High-Speed Steep-Slope GaInAs Impact Ionization MOSFETs (I-MOS) With $SS = 1.25$ mV/dec—Part I: Material and Device Characterization, DC Performance, and Simulation

Daxin Han, *Graduate Student Member, IEEE*, Giorgio Bonomo, Diego Calvo Ruiz, *Graduate Student Member, IEEE*, Akshay Mahadev Arabhavi, *Graduate Student Member, IEEE*, Olivier J. S. Ostinelli, and Colombo R. Bolognesi, *Fellow, IEEE*

***Abstract*— Digital electronics power consumption evolved into a major concern: at the current pace, general-purpose computing energy consumption will exceed global energy production before 2045. The principal approach to curbing energy consumption in digital applications calls for "steep-slope" devices with an inverse subthreshold slope (*SS*) parameter well below the "ln(10)·*kT*/*q*" limit of conventional electronics (60 mV/dec at 300 K). Impact ionization MOSFETs (I-MOS) provide an avenue for steep-slope device realization. High-mobility narrow gap III-V semiconductor channel materials have not yet been investigated for I-MOS applications. We hereby report E-mode narrow bandgap GaInAs-based I-MOS devices with an *SS* of 1.25 mV/dec maintained over five orders of magnitude in drain current and I_{ON}/I_{OFF} ratios $>10^6$ at 300 K ($>10^9$ at 15 K) for a gate length of $L_G = 100$ nm. Part I of this work focuses on the materials and device fabrication and analysis, device dc characterization, and modeling. The present GaInAs devices are the first I-MOS transistors to display a robust steep-slope effect at low voltages $V_{DS} < 1.9$ V at 300 K and <1 V at 15 K. Part II describes the dynamic switching (including clarifications on the role of hysteresis) and RF characteristics of GaInAs I-MOS devices and benchmarks them with respect to other steep-slope technologies.**

***Index Terms*— Atomic layer deposition (ALD), GaInAs, impact ionization, impact ionization MOSFET (I-MOS), inverse subthreshold slope (*SS*), MOSFET, RF performance, steep slope.**

I. Introduction

BY THE early 1980s, CMOS was well underway to become the technology of choice for VLSI applications [1]. The subject of CMOS digital electronics power consumption received relatively little attention until the 1990s, when CPU clock frequencies rose swiftly: from 50 MHz in 1989 [2], clocks reached 200 MHz in 1992 [3] and 1 GHz in 2000 [4], before eventually saturating between 3 and 5 GHz since 2001 [5]–[7]. The concomitant near exponential rise in power density (dissipated power per unit chip area, in W/cm^2) reached $\sim$100 W/cm^2 in 2004 [8], where it also saturated with the introduction of multicore CPU architectures. After guiding the industry for over a decade in scaling between the 500 and 90 nm nodes (1992–2003), Dennard's constant power density scaling rules [9] could no longer be practically maintained, leading to faster reductions in transistor area than in transistor power dissipation—and the situation stands to become even more acute by opening up the third dimension for device manufacturing with 3-D integration [10], [11]. In the span of a decade, power dissipation evolved from a specialty consideration (*i.e.*, battery-based mobile applications) to again become a primary concern in semiconductor digital computing electronics, some 20 years after silicon bipolar junction transistors ran into similar concerns in the 1970s to eventually drive the widespread adoption of CMOS technology [12].

Over the last two decades, the need to curb the increasing power consumption of digital electronics elicited interest in the development of "steep-slope" devices as potential successors to conventional CMOS [13], [14]. Power dissipation concerns become even more pressing with the rapid continuous expansion of data-centric activities, such as high-density video streaming, Internet-based teleconferencing, remote learning/teaching, and work from home in light of related data center electrical needs (with a significant portion dedicated to cooling). Global data center energy consumption carries a carbon footprint that is currently comparable to worldwide air travel, but with a far higher compound annual growth rate. According to scenarios [15], data centers are expected to consume $\sim$300 TWh/year by 2030, and the electricity consumption footprint for information and communication techniques (ICTs), including data centers, is forecasted to amount to 21% of the global electrical needs (up to 51% in the worst case scenario) [15]. Furthermore, at the current pace, the total energy consumption of general purpose computing will exceed global energy production before 2045 [16].

Manuscript received March 4, 2022; revised April 22, 2022; accepted April 28, 2022. Date of publication May 13, 2022; date of current version June 21, 2022. The review of this article was arranged by Editor D. Triyoso. *(Corresponding author: Colombo R. Bolognesi.)*

The authors are with the Department of Information Technology and Electrical Engineering, MWE Group, ETH Zürich, 8092 Zürich, Switzerland (e-mail: colombo@ieee.org).

Digital Object Identifier 10.1109/TED.2022.3171739

In comparison to traditional transistors, steep-slope transistors require lower gate voltage swings than conventional FETs to switch ON/OFF. The slope in question is the inverse subthreshold slope (SS) parameter, which characterizes the transition abruptness between ON and OFF states in terms of the voltage swing required to achieve a tenfold change in current flow. To contextualize, an ideal p-n junction diode is characterized by $SS = 60$ mV/dec at room temperature and low current densities. In comparison, as briefly reviewed below, long-channel MOSFETs already show $SS \geq 60$ mV/dec because the gate control voltage is necessarily divided between the gate insulator and the body depletion before channel inversion. Furthermore, device scaling generally increases the SS parameter, reflecting the increasing role of 3-D fields which make it more difficult to turn short-channel transistors off. Efforts to better electrostatically control these fields to maintain SS as close to 60 mV/dec as possible required major research and design investments, paving the way to advanced 3-D MOSFET architectures such as high-κ gate dielectrics FinFETs. In MOSFETs, the subthreshold current I_D follows:

$$I_D \propto I_{OFF} e^{\frac{q(V_{GS}-V_T)}{nkT}} \tag{1}$$

with

$$n = 1 + \frac{C_d}{C_{ox}} \tag{2}$$

where V_T is the threshold voltage, C_d is the channel depletion capacitance, C_{ox} is the gate oxide capacitance, and the ratio C_d/C_{ox} reflects the division of the gate voltage across the oxide and the underlying depletion region. The OFF-state current is then I_D evaluated at $V_{GS} = 0$ V and $V_{DS} = V_{DD}$ (the power supply voltage). Hence

$$SS = \log\left(\frac{\mathrm{d}\ln I_D}{\mathrm{d}V_{GS}}\right)^{-1} = \ln(10) \cdot n \cdot \frac{kT}{q} \cong n \cdot 60\text{mV/dec} \quad \text{(at 300 K)}. \tag{3}$$

Clearly, long-channel MOSFETs already show $SS \geq 60$ mV/dec at 300 K, even before scaling effects set in to increase SS. The situation is further degraded with operating junction temperatures of ~100 °C.

