Verification of Boltzmann statistics: We numerically calculate the position of the equilibrium Fermi-level, \( E_F \), down to 100 mK relying on Fermi-Dirac statistics in an extrinsic bulk semiconductor, e.g. p-type silicon. In this case, the Poisson equation imposes the charge neutrality, \( p_p = N_A^- \), where \( p_p \) is expressed by Fermi-Dirac statistics [28], [39] and \( N_A^- \) by (2). This yields an implicit equation for \( E_F \), which is solved numerically at each temperature and doping value using an extension of the arithmetic precision. As illustrated in Fig. 2, below 120 K, \( E_F \) remains off the valence band edge with an offset smaller than \( 3kT \) for doping values below the degenerate limit (i.e. \( N_A^- = 4 \times 10^{18} \) cm\(^{-3} \) in Si:B) [28], [39], [45]. Note that this is predicted correctly only when incomplete ionization is taken into account. Complete ionization \( (g_A = 0) \) would predict an offset smaller than \( 3kT \) for \( N_A^- = 10^{18} \) cm\(^{-3} \), and hence a degenerate semiconductor. It should therefore be emphasized that incomplete ionization maintains the non-degeneracy of a highly-doped semiconductor at temperatures down to 100 mK. Furthermore, near zero Kelvin, \( E_F \) tends to saturate at \( (E_A - E_v)/2 \) for all considered doping values. This corresponds to the zero-Kelvin limit in Pierret [39] assuming Boltzmann statistics. Using the now validated Maxwell-Boltzmann description for \( p_p \), i.e. \( N_v \exp[(E_v - E_F)/kT] \), in \( p_p = N_A^- \), leads to a quadratic equation in \( \exp[(E_v - E_F)/kT] \) with as solution,
\[
E_F - E_v = kT \ln N_v \frac{N_A^-}{N_A} + kT \ln \frac{1 + \sqrt{1 + 4g_A N_A^- e^{(E_A - E_i)/q}}}{2}.
\]

Considering the temperature dependency of \( N_v \) [28], [39], while taking the limit of (2) to 0 K, leads to the limit as \( T \rightarrow 0 \) of \( E_F = E_v + (E_A - E_v)/2 \).

Performing the same numerical \( E_F \)-calculation for an intrinsic semiconductor, the extremely small value of \( n_i \) can be verified relying on Fermi-Dirac statistics. The Poisson equa-
tion then imposes the charge neutrality, \( n = p = n_i \), where \( n \) and \( p \) are given by Fermi-Dirac statistics. As illustrated in Fig. 2, this yields \( n_i \)-values lying outside the range of IEEE double-precision arithmetic \((10^{−308}−1)\times10^{308}\), e.g., at 4.2 K, \( \approx 10^{−678} \text{ cm}^{-3} \). Therefore, an extension of the arithmetical precision will also be used in the remainder of this work based on Boltzmann statistics, since the carrier concentrations are then expressed through \( n_i \).

### B. Poisson-Boltzmann equation

Using the Maxwell-Boltzmann approximation of \( n \) and \( p \), validated down to deep-cryogenic temperatures in the previous section, we combine the 1D Poisson equation with Boltzmann statistics, which leads to

\[
\frac{\partial^2 \psi(y)}{\partial y^2} = -\frac{q}{\varepsilon_{\text{si}}} \left( -n_i e^{-\frac{\psi - \psi_{\text{ch}}}{kT}} + n_i e^{-\frac{\psi}{kT}} - N_A \right), \tag{4}
\]

where \( U_T \triangleq kT/q \) is the thermal voltage. The first term on the RHS of equation (4) represents the electron contribution, \( n \), and the second term the hole contribution, \( p \). The intrinsic carrier concentration is given by \( n_i = \sqrt{N_cN_v} \exp(-E_g/2kT) \), with \( E_g \) the bandgap, and \( N_c \) and \( N_v \) the effective density-of-states in the conduction band and valence band respectively. The temperature-dependency of \( E_g \) as described by Varshni [44] is used. The extremely small, but finite value of \( n_i \) at deep-cryogenic temperatures cannot be assumed zero—which would be equivalent to the zero-Kelvin approximation [40] or considering \( f(E) \) as a step function—since this leads to zero mobile carrier concentrations independently of the potential. This is irreconcilable with the observed field-effect and correct functioning of the MOS transistor at 4.2 K [3, 4]. For \( U_T \) small, the exponential factor has a very big dynamic range when \( \psi \) changes during MOS transistor operation, large enough to overrule \( n_i \) in the multiplication.

