Characterization and Compact Modeling of Nanometer CMOS Transistors at Deep-Cryogenic Temperatures

Incandela, Rosario; Song, L; Homulle, Harald; Charbon, Edoardo; Vladimirescu, Andrei; Sebastiano, Fabio

DOI
10.1109/JEDS.2018.2821763

Publication date
2018

Document Version
Publisher's PDF, also known as Version of record

Published in
IEEE Journal of the Electron Devices Society

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.
Characterization and Compact Modeling of Nanometer CMOS Transistors at Deep-Cryogenic Temperatures

Rosario M. Incandela, Student, IEEE, Lin Song, Harald Homulle, Edoardo Charbon, Fellow, IEEE, Andrei Vladimirescu, Fellow, IEEE, and Fabio Sebastiano, Senior Member, IEEE

Abstract—Cryogenic characterization and modeling of two nanometer bulk CMOS technologies (0.16-μm and 40-nm) are presented in this paper. Several devices from both technologies were extensively characterized at temperatures of 4 K and below. Based on a detailed understanding of the device physics at deep-cryogenic temperatures, a compact model based on MOS11 and PSP was developed. In addition to reproducing the device DC characteristics, the accuracy and validity of the compact models are demonstrated by comparing time- and frequency-domain simulations of complex circuits, such as a ring oscillator and a low-noise amplifier (LNA), with the measurements at 4 K.

Index Terms—Cryogenic electronics, CMOS, cryogenic, cryo-CMOS, characterization, modeling, kink, 4 K, LNA.

I. INTRODUCTION

CRYOGENIC electronics plays a fundamental role in several applications, such as spacecraft, high-energy physics experiments, metrology, superconductive astronomical detectors and, with the increased interest in quantum computing, the manipulation of quantum bits (qubits) [1]–[5]. Most qubits are placed in the coldest chamber of dilution refrigerators, where temperatures reach values in the range of 10-100 mK, in order to expose their quantum behavior and extend the lifetime of their quantum state. State-of-the-art quantum processors typically consist of a few qubits that are controlled and read out by general-purpose electronics operating at room-temperature [6]–[8]. Only a few components, such as the first amplification stages, operate at cryogenic temperatures (1-4 K) in order to reduce their noise level [9], [10]. Nonetheless, future quantum computers would require millions of qubits to run any algorithm with practical applications, such as the simulation of a complex molecule [11]. Thus, an enormous amount of room-temperature electronics and, consequently, an unpractical number of wires to connect such electronics to the qubits would be required. A viable alternative is to place the electronics much closer to the qubits [12]–[14]: a few circuits would operate at the qubit temperature (10-100 mK), while most of the electronics would be placed at 4 K, i.e. the lowest temperature at which existing dilution refrigerators can still provide a relatively large cooling power (≈ 1 W). Among electronic technologies operating at those temperatures, only CMOS can provide the high level of integration required to reliably manipulate such a large number of qubits. However, although compact simulation models are crucial for the design of the high-performance circuits necessary for a large-scale quantum computer, the characterization and modeling of nanometer CMOS devices at cryogenic temperatures are not yet adequate. The use of nanometer CMOS nodes is particularly relevant to build low-power circuits for quantum-computing applications, since the handling of high-frequency and large-bandwidth signals is required for qubit control. As shown in Fig. 1, several CMOS technologies have been characterized at temperatures below 77 K [15]–[41], but only [24], [37] and [39] address quantum-computing needs, i.e. nanometer technologies (below 0.18 μm) necessary for high-speed circuits (> 10 GHz), operating at 4 K or below where qubits can, in theory, still operate and the cooling power of state-of-the-art dilution refrigerator enables the placement of large integrated circuits. Furthermore, very few cryogenic models have been developed [29]–[39] with only [37] and [39] (of which this paper is an extension) recently focusing on the region of interest. In this paper,
we present the characterization and SPICE modeling of two CMOS nanometer technologies (0.16-μm and 40-nm bulk CMOS) at 4 K, and for 0.16-μm devices also at 1 K and 100 mK [39]. Unlike prior works (see Fig. 1), we model, for the first time, the behavior of bulk CMOS devices at temperatures as low as 100 mK, and we validate the compact models at 4 K by comparing simulations and measurements of two complex circuits, namely a ring oscillator and a low-noise amplifier (LNA) for spin-qubit readout [42]. The paper is organized as follows: Section II describes the devices characterization; Section III discusses the cryogenic behavior and the related physics in detail; Section IV presents the modeling procedure and the resulting DC characteristics are compared to the measured ones; Section V presents the validation of the models with complex circuits and summarizes the effects of cryogenic cooling on technology parameters and on device figures of merit having an impact on future cryogenic-circuit designs; finally, conclusions are drawn in Section VI.