Interest in steep-slope transistors stems from a consideration of the power dissipation in conventional CMOS

$$P = f \cdot C V_{DD}^2 + I_{Leak} \cdot V_{DD} + I_{Short} \cdot V_{DD} \tag{4}$$

where f is the clock frequency, C represents the total circuit capacitance inclusive of wiring interconnects, V_{DD} is the operating supply voltage, I_{Leak} is the total leakage current consisting of I_{OFF} including the gate and junction leakage currents (where I_{Leak} is typically dominated by I_{OFF} following the introduction of high-κ gate dielectrics to replace SiO_2 in gate oxides), and I_{Short} is the short-circuit current in switching events when both the n- and p-channel devices simultaneously conduct (a small contribution with proper design, and if the V_T/V_{DD} ratio is high enough). The first term in (4) represents the dynamic power consumption of logic switching, and it is responsible for the saturation in clock frequencies and it clearly benefits from reductions in the power supply voltage V_{DD}. However, lowering V_{DD} should be accompanied by a reduction in threshold voltage V_T to maintain a constant transistor ON current drive (and circuit speed, CV_{DD}/I_{ON}), but lowering V_T increases the OFF-state current I_{OFF} and the associated leakage power term $I_{Leak} \cdot V_{DD}$. Indeed, the CMOS OFF-state power has grown with the downscaling of gate length to become comparable to the dynamic power dissipation in traditionally scaled devices, effectively halting traditional CMOS scaling practices [17]. Taking I_T as threshold current, the total leakage current is thereby related to SS

$$I_{Leak} \cong I_{OFF} = I_T \cdot 10^{-\frac{V_T}{SS}}. \tag{5}$$

It is thus the aim of steep-slope technologies to drastically suppress the second term in (4) by engineering devices that can achieve very low I_{OFF} values with reduced voltage swings, ideally, with $SS \ll 60$ mV/dec at room temperature.

The search for the ideal successor to CMOS electronics led to a broad variety of innovative device concepts, including impact-ionization MOSFETs (I-MOS) [14], [18], TFETs [19], negative-capacitance FETs (NCFETs) [20], metal–insulator–transition FETs (MITFETs) [21], antiferromagnetic or ferromagnetic-like FETs [22], nanoelectromechanical FETs (NEMFETs) [23], dirac-source FETs [24], [25], and combinations thereof [26], [27]. Among these concepts, I-MOS devices tend to exhibit weaker short-channel effects and achieve lower OFF currents [28], [29].

Desirable attributes of a steep-slope transistor technology first and foremost include steep-slope behavior over several orders of magnitude in current, as well as operation at low power supply voltage levels, low-leakage current levels I_{OFF}, and a high current drive I_{ON} to rapidly charge/discharge circuit load and interconnect capacitances and maintain/improve circuit performance. Other desirable features include the possibility to form complementary devices to minimize standby power dissipation, as well as large-scale high-yield manufacturability.

The development of I-MOS devices so far largely focused on silicon-based devices, which tends to require a high V_{DD} to achieve steep-slope behavior in lateral devices because of the relatively high electric fields required to induce impact ionization in indirect gap materials. Savio *et al.* [30] reviewed the limitations of Si-based I-MOS transistors at low voltages, and Germanium was suggested as a lower energy gap alternative to facilitate impact ionization at lower drain voltages V_{DS}. Because of challenges associated with high required voltages and the slow temporal take-up of ionization in silicon-based I-MOS devices [30], steep-slope transistor research emphasis largely shifted toward the realization of TFET devices in recent years.

Considering the above, it may be surprising that I-MOS research to date has not yet exploited direct narrow bandgap semiconductors such as high-indium content GaInAs channels in view of their low thresholds for impact ionization. Indeed, much effort continues to be invested in mitigating impact ionization effects in high-indium content high electron mobility transistors (HEMTs) [31], [32], where the resulting parasitic bipolar effect due to generated holes [33] leads to an I_D–V_{DS} characteristics kink and negatively impacts the microwave minimum noise figure (NF_{MIN}) of HEMTs [34]. We recently demonstrated that the use of an Al_2O_3 gate oxide greatly enhances impact ionization artifacts in GaInAs MOS-HEMTs in comparison to conventional Schottky gate HEMTs fabricated with a channel of the same thickness and composition by bolstering the lateral parasitic bipolar effect [35]. Such findings naturally lead one to consider whether high impact ionization materials such as GaInAs (or InAs) channels could profitably be exploited in the realization of I-MOS steep-slope transistors.

In the present work, we report enhancement-mode (E-mode) I-MOS based on a GaInAs channel with an atomic-layer deposited (ALD) amorphous Al_2O_3 gate oxide. Device fabrication is largely similar to that of a conventional HEMT. The process involves no epitaxial regrowth, leading to a fairly straightforward device architecture. The resulting GaInAs MOSFETs feature high I_{ON}/I_{OFF} ratios, high I_{ON} drive currents, and low gate leakage currents and display a robust steep-slope behavior maintained over several orders of magnitude in current. Our devices were characterized at 300 and 15 K: at room temperature, steep-slope behavior is maintained over five orders of magnitude in drain current with low operating drain voltages, with an I_{ON}/I_{OFF} ratio of more than 10^6, while at 15 K, the steep slope is maintained over six orders of magnitude with an I_{ON}/I_{OFF} ratio of more than 10^9 limited by the low-current resolution of the test instrument.

II. Fabrication

A. GaInAs I-MOS Fabrication

GaInAs I-MOS devices were implemented in the ETH Zürich FIRST Center for Micro- and Nanoscience Laboratory, including the molecular beam epitaxial (MBE) growth of the GaInAs I-MOS epitaxial layer structures [see Fig. 1(c)] and subsequent device fabrication. The present I-MOS devices are based on a buried heterostructure $Ga_{0.32}In_{0.68}As$ channel with a thin (5 nm) top-side InP/AlInAs composite barrier. The layers were grown on a 350-nm $Al_{0.48}In_{0.52}As$ buffer layer deposited on an Fe-doped semi-insulating 100 InP substrates. The active region layer stack specifically consists of an undoped 12.5-nm $Ga_{0.32}In_{0.68}As$ n-type quantum well sandwiched between a 5-nm $Al_{0.48}In_{0.52}As$ layer on the bottom and a 2-nm $Al_{0.47}In_{0.53}As$ top-side barrier layer followed by a 3-nm InP selective etch-stop barrier, which enables the formation of a gate recess by selective wet etching. An n^+ composite cap structure doped with Si at 7.3×10^{19} cm^{-3} consisting of 10-nm lattice-matched $Ga_{0.47}In_{0.53}As$ and 10-nm of strained $Ga_{0.32}In_{0.68}As$ is used for the formation of low-resistance source and drain Ohmic contacts. No delta-doping is used here because E-mode operation is desired for I-MOS devices.