1) Derivation of the electric field at the surface: Introducing \( 2N_A \) in Eq. (4), and then multiplying (4) on both sides with \( 2(\partial^2\psi/\partial y) \) gives

\[
\frac{\partial}{\partial y} \left[ \frac{\partial \psi(y)}{\partial y} \right] = \frac{2q}{\varepsilon_{\text{si}}} n_i e^{\frac{\psi - \psi_{\text{ch}}}{kT}} - n_i e^{\frac{\psi}{kT}} + \frac{N_A}{1 + g_Ae^{\frac{\psi - \psi_{\text{ch}}}{kT}}} \psi \tag{5}
\]

Integrating (5) from bulk to surface with \( E = -\partial^2\psi/\partial y \) and \( E_b = 0 \) yields

\[
E_s^2 = \frac{2q}{\varepsilon_{\text{si}}} \int_{\psi_b}^{\psi_s} \left( n_i e^{\frac{\psi - \psi_{\text{ch}}}{kT}} - n_i e^{\frac{\psi}{kT}} + \frac{N_A}{1 + g_Ae^{\frac{\psi - \psi_{\text{ch}}}{kT}}} \right) d\psi. \tag{6}
\]

In (6) the additional potential dependence due to field-assisted ionization of the dopants can be straightforwardly integrated as well, i.e., by replacing \( N_A \) with \( N_A \{1 + g_A \exp[(\psi_A - (\psi - \psi_{\text{ch}}))/U_T] - g_A \exp[(\psi_A - (\psi - \psi_{\text{ch}}))/U_T] \} \) in the numerator of the third term and splitting the resulting integral. This gives an expression for the square of the electric field at the surface,

\[
E_s^2 = \frac{2qn_i U_T}{\varepsilon_{\text{si}}} \left( e^{\frac{\psi - \psi_{\text{ch}}}{kT}} - e^{\frac{\psi - \psi_{\text{ch}}}{kT}} + e^{\frac{\psi}{kT}} - e^{\frac{\psi}{kT}} \right) + 2\frac{qN_A}{\varepsilon_{\text{si}}} \left[ \psi_s - \psi_b - U_T \ln f_s(E_A) \right] \tag{7}
\]