II. CHARACTERIZATION

Several transistors with different dimensions and different oxide thickness from both technologies (SSMC 0.16 μm, ST 40 nm) were tested at 4 K, as shown in Table I. Only a few 0.16-μm devices were cooled down to 1 K and 100 mK, namely one NMOS (W/L = 0.232/1.6) and two PMOS (W/L = 2.32/1.6 and W/L = 0.232/1.6), since other devices did not outlive the bonding required in the sub-Kelvin setup. An ST-500 probe station by Janis Research was used for 300 K and 4 K testing, while a CF-CS81 dilution refrigerator by Leiden Cryogenics was used for sub-Kelvin characterization. $I_D(V_{GS})$ and $I_D(V_{DS})$ characteristics are shown in Fig. 2, 3, 4, 5, Fig. 6, 7, 8, 9 and Fig. 10 for the two technologies.

As mentioned in Section I, mature CMOS processes have been characterized at cryogenic temperatures and the physics underlying several cryogenic effects has been explained. Our measurements closely match to previous observations for several aspects. A general increase in mobility ($\approx 2 \times$), clearly visible in the long-channel devices of Fig. 3b-d and Fig. 7b-d, and threshold voltage ($\approx 30\%$) is observed. The former is due to an overall decrease in electron scattering, while the latter to an increase in ionization energy [43]. From 300 K to 4 K, the subthreshold slope ($SS$) improved by $3.8 \times$ for the 0.16-μm NMOS (Fig. 2) and $3.2 \times$ for the 40-nm NMOS (Fig. 6), thanks to the intrinsic temperature dependence of the diffusion current.

Furthermore, a jump in drain current, which is typically referred to as kink, is observed in the 0.16-μm thick-oxide short-channel NMOS in Fig. 5a-c. The voltage $V_{DS}$ at which the kink occurs, $V_{DS-kink}$ is close to $V_{DD}/2$ and is bias dependent, moving to higher values for larger $V_{GS}$. The kink also causes a very steep subthreshold slope, as shown in Fig. 4a-c for $V_{DS} > 1.7$ V. This jump in current is not observed in all other PMOS and NMOS devices. Although no kink was measured in thin-oxide 0.16-μm NMOS transistors, an onset of the substrate-current body effect (SCBE), common in nanometer CMOS transistors, is present and occurring at a lower $V_{DS}$ at 4 K as compared to 300 K (Fig. 3a-c).

In addition, when cooled down to 100 mK, the 0.16-μm devices showed proper operation, as shown in Fig. 10. In this case, mobility does not significantly change with respect to 4 K and the subthreshold slope improves by only $2.3 \times$ when reducing the temperature from 4 K to 100 mK.

Finally, few devices showed a difference in drain current in the subthreshold region when $V_{GS}$ was swept from low to high or from high to low values. Such hysteresis will be discussed in Section III. For clarity, only the forward sweep is shown in Fig. 2, 3, 4, 5, Fig. 6, 7, 8, 9 and Fig. 10.