Atomic layer deposited (ALD) Al_2O_3 was selected as the gate oxide among other high-κ materials because it offers some of the largest conduction and valence band discontinuities with respect to GaAs, InAs, and InP [36]. Hot carrier injection into SiO_2 was identified as the primary failure mechanism of some Si-based I-MOS transistors [28] because their operating drain biases ($V_{DS} > 10$ V) far exceed barriers for carrier injection into the oxide ($\Delta E_C = 3.1$ eV, $\Delta E_V = 4.8$ eV for SiO_2:Si). For ALD Al_2O_3:InP, the reported $\Delta E_C = 2.18$ eV [37] corresponds to an oxide injection barrier of $\Delta E_C = 2.43$ eV ($\Delta E_V = 2.99$ eV) for electrons (holes) residing in $Ga_{0.47}In_{0.53}As$ once InP/$Ga_{0.47}In_{0.53}As$ band discontinuities are accounted for. As shown next, the oxide injection barriers are higher than the GaInAs I-MOS operating voltages ($V_{DS} < 2$ V), thus eliminating hot carrier injection concerns. Whereas Al_2O_3 deposited directly on GaInAs has been associated with a low device reliability [38], its deposition on InP results in GaInAs FETs exhibiting low $C - V$ and I_D–V_G dispersion properties [39] as well as low interface and border trap densities [40]. Further considerations on the reliability aspects of high-κ gates on GaInAs channels are given in Part II of this work.

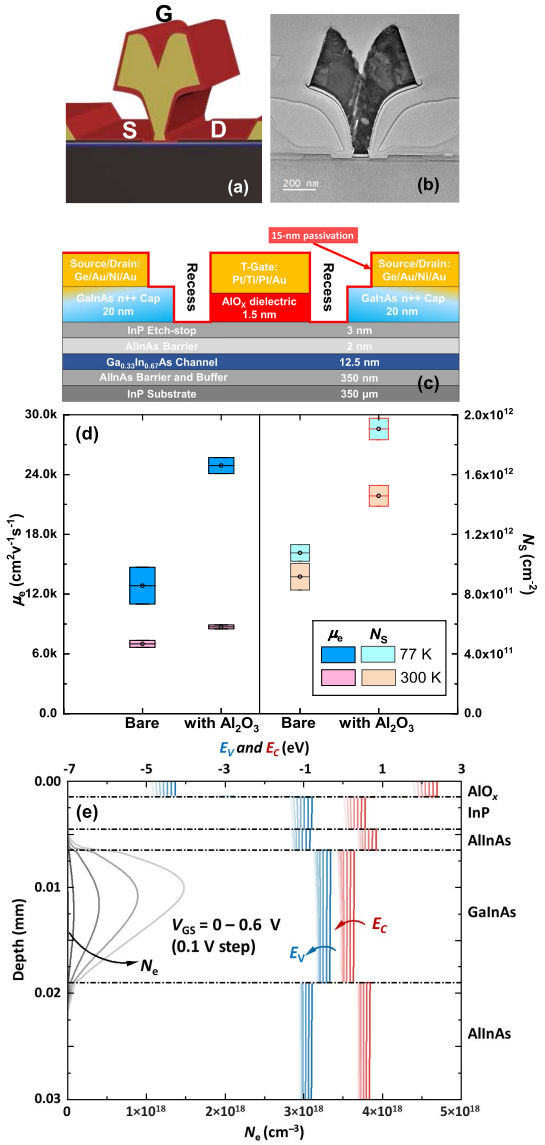

Fig. 1. (a) Three-dimensional schematic representation of the GaInAs I-MOS device active region. (b) STEM cross-sectional image of the active region. (c) MBE-grown epitaxial layer stack and electron-beam deposited metal contact layers. (d) vdP measurement electron mobility and density in the as-grown and Al_2O_3 ALD-passivated epitaxial structure with identical layers to the fabricated devices with exception of an additional delta-doping sheet below GaInAs channel to characterize the influence of oxide surface passivation on 2-DEG transport properties. (e) TCAD simulated cross-sectional band diagrams for various gate voltages (conduction band in red, valence in blue, and electron density in gray).

To characterize the effect of the thin ALD gate oxide deposition on the channel transport properties, van der Pauw (vdP) measurements were performed on as-grown and ALD oxide passivated structures. However, the undoped channel structures used in I-MOS E-mode transistors are not suitable for vdP characterization due to the absence of channel electrons. To enable vdP measurements, a similar epitaxial structure was grown with a δ-doping plane in the lower barrier to supply electrons to the channel. With the n+ GaInAs cap removed, vdP measurements on the δ-doped layers showed an as-grown mobility $\mu = 6660$ (14 700) cm^2/Vs and a sheet carrier density $N_S = 8.3 \times 10^{11}$ (1.13×10^{12}) cm^{-2} at 300 (77) K after 10 min in darkness. Measurements following

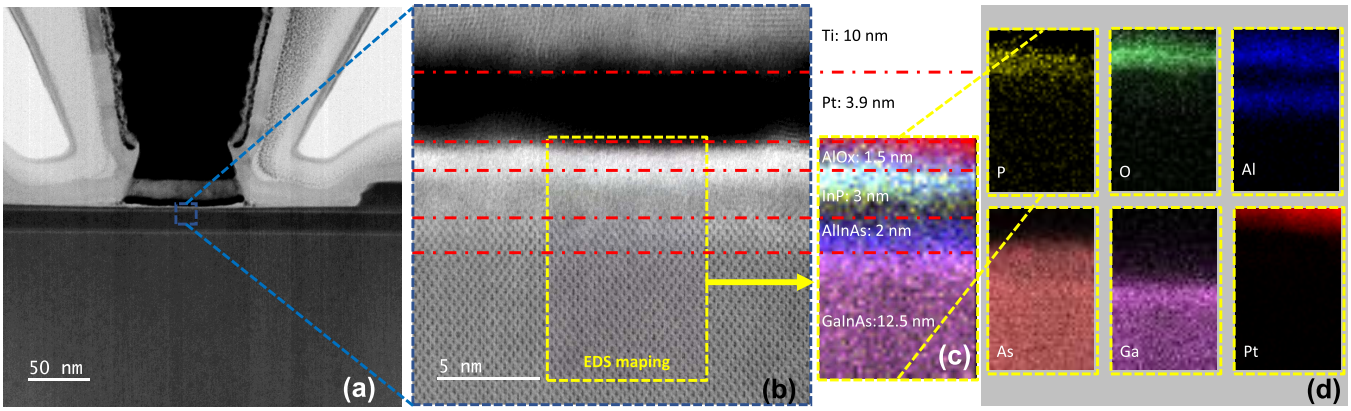

Fig. 2. STEM imaging and energy dispersive spectroscopy (EDS) mapping of the gate region of an GaInAs I-MOS transistor. (a) STEM image of a GaInAs I-MOS gate-electrode showing a gate length $L_G = 100$ nm. (b) Probe-corrected atomic-scale STEM image of the corresponding marked region (blue-box) zoom underneath the gate electrode, showing barrier layers and part of the 12.5-nm GaInAs channel. The ALD process oxidized half of the 3-nm InP etch stop barrier. (c) Composition mapping of the EDS mapping region (yellow box), showing layered elemental distribution in mixed color plots corresponding to individual elements of phosphorus, oxygen, aluminum, arsenic, gallium, and platinum, shown separately in (d).