where \( \psi_b \triangleq (E_{F,b} - E_i)/q \) is the bulk potential and \( \psi_s \triangleq (E_{F,s} - E_i)/q \) the surface potential, as indicated in Fig. 7. \( E_{F,s} \) denotes the Fermi-level at the surface, and \( E_{F,b} \) the Fermi-level in the bulk. The logarithmic term in (7) is the contribution
of incomplete ionization, where
\[
f_s(E_A) \triangleq \frac{1}{1 + g_A e^{E_p - E_F}} = \frac{1}{1 + g_A e^{\psi_s - \psi_{ch}}} \tag{8}\]
the Fermi-Dirac ionization probability at the surface, and
\[
f_b(E_A) \triangleq \frac{1}{1 + g_A e^{E_p - E_F}} = \frac{1}{1 + g_A e^{\psi_s - \psi_b}} \tag{9}\]
the Fermi-Dirac ionization probability in the bulk, assuming that \(V_{ch}\) is zero in the bulk. Both ionization probabilities are qualitatively shown in Fig. 1. If complete ionization is assumed, then \(f_s(E_A) = f_b(E_A) = 1\) and the incomplete ionization term cancels in (7), leading to the expression widely-used at RT [27], [28]. The surface-ionization probability \(f_s(E_A)\) is plotted in Fig. 3 as a function of thermal and field-assisted ionization. Immediately evident is that freeze-out at the surface (arbitrarily defined when \(f_s(E_A) < 0.2\)) is only present when the temperature is below \(\approx 50\) K and the potential is close to the flatband condition \((\psi_s \approx \psi_b)\). Above \(\psi_s\), the ionization probability rapidly transitions to one due to field-assisted ionization. This transition corresponds to the bending of \(E_A\) under \(E_F\) at the surface in Fig. 1. Therefore, complete ionization is a valid approximation even at deep-cryogenic temperatures, although the shift in \(E_F\) due to incomplete ionization (Fig. 2) should be taken into account since it affects the threshold voltage. This \(E_F\)-shift can be quantified by using \(f_b(E_A)\) from (9) in the bulk charge neutrality condition, \(p_p = N_A\), which leads to the quadratic equation \(\exp(2\psi_b/U_T) - (n_i/N_A)\exp(\psi_b/U_T) - (g_A/N_A)\exp(\psi_b/U_T)\) with as solution,
\[
\psi_b = U_T \ln \left( \frac{n_i}{N_A} + U_T \ln \left( 1 + \frac{1 + \frac{4N_A g_A e^{\psi_s}}{2 n_i}}{2} \right) \right) \tag{10}\]
The second term in (10) is the shift of \(E_F\) by including incomplete ionization, which is only dependent on temperature and doping. Assuming complete ionization, i.e. \(g_A = 0\), the well-known \(U_T \ln(n_i/N_A)\) is obtained.

2) Derivation of the charge densities: Applying the Gaussian law over the semiconductor body in Fig. 1, the total semiconductor charge density per unit area, \(Q_{sc}\), is obtained by \(Q_{sc} = -\varepsilon_{si} E_s\), with \(E_s\) given by (7). The obtained \(Q_{sc}\) is plotted in Fig. 4 at RT, 77 K, and 4.2 K. For 77 K and 4.2 K, small kinks are noticeable close to \(\psi_s\) due to the transition from incomplete to complete ionization when \(E_A\) bends under \(E_F\) at the surface (\(E_{F,s}\)), or equivalently, \(\psi_s\) becomes less negative than \(\psi_A\). Above this transition, \(f_s(E_A) \approx 1\) according to (2). At RT, \(E_F\) lies above \(E_A\) in the flatband condition (see Fig. 2) and hence no transitional kink is noticeable. There is however a \(\psi_{ch}\)-shift also at RT due to incomplete ionization according to (11). Note that for complete ionization (dashed lines) no kinks are observed since the logarithmic term cancels in (7). Assuming the charge-sheet and fully-depletion approximations [27], the fixed charge density per unit area, \(Q_f\), is given by
\[
Q_f = -\varepsilon_{si} \sqrt{\frac{2qN_A}{\varepsilon_{si}} (\psi_s - \psi_{ch})} - \frac{2qN_A U_T}{\varepsilon_{si}} \ln \left( \frac{f_s(E_A)}{f_b(E_A)} \right) \tag{11}\]
Relying on the charge neutrality, the mobile charge density per unit area, \(Q_m\), can be obtained from \(Q_{sc} = Q_{sc} - Q_f\), resulting in \(Q_m\). As can be observed in this figure, incomplete ionization does not affect the turn-on rate of \(Q_m\), but contributes a small decrease in the charge threshold voltage, due to the shift of the \(E_F\)-position closer towards the conduction-band edge as shown in Fig. 2, and derived in (10).

Therefore, from this section we conclude that incomplete ionization cannot explain the offset between the measured subthreshold swing at 4.2 K and the thermal limit. As we will show in the next section, the temperature-dependent occupation of interface charge traps can degrade the subthreshold swing down to 4.2 K.