### Table I

**SUMMARY OF CHARACTERIZED DEVICES.**

<table>
<thead>
<tr>
<th>Technology</th>
<th>SSMC 0.16 μm</th>
<th>ST 40 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxide</td>
<td>Thick</td>
<td>Thin</td>
</tr>
<tr>
<td>Nominal $V_{GS}$ [V]</td>
<td>3.3</td>
<td>1.8</td>
</tr>
<tr>
<td>$W/L$ [μm / μm]</td>
<td>2 / 1.6</td>
<td>2.32 / 1.6</td>
</tr>
<tr>
<td></td>
<td>2 / 0.322</td>
<td>2.32 / 0.16</td>
</tr>
<tr>
<td></td>
<td>0.47 / 1.81</td>
<td>0.232 / 1.6</td>
</tr>
<tr>
<td></td>
<td>0.47 / 0.322</td>
<td>0.232 / 0.16</td>
</tr>
</tbody>
</table>

Fig. 2. Measured $I_D(V_{GS})$ characteristics of thin-oxide NMOS in 0.16-μm CMOS. $V_{DS} = [0.1 \text{ V}; 0.95 \text{ V}; 1.8 \text{ V}]$. Solid line: 4 K; dashed line: 300 K.

Fig. 3. Measured $I_D(V_{DS})$ characteristics of thin-oxide NMOS in 0.16-μm CMOS. $V_{GS} = [0.68 \text{ V}; 1.24 \text{ V}; 1.8 \text{ V}]$. Solid line: 4 K; dashed line: 300 K.
A. Kink Effect

For bulk CMOS transistors at 4 K, the kink (see Fig. 5a-c) was first explained in [18]. At large $V_{DS}$, the generation of electron-hole pairs due to impact ionization leads to a multiplication current, flowing in part to the drain and the rest to the bulk ($I_{bulk}$ in Fig. 11c). At the same time, an increase in the substrate resistance $R_{bulk}$ is observed due to carrier freeze-out, which appears at temperatures below 70 K and is exemplified by the large resistance of the n-well resistor in Fig. 11b. The multiplication current flowing through $R_{bulk}$ causes the bulk potential to rise, thus producing a threshold voltage drop and, consequently, the current jump measured in Fig. 5a-c. Nonetheless, after reaching the kink, the current jump tends to flatten out as the drain-bulk-source lateral bipolar transistor starts conducting, with the bulk-source junction equivalent to the forward-biased base-emitter junction, thus diverting most of the drain ionization current and leading to a gradual saturation of $I_{bulk}$, as can be seen in the bulk current plot of Fig. 11c. Furthermore, the voltage at which the kink occurs, $V_{DS-kink}$, shifts to higher $V_{DS}$ for larger gate voltages $V_{GS}$. This can be explained by considering that a higher $V_{GS}$ leads to a larger surface scattering and, hence, to a more pronounced mobility degradation, which mitigates the impact ionization and, consequently, the value of $V_{DS}$ where the kink occurs.

It is important to note that the kink is observed only in two thick-oxide NMOS transistors, namely W/L = 2/0.322 and W/L = 0.4/0.322 in the 0.16-μm CMOS technology. Prior works shown in Fig. 1 report the presence of the kink only in mature technologies, mostly for feature sizes larger than 0.35 μm and at temperatures below 100 K. The reason for

III. ANALYSIS OF THE CRYOGENIC BEHAVIOR

In this section, the differences between room-temperature and cryogenic behavior are discussed.

Fig. 4. Measured $I_D(V_{GS})$ characteristics of thick-oxide NMOS in 0.16-μm CMOS. $V_{DS} = [0.1 \text{ V}; 1.7 \text{ V}; 3.3 \text{ V}]$. Solid line: 4 K; dashed line: 300 K.

Fig. 5. Measured $I_D(V_{DS})$ characteristics of thick-oxide NMOS in 0.16-μm CMOS. $V_{GS} = [1.05 \text{ V}; 2.17 \text{ V}; 3.3 \text{ V}]$. Solid line: 4 K; dashed line: 300 K.

Fig. 6. Measured $I_D(V_{GS})$ characteristics of thin-oxide NMOS in 40-nm CMOS. $V_{DS} = [0.1 \text{ V}; 0.6 \text{ V}; 1.1 \text{ V}]$. Solid line: 4 K; dashed line: 300 K.