the ALD deposition of a 1.5-nm amorphous Al_2O_3 layer yielded an improved mobility $\mu = 8920$ (25 700) cm^2/Vs and a sheet carrier density $N_S = 1.39 \times 10^{12}$ (1.98×10^{12}) cm^{-2} at 300 (77) K, also after 10 min in darkness. The ALD passivation therefore improves both the mobility and sheet carrier density of near-surface AlInAs/GaInAs two-dimensional electron gas (2DEG) channels. Similar findings were recently reported by Södergren *et al.* [41].

The present GaInAs I-MOS device fabrication process does not involve epitaxial regrowth steps. First, a Ge/Au/Ni/Au (18/48/10/50 nm) metal stack for Ohmic contacts was e-beam deposited following an ion-beam cleaning of the semiconductor surface. Following liftoff, the Ohmic contact metal stack was annealed at 273 °C. The composite GaInAs cap was then selectively etched in a mixture of succinic acid and hydrogen peroxide to form a 300-nm wide gate recess opening defined by electron beam lithography. A ~1.5-nm amorphous Al_2O_3 layer was deposited by atomic layer deposition (ALD) according to postdeposition ellipsometric measurements. Its thickness was confirmed by scanning transmission electron microscopy (STEM) on the completed devices (see Fig. 2). T-gate electrodes with a (2×75) μm width and a footprint $L_G = 100$ and 150 nm were centered in the gate recess by electron beam lithography and e-beam evaporation of Pt/Ti/Pt/Au (3/20/20/357 nm). Following the gate metal stack deposition and liftoff, the device active region was passivated with an additional 15 nm layer of ALD Al_2O_3.

III. Results

A. GaInAs I-MOS Steep-Slope Switching Characteristics

At higher V_{DS}, impact ionization creates secondary electron–hole pairs near the drain end of the gate: whereas generated electrons flow out the drain contact with the initiating electrons, the resulting holes drift back toward the source region, where their positive space charge lowers the potential barrier at the source end of the channel, thus allowing more electrons to enter the channel through a bipolar transistor-like positive feedback mechanism on the drain current [33]. The gate oxide suppresses hole leakage through the gate with respect to the Schottky gate in an HEMT [35], enhancing the positive feedback on the drain current I_D. For sufficiently high applied drain voltages, a steep-slope behavior develops in the I_D–V_G transistor characteristics, giving rise to a rapid current increase in drain current. The use of a narrow direct bandgap channel material facilitates a rapid onset of impact ionization at lower voltages than possible in silicon-based devices.

Fig. 3 shows the dc characteristics of the fabricated GaInAs I-MOS transistors measured at 300 and 15 K. At 300 K, the inverse *SS* for MOSFETs with a gate length of $L_G = 100$ nm begins to break through the $2.3kT/q \approx 60$ mV/dec Boltzmann limit at a drain bias of $V_{DS} = 0.9$ V, where it shows $SS = 47.6$ mV/dec due to increasing levels of impact ionization. The I_D–V_{DS} characteristics show the impact ionization-induced kink for $V_{DS} > 0.7$ V. The maximum steep slope of $SS = 1.25$ mV/dec over six orders of magnitude in drain current is observed with $V_{DS} = 1.9$ V in the gate turn-on sweep for a 100-nm gate. For a gate length of $L_G = 150$ nm, *SS* is degraded to 15.8 mV/dec under the same bias due to the reduced channel electric field associated with the longer gate length. For the turn-off gate sweep, the 100-nm gate shows a slightly degraded $SS = 6.65$ mV/dec. The above steep-slope characteristics compare favorably to the report of simulated Ge transistors predicted to achieve I-MOS operation at $V_{DS} = 1.8$ V for a 60-nm gate length, as well as to an experimental $SS = 4.5$ mV/dec maintained over less than three orders of magnitude in current at $V_{DS} = 8.5$ V for a 100-nm gate length and a $Si_{0.85}Ge_{0.15}$ impact ionization region [42]. While Si-based I-MOS transistors show significant threshold voltage shifts after a few switching cycles (as few as five cycles [28]), the present devices operate stably: Part II of this article shows stable operation over tens of billions switching cycles. Further comments on device reliability are given in Part II.

At 15 K, clear steep-slope behavior is observed with $V_{DS} \leq 1$ V for both gate lengths, with a significantly increased I_{ON}/I_{OFF} ratio of over nine orders of magnitude in current. The measured I_{OFF} is limited by the curve tracer minimum resolution. A steep slope with $SS < 3$ mV/dec is observed at a lower $V_{DS} > 0.7$ V because of enhanced impact ionization at low temperatures. These observations suggest that subvolt operation should be possible at 300 K if impact ionization levels can be increased with the use of InAs channels and shorter gate lengths. To the best of our knowledge, the present 15 K data are the first to display steep-slope behavior with $V_{DS} \leq 1$ V for an I-MOS device.