3) Interface charge traps: Defects and lattice breaking at the oxide-semiconductor interface introduce trap energy levels, \(E_t,\) in the bandgap which degrade the control of the gate-to-bulk voltage, \(V_{GB},\) over the channel. In what follows, the Fermi-Dirac occupation of interface traps, \(f(E_t)\), is included in the surface-boundary condition and the effect on the \(Q_m\) turn-on rate is analyzed at 4.2 K. The surface-boundary condition, i.e. the link between \(V_{GB}\) and \(\psi_s\), is given by \(V_{GB} = V_{FB} + \varepsilon_s E_s/C_{ox} + (\psi_s - \psi_{ch})\) where \(C_{ox}\) is the oxide capacitance per unit area, and \(V_{FB}\) is the flatband voltage, given by \(V_{FB} \equiv \phi_{ms} - Q_{it}/C_{ox}\) [27], [28]. Here \(Q_{it}\) is the interface trap charge density per unit area. We consider the summation of the discrete acceptor trap energy levels (all donor states are occupied and neutral during turn-on in \(n\)MOS [28]). Each discrete trap-energy-level, \(E_{t,j},\) at position \(j\) in the bandgap has its particular \(N_{it,j}\)-value assigned to it, where \(N_{it}\) is the density-of-interface-traps per unit area. \(Q_{it}\) can then be expressed as \(Q_{it} = -q \sum_{j} N_{it,j} f_s(E_{t,j})\) where \(N_{it}\) is the number of interface traps, and
\[
f_s(E_{t,j}) = \frac{1}{1 + g_A e^{E_{t,j} - E_F}} = \frac{1}{1 + g_A e^{\psi_s - \psi_{ch}}} \tag{13}\]
where:

\[ Q_m = -\varepsilon_{ai} \left[ \frac{2q_n U_T}{\varepsilon_{ai}} \left( e^{\frac{\psi_s - \psi_b}{U_T}} - e^{\frac{-\psi_b}{U_T}} \right) + \frac{2qN_A}{\varepsilon_{ai}} \left( \psi_s - \psi_b - U_T \ln \frac{f_s(E_A)}{f_b(E_A)} \right) \right] + \varepsilon_{ai} \left[ \frac{2qN_A}{\varepsilon_{ai}} \left( \psi_s - \psi_b - U_T \ln \frac{f_s(E_A)}{f_b(E_A)} \right) \right] \]

is the Fermi-Dirac occupation probability of the trap energy level \( E_{t,j} \). The RHS of Eq. (13) is obtained by defining the trap potentials, \( \psi_{t,j} \triangleq (E_{t,j} - E_i)/q \) [46]–[48]. This leads to the flatband voltage

\[ V_{FB} = \phi_{ms} + \frac{q}{C_{ox}} \sum_j N_{t,j} \left( 1 + g_t \exp \left[ \left( \psi_{t,j} - (\psi_s - V_{ch}) \right)/U_T \right] \right) \]

(14)

Plotting \( Q_m \) from (12) versus \( V_{GB} \) at 4.2 K in Fig. 6, including four interface traps close to the conduction band, reveals how each interface trap degrades the turn-on of \( Q_m \) separately, as well as the combined effect of the sum of the interface traps.

IV. CURRENT DERIVATION

To derive the current in the linear regime, this core-model assumes drift-diffusion transport, and does not include ballistic nor quantum transport. To verify the drift-diffusion transport mechanism at cryogenic temperatures, the proposed model for the drain-to-source current will be experimentally validated in Section V. Neglecting the hole-contribution to the current, the expression for the total drain-source current is given by

\[ I_{DS} = -\mu_n (W/L) \int_{V_{GB}} V_m(V_{ch})dV_{ch}, \]

where the electron mobility \( \mu_n \) is assumed constant along the channel, and \( W/L \) is the device aspect-ratio, as indicated in Fig. 1. In the linear regime, \( Q_m \) can be assumed independent of \( V_{ch} \). In this case, the total drain-source current is given by \( I_{DS} = -\mu_n (W/L) Q_m V_{DS} \). In saturation, the integral over \( V_{ch} \) cannot be readily solved.