Fig. 7. Measured $I_D(V_{DS})$ characteristics of thin-oxide NMOS in 40-nm CMOS. $V_{GS} = [0.43 \text{ V}; 0.76 \text{ V}; 1.1 \text{ V}]$. Solid line: 4 K; dashed line: 300 K.
this technology dependence is strictly related to transistor scaling. Traditional scaling rules require thinner oxide thick-
ness to keep a constant vertical electric field. However, the vertical electric field in nanometer technologies has inevitably increased in spite of a proportional scaling of supply and gate insulator thickness, leading to higher mobility degradation (due to surface scattering, for example) that mitigates the impact ionization effect and, hence, the occurrence of the kink. Additionally, the substrate doping of nanometer technologies is considerably higher due to channel engineering, threshold-voltage control, lightly doped drain (LDD) [20], pocket and deep retrograde implants, which help to reduce the drastic increase of the bulk resistance and, hence, the kink effect. Finally, supply voltages below 1.2 V inevitably prevent the kink from occurring because carriers cannot acquire enough energy through impact ionization to overcome the silicon bandgap. All these effects combined explain why we observe the kink only in thick-oxide short-channel 0.16-μm NMOS devices. This type of transistor, in fact, closely resembles a 0.35-μm-CMOS device and does not yet “benefit” from the scaling consequences explained above. Long-channel devices (Fig. 5b-d) do not show a kink because the longitudinal electric field never reaches the required critical electric field to trigger impact ionization. Finally, no kink is observed in 0.16-μm thick-oxide PMOS because their intrinsic lower mobility suffices for the complete suppression of the kink.

B. Sub-Kelvin regime

Measurements at 100 mK∗ showed proper transistor operation, as mentioned in Section II. Fig. 10 shows the $I_D(V_{GS})$ characteristics at 4 K, 1 K and 100 mK of the three transistors listed in Section II normalized to the 300-K measurement to remove any chip-to-chip spread between the devices measured at 4 K in the probe station and those measured in the dilution refrigerator down to 100 mK. It can be observed that the three characteristics in Fig. 10 are almost overlapping for the three transistors, meaning that mobility and threshold voltage do not change significantly at temperatures below 4 K. For the PMOS, the current slightly diminishes below

Fig. 10. Measured characteristics of the three devices tested at 1 K and 100 mK compared to 4 K: (a) NMOS, $V_{DS} = 1.8$ V; (b)-(c) PMOS, $V_{DS} = -1.8$ V. Every curve is normalized to 300 K to remove chip-to-chip spread since the 4 K measurements were done on a different chip placed in the probe station.

Fig. 11. (a) Cross-Section and schematic representation of the cause of kink effect at 4 K; (b) N-well resistance versus current at 4 K; the resistance value at 300 K is 3.5 kΩ; (c) $I_{bulk}(V_{DS})$ of thick-oxide NMOS, W/L = 2.32/0.322 at 4 K.
4 K. The mobility degradation below 4 K is attributed to the strong contribution of carrier-to-carrier and neutral-impurity scattering, which increase at low temperature, as also observed in [44], [45]. Finally, the saturation in threshold voltage is most likely due to the fact that, once freeze-out is reached, the amount of ionized atoms does not significantly change and, hence, the threshold voltage $V_T$ is not notably influenced.

C. Hysteresis

Although it has been a major concern in the past for several technologies operating at cryogenic temperatures [23], [46], [47], no significant hysteresis was observed in most of the tested devices, with the exception of those that showed a kink at 4 K and for the long-channel 0.16-$\mu$m PMOS at 100 mK. 

For the thick-oxide NMOS (W/L = 2/0.16), hysteresis occurs in the subthreshold region, as can be observed in Fig. 12 where a voltage shift in $V_{GS}$ of around 0.15 V is measured. This happens only for $V_{DS} > 1.7$ V, i.e. after the kink. This can be explained by considering that before $V_{GS}$ overcomes the threshold voltage, negligible free charge is present at the surface, thus preventing significant impact ionization. As $V_{GS}$ rises and the channel is being formed, the mechanism of avalanche due to impact ionization builds up as well. Once the channel is formed and the current is not negligible (at $\approx 0.4$ V as shown in the forward-sweep curve in Fig. 12), the threshold voltage decreases due to the kink effect (see Section III-A), reaching the final value once the avalanche current is maximum, which happens when the channel is in strong inversion. When $V_{GS}$ is swept in the other direction (red-circled line in Fig. 12), impact ionization is already occurring at large $V_{GS}$ and, consequently, the threshold voltage $V_T$ is already approximately 0.22 V. The $V_{GS}$ must diminish below this value to completely turn off the transistor.