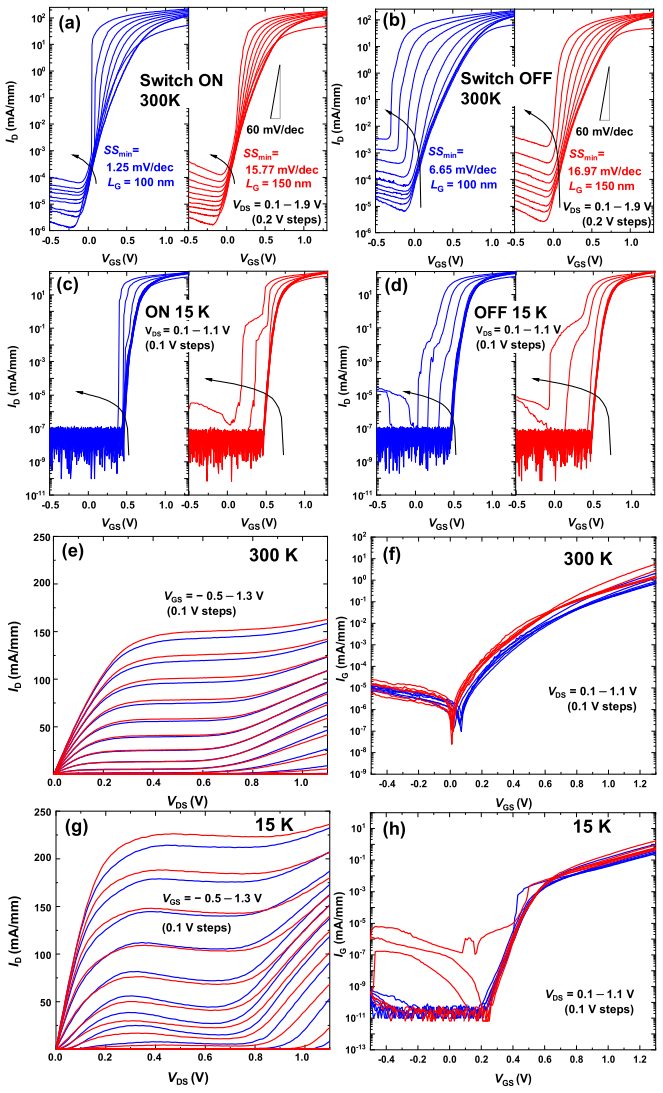

Fig. 3. Static characteristics of GaInAs I-MOS devices with L_G = 100/150 nm gate lengths (blue/red traces, respectively). (a) and (c) OFF-to-ON switching at 300 and 15 K. (b) and (d) ON-to-OFF switching at 300 and 15 K. At 300 K, V_{DS} = 0.1–1.9 V (0.2 V steps) with a minimum *SS* of 1.25 mV/dec. At 15 K, V_{DS} = 0.1–1.1 V (0.1 V steps). Drain characteristics with V_{GS} = −0.5–1.3 V (0.1 V steps) showing kink effect drain for V_{DS} > 0.7 V at both (e) 300 and (g) 15 K. Gate leakage current with V_{DS} = 0.1–1.1 V (0.1 V step) at both (f) 300 and (h) 15 K.

B. GaInAs I-MOS TCAD Simulation

TCAD drift-diffusion simulations were performed to qualitatively support our interpretation of the steep-slope behavior observed in GaInAs I-MOS transistors. With a 2-D device layout corresponding to the cross section of the fabricated transistors shown in Fig. 1(b), transistor dc characteristics were modeled at 300 K using the Silvaco Atlas TCAD simulation tool to qualitatively reproduce the steep ON-switching current characteristics [see Fig. 4(b)] and the kink in the I_D–V_{DS} drain characteristics [see Fig. 4(c)]. Specifically, the simulations of the ON-switching drain current in Fig. 4(b) show a steep-slope onset at a similar V_{GS} as experimentally measured on our I-MOS transistors and a clear increase of the steep-slope behavior at higher V_{DS}. The simulated I_D–V_{DS} curves show a strong kink effect due to impact ionization for V_{DS} > 0.7 V, in accordance with measurements. E-mode operation is achieved without involving channel inversion as in a Si-based MOSFET: undoped GaInAs layers are weakly n-type, and

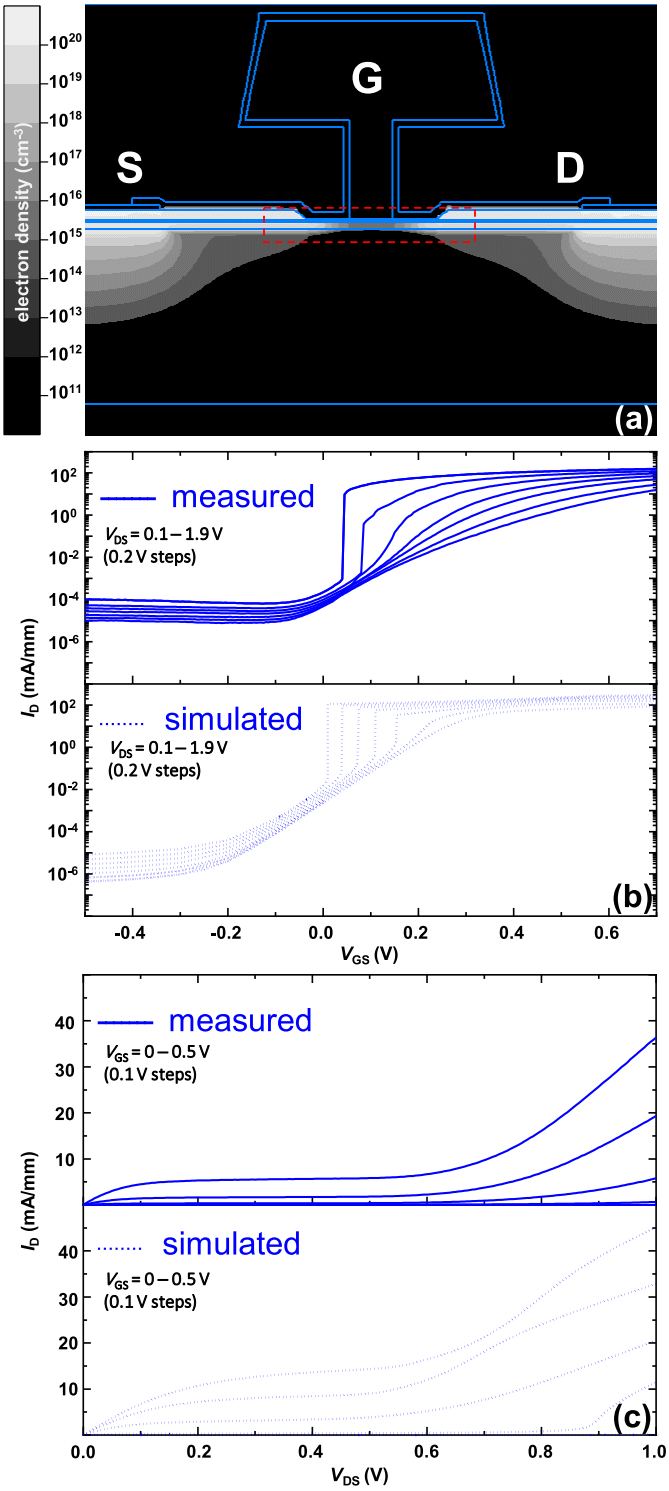

Fig. 4. 300 K TCAD simulation of the dc characteristics of a GaInAs I-MOS device with L_G = 100 nm at 300 K. (a) Simulated electron density in an unbiased device showing a maximum density of 1.21 × 10^{20} cm^{-3}. (b) Logarithmic scale plot of simulated (dotted traces) and measured (solid traces) OFF-to-ON switching characteristics for V_{DS} = 0.1–1.9 V (0.2 V steps). (c) Corresponding drain *I*–*V* characteristics with V_{GS} = 0–0.5 V (0.1 V steps). The red rectangle in (a) represents the simulation domain for the impact ionization rate, recombination rate, and hole density profiles shown in Fig. 5.