Therefore, starting from the drift-diffusion equation gives

\[ I_{DS} = -\frac{W}{L} \int_{\psi_{s,S}}^{\psi_{s,D}} \mu_n Q_m d\psi + W \int_{\psi_{M,S}}^{\psi_{M,D}} \mu_n U_T dQ_m. \]

(15)

Assuming a linearization of the mobile charge density with respect to the surface potential at constant gate voltage [49] in [15], i.e., \( Q_m = mC_{ox} (\psi_s - \psi_P) \), with \( m \triangleq \partial Q_m(C_{ox})/\partial \psi_s \) and \( \psi_P \) the pinch-off potential, and integrating, results in an expression for the total drain-source current in saturation,

\[ I_{DS} = \frac{W}{L} \mu_n \left[ -\frac{Q_{m,D}^2 - Q_{m,S}^2}{2mC_{ox}} + U_T (Q_{m,D} - Q_{m,S}) \right]. \]

(16)

\( Q_{m,S} \) and \( Q_{m,D} \) are obtained from (12), setting \( V_{ch} \) to zero and \( V_{DS} \) respectively. At cryogenic temperature an improvement in the low-field mobility, \( \mu_0 \), is observed due to a reduction of the phonon scattering [1], [6]. In addition, the mobility reduces at higher gate voltage due to surface roughness scattering at high vertical electric field [1]. This mobility reduction can be modeled by \( \mu_n = \mu_0/(1 + \theta V_{GB}) \) where \( \theta \) is the mobility reduction factor.

V. EXPERIMENTAL RESULTS AND DISCUSSION

Room temperature and cryogenic measurements were performed on devices fabricated in a 28-nm bulk CMOS process. The full set of measurements, measurement set-up, and characterization were previously reported in [3], [4]. After measuring at RT, the samples were immersed into liquid helium (4.2 K) and liquid nitrogen (77 K) baths with a dipstick. Fig. 7a favorably compares the model with the linear transfer characteristics (\( V_{DB} = 20 \text{ mV} \)) measured at RT and 4.2 K on a long \( n \text{MOS} \) device with \( W/L = 3 \mu m/1 \mu m \), in linear and logarithmic scales. The extracted \( \mu_0 \)-values from the model are in accordance with the characterization performed in [3]. Furthermore Fig. 7b analyzes the effect of incomplete ionization, interface traps, and mobility, on the current at 4.2 K. Note that incomplete ionization reduces the threshold voltage, and interface traps can degrade the subthreshold swing (SS) to \( \approx 10 \text{ mV/decade} \). The strong increase in the mobility increases the on-state current at 4.2 K. Fig. 8 validates the model for the current in saturation (\( |V_{DB}| = 0.9 \text{ V} \)) using the measurements performed at RT, 77, and 4.2 K on a long \( p \text{MOS} \) device with \( W/L = 3 \mu m/1 \mu m \), in linear and logarithmic scales. The metal-semiconductor work function difference, \( \phi_{ms} \), increases in absolute value at lower temperatures according to the change in \( E_F \)-position (Fig. 2).

VI. SUBTHRESHOLD-SWING DERIVATION

In this section a SS-expression including incomplete ionization and temperature-dependent interface trapping is derived. Incomplete ionization is included to prove the minimal influence on SS shown in Fig. 7b. The temperature-dependency
of interface-trap occupation, \( f_s(E_i) \), allows to obtain the \( \Delta_{SS} \)-offset of \( \approx 10 \text{ mV/dec} \) above the thermal limit, \( U_T \ln 10 \), previously observed on long-channel devices [3], [4].