At sub-Kelvin temperature, a softer hysteresis can be seen in Fig. 13. However, although the temperature of the dilution refrigerator was set to 40 mK, the temperature of the sample varied between 40 mK and 100 mK due to self-heating. This could partially explain the presence of hysteresis in Fig. 13 as a measurement artifact. Further investigations are necessary though, to fully explain the cause of a possible hysteresis in the sub-Kelvin regime. In conclusion, we recognize that hysteresis is not a significant issue in nanometer nodes, since it emerged in very few tested devices.

D. Subthreshold Slope

The subthreshold slope is expected to become steeper with decreasing temperature due to the exponential dependence of $I_D$,

$$I_D \approx I_0 \cdot e^{\frac{(V_{GS}-V_T)}{nk_BT}}$$

(1)

where $I_0$ is the saturation current, $q$ the electron charge, $k_B$ the Boltzmann constant and $T$ the temperature. The expression of the subthreshold slope results, therefore:

$$SS(T) = \left[ \frac{\partial \log(I_D)}{\partial V_{GS}} \right]^{-1} = \ln(10) \frac{nk_BT}{q}$$

(2)

From the Eq. 2, it can be concluded that the subthreshold slope should be linearly dependent on $T$. However, according to the literature [15], the subthreshold slope follows the proportionality with the temperature down to $\approx 77$ K, as predicted by Eq. 2, but it becomes strongly sub-linear at lower temperatures. The sub-linear trend matches our observation of the NMOS $SS$ (Fig. 2), which improves by only $3.8 \times$ from 300 K to 4 K. This can be attributed to the factor $n$ and its dependence on $T$. As a proxy for the subthreshold slope, $n$ is

$$n = 1 + \frac{C_{dep}}{C_{ox}} + \frac{C_{it}}{C_{ox}}$$

(3)

where $C_{dep}$, $C_{it}$ and $C_{ox}$, are the depletion, interface-states and oxide capacitances per device area, respectively. The temperature dependence of $n$ can be explained by the non-negligible increase of interface states at cryogenic temperatures, as reported in [48]. By extracting the value of $n$ from the measurements at temperature $T^*$ as

$$SS(300 \text{ K}) = \frac{n_{300 \text{ K}}}{n_{T^*}} \cdot \frac{300 \text{ K}}{T^*}$$

(4)

we obtain the plot of Fig. 14 for the 0.16-$\mu$m PMOS (W/L = 2.32/1.6). In the figure, we can observe a drastic increase of $n$ below 1 K, which is in agreement with the theory in [48]. In conclusion, a summary of subthreshold slope and $n$ of a PMOS (W/L = 2.32/1.6) versus temperature is shown in Fig. 14.

IV. Modeling

No commercial compact model can predict the behavior of CMOS at cryogenic temperatures and hence, commercial
models cannot be used for the design of circuits operating at cryogenic temperatures. To overcome that limitation, we propose a model for cryogenic CMOS devices based on existing compact models. Our modeling effort aims at demonstrating the capability of standard compact models to cover cryogenic operation without significant modifications. More specifically, we adopted MOS11 [49] and PSP [50] for the 0.16-μm and the 40-nm CMOS technology, respectively, since they were provided by the respective foundry for the standard temperature range. The development of PSP was inspired by MOS11 and therefore the modeling procedure could be mirrored, as will be presented below. Moreover, being MOS11 and PSP derived from surface-potential equations, they are possibly better suited to cover device behaviors in extreme conditions without trading off accuracy for model continuity.

As shown in Section II, the 40-nm CMOS technology did not show any specific cryogenic non-ideality. Therefore, the modeling procedure consisted in updating the parameters in the foundry-provided models with the new values extracted from the cryogenic characterization. On the contrary, some 0.16-μm transistors showed strong non-idealities, such as the kink, which are not included in standard models for bulk CMOS. As a consequence, the modeling of those devices required the addition of extra electrical components to capture semiconductor physics at 4 K. In particular, a non-linear oxide 0.16-μm NMOS transistor was connected to a high-impedance non-linear resistor implemented as a look-up table.