surface depletion effects suffice to deplete the channel due to the close proximity of the gate metal and channel. As such, the present devices operate similar to E-mode HEMTs. The simulated gray-scale curves in Fig. 1(e) illustrate gate control

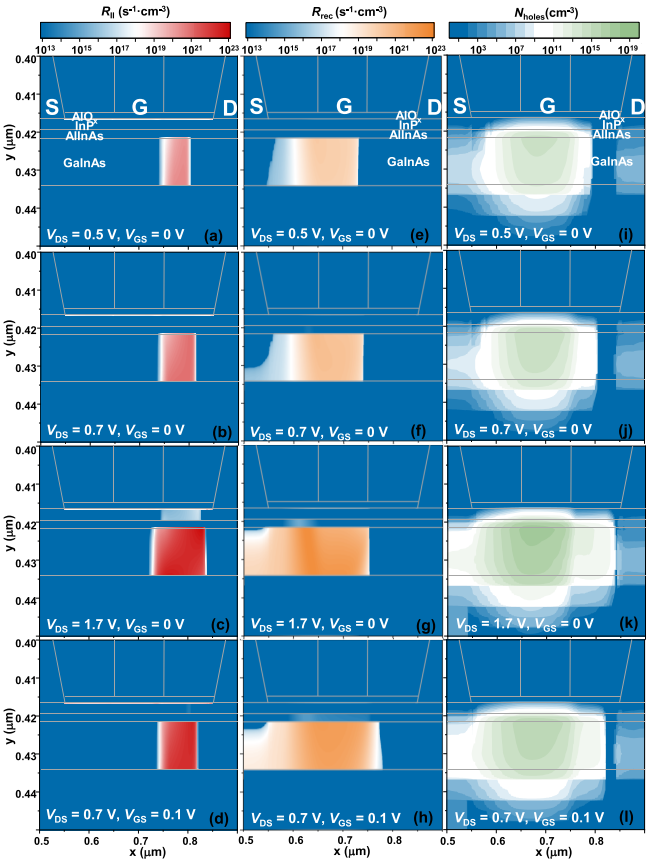

Fig. 5. 300 K TCAD simulation of GaInAs I-MOS operation with L_G = 100 nm. (a)–(d) Contour maps of impact ionization rate R_{II}. (e)–(h) Recombination rate R_{rec}. (i)–(l) Hole density N_{holes} in the device active region marked by red rectangle box in Fig. 4(a). Simulation biases: (a), (e), and (i) V_{DS} = 0.5 V and V_{GS} = 0 V; (b), (f), and (j) V_{DS} = 0.7 V and V_{GS} = 0 V. (c), (g), and (k) V_{DS} = 0.5 V and V_{GS} = 0 V. (d), (h), and (l) V_{DS} = 0.7 V and V_{GS} = 0.1 V. For all quantities, the color blue indicates negligible values.

of the I-MOS channel electron concentration under the gate. The measured I_{ON}/I_{OFF} > six orders of magnitude at 300 K [see Fig. 3(a)] confirms a very good E-mode charge control in the present I-MOS structure.

The simulations, however, underestimate impact ionization at low current levels and overestimate it at high current levels. The discrepancy between simulated and measured characteristics is attributed to the highly nonuniform electric field distribution along the channel, as well as to the limitations of the model for impact ionization rates. The use of a local electric field to model impact ionization phenomena in drift-diffusion simulations normally overestimates the impact ionization rate. This is because the model assumes that impact ionization rates depend solely on the local electric field, whereas carriers must travel a certain distance (the so-called dead-space distance) in an electric field to acquire sufficient energy for ionization [43]. Beyond nonlocality, experimental impact ionization rates also depend on the carrier history and on the electric field gradient [43]: as pointed out by McIntyre [43], Monte Carlo simulations are necessary to correctly calculate impact ionization rates in complex electric field profiles.

Fig. 4(a) shows a simulated electron density distribution in a device at equilibrium. The impact ionization generation rate R_{II} was computed for various device bias conditions. Fig. 5(a)–(d) shows the computed R_{II} in the GaInAs channel under biases of (V_{DS}, V_{GS}) = (0.5, 0 V), (0.7, 0 V), (1.7, 0 V), and (0.7, 0.1 V). Fig. 5 also shows the (i)–(l) hole density and (e)–(h) recombination rate in the device, also in logarithmic scale (where the color blue indicates negligible values in all cases). Simulation shows that ionization mostly occurs in the gate-to-drain space, well past the drain end of the gate electrode foot. Although generation takes place in the gate–drain space, holes drift back toward the source and mostly accumulate under the gate: the associated positive space charge allows more electrons to enter the channel (*i.e.*, parasitic bipolar effect [33]). The recombination rate R_{rec} is generally maximized near the source end of the gate. As to be expected from the range of applied voltages, impact ionization takes place within the GaInAs channel and not in the wider gap AlInAs/InP or Al_2O_3 layers. At (V_{DS}, V_{GS}) = (0.7, 0.1 V), impact ionization is much stronger, reaching $R_{II} = 2.17 \times 10^{22}$ s^{-1}·cm^{-3}, and causing *SS* to break through the limit of 60 mV/dec. In the case of extreme impact ionization at V_{DS} = 1.7 V and V_{GS} = 0 V, a maximum $R_{II} = 7.44 \times 10^{22}$ s^{-1}·cm^{-3} was reached, remaining above 10^{22} s^{-1}·cm^{-3} over most of the gate-to-drain space and enabling the super steep-slope switching characteristics. Although the simulations are not in strict quantitative agreement with experimentally measured results, they do provide a compelling illustration of GaInAs I-MOS operation.