The subthreshold swing, \( SS \), is usually expressed as \( nU_T \ln 10 \), where the non-ideality factor or slope factor, \( n \), is given by \( (\partial V_G/\partial \psi_s) \), describing the deviation from the thermal limit. Assuming \( f_s(E_i) \) from [13] to be one, \( (\partial V_G/\partial \psi_s) \) yields \( 1 + \sqrt{2}Q_{it}\varepsilon_{si}/[C_{ox}(2\sqrt{\varepsilon_s - \psi_b}) + qN_{it}/C_{ox}] [27], [50] \). However, at 4.2 K, and assuming the highest possible doping value below the degenerate limit, a large \( N_{it} \)-value in the order of \( 10^{13} \text{ cm}^{-2} \) is extracted to accommodate for a \( SS \) of \( \approx 10 \text{ mV/dec} \) [33], [51], since \( N_{it} \) becomes multiplied with \( U_T \) in this expression. However, it should be emphasized that in the used expression for \( SS \), the temperature dependency of interface-trap occupation is not taken into account. Relying on drift-diffusion transport in the linear regime and assuming \( \mu \) independent of \( V_G \), the subthreshold slope, \( SS^{-1} \), is given by

\[
SS^{-1} = \frac{1}{\ln 10} \frac{1}{Q_m} \frac{\partial Q_m}{\partial \psi_s} \frac{\partial \psi_s}{\partial V_G}.
\]

The factor \( (1/Q_m)(\partial Q_m/\partial \psi_s) \) is found from [12] by considering \( Q_m \ll Q_f \) or \( Q_{ac} \approx Q_f \) in the subthreshold region. After some mathematical manipulation, we find

\[
\frac{1}{Q_m} \frac{\partial Q_m}{\partial \psi_s} = q\varepsilon_{si} \left[ \frac{1}{Q_m} \frac{\partial \psi_s}{\partial V_G} + \frac{1}{Q_f} \frac{\partial \psi_s}{\partial V_G} \right] - \frac{1}{Q_T^2} \frac{N_A}{q} \left( 1 - \frac{U_T}{f(E_A)} \frac{\partial f(E_A)}{\partial \psi_s} \right).
\]

Merging [11] and [18], and inverting, gives

\[
SS = 2N_A \frac{\partial V_G}{Q_m n_s e^{-\psi_b}} \left( \frac{\psi_s - \psi_b}{U_T} - \ln \frac{f(E_A)}{f_s(E_A)} \right) \frac{\partial \psi_s}{\partial V_G}.
\]

The following relation can be derived for (a) in the above equation (see Appendix),

\[
n_s e^{-\psi_b} = \frac{Q_m}{Q_f} 2N_A \left( \frac{\psi_s - \psi_b}{U_T} - \ln \frac{f(E_A)}{f_s(E_A)} \right).
\]
where

\[
qN_{it,j} \frac{\partial f_s(E_{t,j})}{\partial \psi_s} \]

for \( i \) from 1 to \( N_{it} \), at each \( \psi_s \) value in the subthreshold region, and assuming a uniform \( N_{it,j} \) value for each trap, leads to

\[
\frac{\partial V_{GB}}{\partial \psi_s} = 1 + \sqrt{2qN_{it} \varepsilon_{si}} \frac{1}{C_{ox}} 2\sqrt{\psi_s - \psi_b} + \frac{qN_{it}}{C_{ox}} \frac{1}{U_T(1 + g_t)^2}.
\]

The first two terms in (25) yield the non-ideality or slope factor without interface traps, \( n_0 \). The \( SS \)-expression becomes

\[
SS = n_0 U_T \ln 10 + \frac{qN_{it}}{C_{ox}} \frac{g_t}{(1 + g_t)^2} \ln 10
\]

where the second term on the RHS is the sought \( \Delta SS \)-offset. The non-ideality factor \( n_0 \) has an upper bound of 2 mainly related to doping. Therefore, at 4.2 K, the first term on the RHS of (26) is limited by 1.6 mV/decade. Note that \( N_{it} \) does not become multiplied with \( U_T \) in (26). Therefore, assuming a reasonable value for \( N_{it} = 3 \times 10^{11} \text{cm}^{-2} \), the second term gives \( \approx 9 \text{mV/decade} \) (with \( C_{ox} = 20 \text{mF m}^{-2} \) and \( g_t = 4 \)).
Together they yield the 5.5-degradation observed on a long nMOS device at 4.2 K. At 77 K, a similar calculation using \( n_0 = 1.08 \) and \( N_{it} = 3 \times 10^{10} \text{cm}^{-2} \) gives 25 mV/dec, corresponding to the subthreshold swing measured \([3]\) on a long pMOS device at 77 K (Fig. 8).