As observed from the measurements. After this, the parameters that impact mobility degradation were modified; in particular, those related to surface scattering (THESATR for MOS11, THEMUO for PSP), and to velocity saturation (THESATR for MOS11, THESATO for PSP). This enabled us to match the curves at high $V_{GS}$ where the characteristics are degraded by these effects. In combination to mobility degradation, impact ionization was enhanced through the parameters A1R, A2R and A3R for the 0.16-μm technology in order to reproduce the kink or the onset of SCBE, while this was not necessary for the 40-nm technology. At this point, the relative error between simulation and measurement is below 20% for most of the curves. The remaining parameters listed in Table II but not mentioned above were modified to fine-tune the models, to compensate for other second-order effects (e.g. channel-length modulation) and, hence, to reduce the mismatch further below 10% in almost all the bias regions.

During the whole procedure, the bulk contact of the thick-oxide 0.16-μm NMOS transistor was connected to a high-impedance non-linear resistor implemented as a look-up table. The data of such look-up table were extracted from the measurement of the n-well resistor in Fig. 11b and scaled by a constant coefficient at the end of the modeling procedure to tune the magnitude of the current jump and $V_{DS}-kink$. The final implemented resistor has an impedance of 100 kΩ at $I_{bulk} = 1$ nA.

The modeling of the devices at 100 mK followed the same procedure, proving that the models are scalable with temperature. It has to be noted that the parameters related to the dynamic behavior of the transistor were not modified and hysteresis was not modeled since the majority of the devices did not show this effect. Finally, although the extraction procedure described above has not yet been automated, we believe that, especially for the majority of the devices that do not show kink, the standard extraction procedure for the respective model can be adopted.

### B. Modeling results

The results of this parameter-fitting procedure are shown in Fig. 15 and Fig. 16 where both 4 K and 100 mK measurements (dashed lines) and simulations based on the newly proposed models (solid lines) are superimposed. Good matching (relative error < 10%) of simulations and experimental data is obtained at every bias point, in both weak and strong inversion, with a single set of parameters for each device. However, the accuracy of the model is limited in the moderate-inversion region, e.g. in Fig. 15f for $V_{GS} = 0.9$ V → 1 V, because both models employ a smoothing function to continuously merge the equations for weak and strong inversion, which was not modified in this work.

![Graph](image-url)

**Fig. 14.** Subthreshold slope ($SS$) and extracted ideality factor $n$ of the PMOS ($W/L = 2.32/1.6$) across temperature.
of 0.16-µm thick-oxide NMOS; (b,f) $I_{DS}$ of 0.16-µm thick-oxide PMOS (W/L = 2.32/1.6) at 100 mK. (a) $I_{DS}$, $V_{GS} = -1.8 \rightarrow -0.09$ V; (b) $I_{DS}$, $V_{GS} = -0.4 \rightarrow -1.8$ V. 

V. MODEL VALIDATION

A. Circuits at cryogenic temperature

A 2703-stages ring oscillator in 0.16-µm CMOS has been tested at 4 K. Its measured oscillation frequency is shown in Fig. 17 for both 300 K and 4 K for different supply voltages and compared to the simulation employing the proposed model. The higher driving capability reduces the gate delay by almost 30%. The simulation was performed on a different model. The higher driving capability reduces the gate delay and compared to the simulation employing the proposed model. The observed discrepancy is well within the margin due to process spread, which has not been modeled in this work due to lack of the required corner-lot samples. In conclusion, both experiments above confirm the effectiveness of the parameter extraction and the modeling procedure presented above.

B. Impact of cryogenic CMOS on circuit design

Device performance at 300 K and 4 K is compared and summarized in Table III for minimum-length devices for both technologies. The higher threshold voltage $V_T$ combined

![Fig. 15. Models (solid) and measurements (dashed) at 4 K; W/L in µm. (a,c) $I_{DS}(V_{GS})$ and $I_{DS}(V_{DD})$ of 0.16-µm thick-oxide NMOS; (b,f) $I_{DS}(V_{GS})$ and $I_{DS}(V_{DD})$ of 0.16-µm thick-oxide NMOS; (c,g) $I_D(V_{GS})$ and $I_D(V_{DD})$ of 40-nm thin-oxide NMOS; (d,h) $I_D(V_{GS})$ and $I_D(V_{DD})$ of 40-nm thick-oxide NMOS.