IV. Conclusion

We reported the first fabrication and characterization of GaInAs-based I-MOS steep-slope transistors. The ALD passivation of an MBE-grown n-type GaInAs quantum well structure on InP with an ultrathin 1.5-nm Al_2O_3 oxide improves the electron mobility and the electron sheet density, effectively decreasing the channel sheet resistance by a factor of 2×. Such passivated quantum well channels were used to fabricate nanometric E-mode I-MOS transistors with pronounced steep-slope characteristics at both room and cryogenic temperatures. Steep-slope behavior was observed for drain biases $V_{DS} \geq 0.9$ V at room temperature. A record steep-slope *SS* = 1.25 mV/dec (limited by the instrument resolution) maintained over several orders of magnitude in drain current, with I_{ON}/I_{OFF} ratios of more than 10^6 (10^9) at 300 (15) K. The present GaInAs devices are the first I-MOS transistors to display a robust steep-slope effect at low voltages V_{DS} < 1.9 V at 300 K and <1 V at 15 K. The measured room temperature breakdown voltages <1.9 V with L_G = 100 nm which compare well to the 1.8 V simulated for L_G = 60 nm Ge I-MOS transistors proposed as low-voltage alternatives to Si I-MOS transistors [42]. More extensive benchmarking of the steep-slope device performance of GaInAs I-MOS transistors is presented in Part II, along with the characterization of their time-domain high-speed switching and RF properties.

References

[1] M. H. Woods, "MOS VLSI reliability and yield trends," *Proc. IEEE*, vol. 74, no. 12, pp. 1715–1729, Dec. 1986, doi: 10.1109/PROC.1986.13687.

[2] B. J. Benschneider *et al.*, "A pipelined 50-MHz CMOS 64-bit floating-point arithmetic processor," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1317–1323, Oct. 1989, doi: 10.1109/JSSC.1989.572606.

[3] D. W. Dobberpuhl *et al.*, "A 200-MHz 64-b dual-issue CMOS microprocessor," *IEEE J. Solid-State Circuits*, vol. 27, no. 11, pp. 1555–1567, Nov. 1992, doi: 10.1109/4.165336.

[4] P. Hofstee *et al.*, "A 1 GHz single-issue 64 b PowerPC processor," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, San Francisco, CA, USA, Feb. 2000, pp. 92–93, doi: 10.1109/ISSCC.2000.839705.

[5] G. Hinton *et al.*, "A 0.18-μm CMOS IA-32 processor with a 4-GHz integer execution unit," *IEEE J. Solid-State Circuits*, vol. 36, no. 11, pp. 1617–1627, Nov. 2001, doi: 10.1109/4.962281.

[6] D. Deleganes, J. Douglas, B. Kommandur, and M. Patyra, "Designing a 3 GHz, 130 nm, Intel ® Pentium® 4 processor," in *Symp. VLSI Circuits, Dig. Tech. Papers*, Honolulu, HI, USA, Jun. 2002, pp. 130–133, doi: 10.1109/VLSIC.2002.1015065.

[7] S. Vangal *et al.*, "5-GHz 32-bit integer execution core in 130-nm dual-V_T CMOS," *IEEE J. Solid-State Circuits*, vol. 37, no. 11, pp. 1421–1432, Nov. 2002, doi: 10.1109/JSSC.2002.803944.

[8] E. Pop, S. Sinha, and K. E. Goodson, "Heat generation and transport in nanometer-scale transistors," *Proc. IEEE*, vol. 94, no. 8, pp. 1587–1601, Aug. 2006, doi: 10.1109/JPROC.2006.879794.

[9] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE J. Solid-State Circuits*, vol. SSC-9, no. 5, pp. 256–268, Oct. 1974, doi: 10.1109/JSSC.1974.1050511.

[10] D. A. Antoniadis, A. Wei, and A. Lochtefeld, "SOI devices and technology," in *Proc. 29th ESSDERC*, Leuven, Belgium, Sep. 1999, pp. 81–87. [Online]. Available: https://ieeexplore.ieee.org/document/1505452

[11] J. W. Joyner and J. D. Meindl, "Opportunities for reduced power dissipation using three-dimensional integration," in *Proc. Int. Interconnect Technol. Conf.*, Burlingame, CA, USA, Jun. 2002, pp. 148–150, doi: 10.1109/IITC.2002.1014915.

[12] F. Faggin, "Trends in microcomputers," *Microprocessing Microprogram.*, vol. 5, no. 6, pp. 344–349, 1979, doi: 10.1016/0165-6074(79)90078-4.

[13] R. Chau *et al.*, "A 50 nm depleted-substrate CMOS transistor (DST)," in *IEDM Tech. Dig.*, Washington, DC, USA, Dec. 2001, pp. 29.1.1–29.1.4, doi: 10.1109/IEDM.2001.979585.

[14] K. Gopalakrishnan, P. B. Griffin, and J. D. Plummer, "I-MOS: A novel semiconductor device with a subthreshold slope lower than kT/q," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2002, pp. 289–292, doi: 10.1109/IEDM.2002.1175835.

[15] A. S. G. Andrae and T. Edler, "On global electricity usage of communication technology: Trends to 2030," *Challenges*, vol. 6, no. 1, pp. 117–157, 2015, doi: 10.3390/challe6010117.

[16] T. T. J. Neuffer, "Decadal plan for semiconductors," Semiconducot Res. Corp., Durham, NC, USA, Tech. Rep., 2021, pp. 122–123. [Online]. Available: https://www.src.org/about/decadal-plan/

[17] W. Haensch *et al.*, "Silicon CMOS devices beyond scaling," *IBM J. Res. Develop.*, vol. 50, nos. 4–5, pp. 339–361, Jul. 2006, doi: 10.1147/rd.504.0339.

[18] E.-H. Toh *et al.*, "Impact ionization nanowire transistor with multiple-gates, silicon-germanium impact ionization region, and sub-5 mV/decade subtheshold swing," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2007, pp. 195–198, doi: 10.1109/IEDM.2007.4418900.

[19] Y. Khatami and K. Banerjee, "Steep subthreshold slope n- and p-type tunnel-FET devices for low-power and energy-efficient digital circuits," *IEEE Trans. Electron Devices*, vol. 56, no. 11, pp. 2752–2761, Nov. 2009, doi: 10.1109/TED.2009.2030831.

[20] S. Salahuddin and S. Datta, "Use of negative capacitance to provide voltage amplification for low power nanoscale devices," *Nano Lett.*, vol. 8, no. 2, pp. 405–410, 2007, doi: 10.1021/nl071804g.

[21] J. Frougier *et al.*, "Phase-transition-FET exhibiting steep switching slope of 8 mV/decade and 36% enhanced ON current," in *Proc. IEEE Symp. VLSI Technol.*, Honolulu, HI, USA, Jun. 2016, pp. 1–2, doi: 10.1109/VLSIT.2016.7573445.

[22] C. H. Ahn, J.-M. Triscone, and J. Mannhart, "Electric field effect in correlated oxide systems," *Nature*, vol. 424, no. 6952, pp. 1015–1018, 2003, doi: 10.1038/nature01878.