**VII. CONCLUSION**

A theoretical MOS transistor model is developed valid from room temperature down to liquid-helium temperature. The model relies on Boltzmann statistics, verified in the limit to zero Kelvin, and includes incomplete ionization, interface traps, bandgap temperature dependency, and mobility reduction. It is evidenced that incomplete ionization maintains the non-degeneracy of a semiconductor at deep-cryogenic temperatures, and leads to a decrease in the threshold voltage on top of the overall increase due to Fermi-Dirac distribution scaling. The Fermi-Dirac temperature-dependency of interface-trap occupation degrades the subthreshold swing down to 4.2 K. An expression for the subthreshold swing including incomplete ionization and temperature-dependent interface trapping is derived. The proposed model builds the indispensable physical foundation for future low-temperature CMOS circuit design.

**APPENDIX**

Starting from \( Q_m + Q_f = Q_{sc} \), we can write \( (Q_m + Q_f)^2 = \varepsilon_e^2 E_s^2 \), with \( E_s \) given by \([7]\). Solving a quadratic equation for \( Q_m \) leads to

\[
\frac{Q_m}{Q_f} = -1 + \sqrt{1 + \frac{2 \varepsilon_e U_T \varepsilon_s}{Q_f^2} e^{-\frac{\varphi_b - \varphi_T}{kT}}} \tag{27}
\]

where we neglected the exponential term in \( \varphi_b/U_T \). In the subthreshold region \( Q_m \ll Q_f \), \( \sqrt{1 + \frac{1}{x}} \) can be approximated by \( 1 + x/2 \) for \( x \to 0 \). Using \([11]\) for \( Q_f^2 \) leads to \([20]\).

**REFERENCES**

[1] F. Balestra and G. Ghiaudo, “Physics and performance of nanoscale semiconductors at cryogenic temperatures,” *Semiconductor Science and Technology*, vol. 32, no. 2, p. 025002, Feb. 2017.

[2] M. de Souza, V. Khkhtyska, D. Flandre, and M. A. Pavanello, “Liquid helium temperature analog operation of asymmetric self-cascode FD SOI MOSFETs,” in 2012 IEEE International SOI Conference (SOI), Oct 2012, pp. 1–2.

[3] A. Beckers, F. Jazaeri, A. Ruffino, C. Bruschini, A. Baschirotto, and A. Morello, “A single-atm electron spin qubit in silicon,” *Nature*, vol. 489, no. 7417, pp. 541–545, Jul. 2012.

[4] A. Maurand, X. Jehl, D. Kotek-Patil, A. Corna, H. Bohuslavskyi, R. Lavielle, L. Hutin, S. Barraud, M. Vinet, M. Sanquer, et al., “A CMOS silicon spin qubit,” *Nature Communications*, vol. 7, 2016.

[5] L. M. K. Vandersypen, H. Bihlm, J. S. Clarke, A. S. Dzurak, A. Ishihara, A. Morello, D. J. Reilly, L. R. Schreiber, and M. Veldhorst, “Interfacing spin qubits in quantum dots and donors: hot, dense, and coherent,” *npj Quantum Information*, vol. 3, no. 1, p. 34, Sep. 2017.

[6] 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 Transactions on Electron Devices*, vol. 57, no. 2, pp. 539–547, Feb. 2010.

[7] D. J. Reilly, “Engineering the quantum-classical interface of solid-state qubits,” *npj Quantum Information*, vol. 1, p. 15011, Oct. 2015.