![Fig. 16. Models (solid) and measurement (dashed) of 0.16-µm thin-oxide PMOS (W/L = 2.32/1.6) at 100 mK. (a) $I_{DS}$, $V_{GS} = -1.8 \rightarrow -0.09$ V; (b) $I_{DS}$, $V_{GS} = -0.4 \rightarrow -1.8$ V.

![Fig. 17. Ring oscillator frequency versus supply voltage ($V_{DD}$) at 300 K and 4 K. Measurements (dotted lines) and simulations (solid lines) are compared with the compact model presented in Section IV.](http://www.ieee.org/publications_standards/publications/rights/index.html for more information.)
with a steeper $SS$ is beneficial to minimize leakage currents. However, a high $V_T$ reduces the voltage headroom available for circuits, especially when operating at low supply voltages. This may force the adoption of low-voltage circuit topologies typically showing a lower energy efficiency [52]. This can be particularly detrimental in cryogenic applications, such as quantum-computing controllers, because of the tight power budget imposed by the cooling capabilities of dilution refrigerators.

The effect of the higher $V_T$ is partially compensated by the increase in mobility, which results in an overall increase of the maximum drain current (see $I_{on}$ in Table III) and, consequently, in a faster switching speed (see Fig. 17). The larger driving current and lower leakage bring an improved $I_{on}/I_{off}$ ratio (better than $100 \times$ for 40 nm), which is advantageous for the implementation of high-energy-efficiency digital logic. Table III also shows the transconductance efficiency in terms of $g_m/I_D$ ratio in both weak and strong inversion, extracted from the measurements shown in Fig. 2, 3, 4, 5, Fig. 6, 7, 8, 9. A significant improvement ($\approx 3 \times$) is observed at 4 K in weak inversion while a slightly lower efficiency appears in strong inversion. This would lead to a major improvement in power efficiency for cryogenic circuits biased in weak inversion. Finally, a lower intrinsic gain is observed, due to the large increase in the channel-length modulation factor $\lambda$, in agreement to what was observed in [44].

VI. CONCLUSIONS

This work presents the extensive DC characterization of transistors fabricated in 0.16-µm and 40-nm CMOS technologies at deep-cryogenic temperatures (4 K, 1 K and 100 mK). A detailed understanding of the device physics at cryogenic temperatures was developed and captured in a compact model that was validated both via device DC characterization and through the design and testing of complex circuits.

It was demonstrated that nanometer bulk CMOS devices can operate reliably down to 100 mK and that, for the first time, their behavior can be modeled at such low temperature. Moreover, nanometer CMOS technologies are not affected by strong cryogenic non-idealities, such as kink, and several device figures of merit, such as $I_{on}/I_{off}$ and $g_m/I_D$, significantly improve at 4 K, which is expected to enhance the performance of both analog and digital circuits at cryogenic temperatures.

Thus, the proposed models will enable the design and simulation of circuits in nanometer CMOS, which is a viable technology for the implementation of high-performance and power-efficient cryogenic circuits, as required in demanding applications such as quantum computing.

ACKNOWLEDGMENT

The authors would like to thank Intel Corp. for funding, XNP and ST Microelectronics for chip fabrication and dr. Leonardo DiCarlo for help in the 100-mK characterization.

REFERENCES

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2018.2821763, IEEE Journal of the Electronic Devices Society
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JEDS.2018.2821763, IEEE Journal of the Electron Devices Society


Rosario M. Incandela (S'17) was born in Trapani, Italy in 1992. He received the B.Sc. in Electronics Engineering from Politecnico di Milano, Italy and the M.Sc (cum laude) in Microelectronics from Delft University of Technology, The Netherlands, in 2014 and 2016, respectively. In 2016, he started the Ph.D. in cryogenic electronics for quantum applications at Delft University of Technology. His main research interests include low-noise cryogenic CMOS readout of spin qubits and modeling of CMOS technologies at cryogenic temperature.