[23] H. Kam, D. T. Lee, R. T. Howe, and T.-J. King, "A new nano-electro-mechanical field effect transistor (NEMFET) design for low-power electronics," in *IEDM Tech. Dig.*, Washington, DC, USA, Dec. 2005, pp. 463–466, doi: 10.1109/IEDM.2005.1609380.

[24] Z. Tang *et al.*, "A steep-slope MoS_2/graphene dirac-source field-effect transistor with a large drive current," *Nano Lett.*, vol. 21, no. 4, pp. 1758–1764, Feb. 2021, doi: 10.1021/acs.nanolett.0c04657.

[25] C. Qiu *et al.*, "Dirac-source field-effect transistors as energy-efficient, high-performance electronic switches," *Science*, vol. 361, no. 6400, pp. 387–392, Jul. 2018, doi: 10.1126/science.aap9195.

[26] W. Y. Choi, J. Y. Song, J. D. Lee, Y. J. Park, and B.-G. Park, "70-nm impact-ionization metal-oxide-semiconductor (I-MOS) devices integrated with tunneling field-effect transistors (TFETs)," in *IEDM Tech. Dig.*, Washington, DC, USA, Dec. 2005, pp. 955–958, doi: 10.1109/IEDM.2005.1609519.

[27] S. Ramaswamy and M. J. Kumar, "Junctionless impact ionization MOS: Proposal and investigation," *IEEE Trans. Electron Devices*, vol. 61, no. 12, pp. 4295–4298, Dec. 2014, doi: 10.1109/TED.2014.2361343.

[28] K. Gopalakrishnan, R. Woo, C. Jungemann, P. B. Griffin, and J. D. Plummer, "Impact ionization MOS (I-MOS)—Part II: Experimental results," *IEEE Trans. Electron Devices*, vol. 52, no. 1, pp. 77–84, Jan. 2005, doi: 10.1109/TED.2004.841345.

[29] K. Gopalakrishnan, P. B. Griffin, and J. D. Plummer, "Impact ionization MOS (I-MOS)—Part I: Device and circuit simulations," *IEEE Trans. Electron Devices*, vol. 52, no. 1, pp. 69–76, Jan. 2005, doi: 10.1109/TED.2004.841344.

[30] A. Savio, S. Monfray, C. Charbuillet, and T. Skotnicki, "On the limitations of silicon for I-MOS integration," *IEEE Trans. Electron Devices*, vol. 56, no. 5, pp. 1110–1117, May 2009, doi: 10.1109/TED.2009.2015163.

[31] T. Suemitsu, T. Enoki, and Y. Ishii, "Body contacts in InP-based InAlAs/InGaAs HEMTs and their effects on breakdown voltage and kink suppression," *Electron. Lett.*, vol. 31, no. 9, pp. 758–759, Apr. 1995, doi: 10.1049/el:19950496.

[32] C. R. Bolognesi, M. W. Dvorak, and D. H. Chow, "Impact ionization suppression by quantum confinement: Effects on the DC and microwave performance of narrow-gap channel InAs/AlSb HFET's," *IEEE Trans. Electron Devices*, vol. 46, no. 5, pp. 826–832, May 1999, doi: 10.1109/16.760386.

[33] B. J. V. Zeghbroeck, W. Patrick, H. Meier, and P. Vettiger, "Parasitic bipolar effects in submicrometer GaAs MESFET's," *IEEE Electron Device Lett.*, vol. EDL-8, no. 5, pp. 188–190, May 1987, doi: 10.1109/EDL.1987.26598.

[34] D. C. Ruiz, T. Saranovac, D. Han, O. Ostinelli, and C. R. Bolognesi, "Impact ionization control in 50 nm low-noise high-speed InP HEMTs with InAs channel insets," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2019, pp. 9.3.1–9.3.4, doi: 10.1109/IEDM19573.2019.8993654.

[35] D. Han, D. C. Ruiz, G. Bonomo, T. Saranovac, O. J. S. Ostinelli, and C. R. Bolognesi, "Low-noise microwave performance of 30 nm GaInAs MOS-HEMTs: Comparison to low-noise HEMTs," *IEEE Electron Device Lett.*, vol. 41, no. 9, pp. 1320–1323, Sep. 2020, doi: 10.1109/LED.2020.3012017.

[36] J. Robertson and B. Falabretti, "Band offsets of high K gate oxides on III–V semiconductors," *J. Appl. Phys.*, vol. 100, no. 1, Jul. 2006, Art. no. 014111, doi: 10.1063/1.2213170.

[37] K. Xu *et al.*, "Band offset determination of atomic-layer-deposited Al_2O_3 and HfO_2 on InP by internal photoemission and spectroscopic ellipsometry," *J. Appl. Phys.*, vol. 113, no. 2, Jan. 2013, Art. no. 024504, doi: 10.1063/1.4774038.

[38] J. Franco *et al.*, "Suitability of high-k gate oxides for III–V devices: A PBTI study in $In_{0.53}Ga_{0.47}As$ devices with Al_2O_3," in *Proc. IEEE Int. Rel. Phys. Symp.*, Waikoloa, HI, USA, Jun. 2014, pp. 6A.2.1–6A.2.6, doi: 10.1109/IRPS.2014.6861098.

[39] M. Radosavljevic *et al.*, "Advanced high-K gate dielectric for high-performance short-channel $In_{0.7}Ga_{0.3}As$ quantum well field effect transistors on silicon substrate for low power logic applications," in *IEDM Tech. Dig.*, Washington, DC, USA, Dec. 2009, pp. 319–322, doi: 10.1109/IEDM.2009.5424361.

[40] Y. Li *et al.*, "Interface and electrical properties of buried InGaAs channel MOSFET with an InP barrier layer and $Al_2O_3/HfO_2/Al_2O_3$ gate dielectrics," *Appl. Phys. Exp.*, vol. 13, no. 1, Jan. 2020, Art. no. 011004, doi: 10.7567/1882-0786/ab5acf.

[41] L. Södergren, N. S. Garigapati, M. Borg, and E. Lind, "Mobility of near surface MOVPE grown InGaAs/InP quantum wells," *Appl. Phys. Lett.*, vol. 117, no. 1, Jul. 2020, Art. no. 013102, doi: 10.1063/5.0006530.

[42] E.-H. Toh *et al.*, "A novel CMOS compatible L-shaped impact-ionization MOS (LI-MOS) transistor," in *IEDM Tech. Dig.*, Washington, DC, USA, Dec. 2005, pp. 951–954, doi: 10.1109/IEDM.2005.1609518.

[43] R. J. McIntyre, "A new look at impact ionization—Part I: A theory of gain, noise, breakdown probability, and frequency response," *IEEE Trans. Electron Devices*, vol. 46, no. 8, pp. 1623–1631, Aug. 1999, doi: 10.1109/16.777150.