[8] M. Rahman and T. Lehmann, “A cryogenic DAC operating down to 4.2 K,” *Cryogenics*, vol. 57, pp. 47–55, Apr. 2016.

[9] D. P. D’Inverno, “The physical implementation of quantum computation,” arXiv preprint quant-ph/0002077, 2000.

[10] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[11] V. Revèret, X. De La Broise, C. Fermon, M. Pannetier-Lecoeur, C. Pigot, L. Rodriguez, J.-L. Sauvageot, Y. Jin, S. Marnieros, D. Bouchet et al., “CESAR: Cryogenic electronics for space applications,” *Journal of Low Temperature Physics*, vol. 176, no. 3-4, pp. 446–452, 2014.

[12] P. Merken, T. Souverijns, J. Putzeys, Y. Creten, and C. V. Hoof, “Flight Qualification and Circuit Development of Sensor Front-End Electronics for PACS/Ros体现 Liquid Helium Temperature,” *Journal of Microelectronics and Electronics Packaging*, vol. 4, no. 4, 2007.

[13] J. R. Hoff, G. W. Deptuch, G. Wu, and P. Gui, “Cryogenic Lifetime Studies of 130 nm and 65 nm MOS Transistors for High-Energy Physics Experiments,” *IEEE Transactions on Nuclear Science*, vol. 62, no. 3, pp. 1255–1261, Jun. 2015.

[14] B. Okcan, P. Merken, G. Gielen, and C. Van Hoof, “A cryogenic analog to digital converter operating from 300 K down to 4.4 K,” *Review of Scientific Instruments*, vol. 81, no. 2, p. 024702, Feb. 2010.

[15] F. R. Ihmig, S. G. Shirley, R. K. Kirschman, and H. Zimmermann, “Frozen Cells and Bits: Cryoelectronics Advances Biopreservation,” *IEEE Pulse*, vol. 4, no. 5, pp. 35–43, Sep. 2013.

[16] P. Martin, M. Cavaleri, R. Fascio, G. Ghiaudo, and M. Bucher, “EKV3 compact modeling of MOS transistors from a 0.18um CMOS technology for mixed analog/digital circuit design at low temperature,” *Cryogenics*, vol. 49, no. 11, pp. 595–598, 2009.

[17] Y. Creten, P. Merken, W. Sansen, R. Mertens, and C. Van Hoof, “A cryogenic analog ADC operating down to 4.2 K,” in 2007 IEEE International Solid-State Circuits Conference. Digest of Technical Papers, Feb 2007, 468–469.

[18] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*. Oxford Univ. Press, 2011.

[19] S. M. Sze and K. K. Ng, *Physics of semiconductor devices*. John Wiley & Sons, 2006.

[20] A. Akturk, M. Holloway, S. Potbhare, D. Gundlach, B. Li, N. Goldsman, A. Beckers, F. Jazaeri, H. Bohuslavskyi, L. Hutin, S. D. Franceschi, and C. Enz, “Design-oriented modeling of 28 nm FDSOI CMOS technology down to 4.2 K for quantum computing,” in 2018 Joint...
International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS), March 2018, pp. 1–4.

[33] 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,” Journal of Applied Physics, vol. 67, no. 4, pp. 1950–1952, Feb. 1990.

[34] A. K. Jonscher, “Semiconductors at cryogenic temperatures,” Proceedings of the IEEE, vol. 52, no. 10, pp. 1092–1104, Oct 1964.

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

[36] A. K. Jonscher, “Semiconductors at cryogenic temperatures,” Proceedings of the IEEE, vol. 52, no. 10, pp. 1092–1104, Oct 1964.

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

[38] A. K. Jonscher, “Semiconductors at cryogenic temperatures,” Proceedings of the IEEE, vol. 52, no. 10, pp. 1092–1104, Oct 1964.

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

[40] A. K. Jonscher, “Semiconductors at cryogenic temperatures,” Proceedings of the IEEE, vol. 52, no. 10, pp. 1092–1104, Oct 1964.

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