Lin Song received the B.S., M.S. degrees in electrical engineering from Tsinghua University, Beijing, China and Delft University of Technology, Delft, the Netherlands, in 2014 and 2016, respectively. He is currently an analog designer with Analog Devices Inc., in Beijing design office, Beijing, China.

Andrei Vladimirescu (F’17) received the M.S. and Ph.D. degrees in EECS from the University of California at Berkeley, California, where he was a key contributor to the SPICE simulator, releasing the SPICE2G6 production-level SW in 1981. He pioneered electrical simulation on parallel computers with the CLASSIE simulator as part of his PhD. He is the author of “The SPICE Book” published by J. Wiley and Sons.

For many years Andrei was R&D director leading the design and implementation of innovative software and hardware Electronic Design Automation products for Analog Devices Inc., Daisy Systems, Analog Design Tools, Valid Logic and Cadence. Currently he is Professor involved in research projects at the University of California at Berkeley, Delft University of Technology and the Institut Supérieur d’Électronique de Paris, ISEP, as well as consultant to industry. His research activities are in the areas of ultra-low-voltage (ULV) CMOS, design, simulation and modeling of circuits with new devices and circuits for quantum computing.

Andrei is an IEEE Fellow.

Edoardo Charbon (SM’00–F’17) received the Diploma from ETH Zurich, the M.S. from the University of California at San Diego, and the Ph.D. from the University of California at Berkeley in 1988, 1991, and 1995, respectively, all in electrical engineering and EECS. He has consulted with numerous organizations, including Bosch, X-Fab, Texas Instruments, Maxim, Sony, Agilent, and the Carlyle Group. He was with Cadence Design Systems from 1995 to 2000, where he was the Architect of the company’s initiative on information hiding for intellectual property protection. In 2000, he joined Canesta Inc., as the Chief Architect, where he led the development of wireless 3-D CMOS image sensors. Since 2002 he has been a member of the faculty of EPFL, where is a full professor since 2015. From 2008 to 2016 he was with Delft University of Technology’s as Chair of VLSI design. He has been the driving force behind the creation of deep-submicron CMOS SPAD technology, which is mass-produced since 2015 and is present in telemeters, proximity sensors, and medical diagnostics tools. His interests span from 3-D vision, FLIM, FCS, NIROT to super-resolution microscopy, time-resolved Raman spectroscopy, and cryo-CMOS circuits and systems for quantum computing. He has authored or co-authored over 250 papers and two books, and he holds 20 patents. Dr. Charbon is a distinguished visiting scholar of the W. M. Keck Institute for Space at Caltech, a fellow of the Kavli Institute of Nanoscience Delft, a distinguished lecturer of IEEE Photonics Society, and a fellow of the IEEE.

Harald Homulle received the B.S. and M.S. degrees in electrical engineering from Delft University of Technology, The Netherlands, in 2012 and 2014, respectively. During the Master’s project, he did an internship at EPFL, Lausanne, Switzerland, for the development of a multi-channel TDC on FPGA. In 2014 he joined the research group of Edoardo Charbon as a Ph.D. student on the topic of cryogenic electronics, with the main focus on quantum computing applications in collaboration with other QuTech researchers.
Fabio Sebastiano (S’09–M’10–SM’17) was born in Teramo, Italy in 1981. He received the B.Sc. (cum laude) and M.Sc. (cum laude) degrees in electrical engineering from University of Pisa, Italy, in 2003 and 2005, respectively, the M.Sc. degree (cum laude) from Sant’Anna school of Advanced Studies, Pisa, Italy, in 2006 and the Ph.D. degree from Delft University of Technology, The Netherlands, in 2011. From 2006 to 2013, he was with NXP Semiconductors Research in Eindhoven, The Netherlands, where he conducted research on fully integrated CMOS frequency references, deep-submicron temperature sensors and area-efficient interfaces for magnetic sensors. In 2013, he joined Delft University of Technology, where he is currently an Assistant Professor. He has authored or co-authored one book, nine patents, and over 40 technical publications. His main research interests are sensor read-outs, fully-integrated frequency references, and cryogenic electronics for quantum applications.