Performance Analysis of Ultra-Scaled InAs HEMTs

Neerav Kharche[1], Gerhard Klimeck[1], Dae-Hyun Kim[2,3], Jesús. A. del Alamo[2], and Mathieu Luisier[1]
[1]Network for Computational Nanotechnology, Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907;
[2]Microsystems Technology Labs, Massachusetts Institute of Technology, Cambridge, MA 02139;
[3]Teledyne Scientific & Imaging, LLC, Thousand Oaks, CA, 91360;
Email: nkharche@purdue.edu, Ph: (765) 494-9050

Abstract

The scaling behavior of ultra-scaled InAs HEMTs is investigated using a 2-dimensional real-space effective mass ballistic quantum transport simulator. The simulation methodology is first benchmarked against experimental I_d-V_{gs} data obtained from devices with gate lengths ranging from 30 to 50 nm, where a good quantitative match is obtained. It is then applied to optimize the logic performance of not-yet-fabricated 20nm InAs HEMT. It is demonstrated that the best performance is achieved in thin InAs channel devices by reducing the insulator thickness to improve the gate control while increasing the gate work function to suppress the gate leakage.

Introduction

III-V High Electron Mobility Transistors (HEMTs) have emerged as potential candidates for high-speed, low-power logic applications beyond Si-CMOS technology [1-5]. Device simulation may help experimentalists scale the gate length of HEMTs below 30 nm to keep following Moore's law. However, performances predictions of not-yet-fabricated transistors require a simulator that reproduces available experimental data. We demonstrate very good quantitative agreement of our simulations with experimental data [4] on InAs HEMTs with gate lengths ranging from 30 to 50 nm. As opposed to other quantum transport tools our real-space simulator captures both the OFF- and the ON-state currents of transistors by injecting carriers from all contacts [6].

The calibration of the simulator with experimental data at different gate lengths allows us to study the scalability of InAs HEMTs. We find that (i) such devices are perfectly viable at a gate length of 20nm, (ii) source-to-drain tunneling remains manageable at 20nm despite the low InAs effective mass, (iii) the InAlAs dielectric layer separating the multi-quantum-well InGaAs-InAs-InGaAs channel from the gate contact should be reduced to 3nm to maintain a good electrostatic control as well as low subthreshold slopes and DIBL, (iv) the metal work function of the gate contact should be increased to keep the gate leakage current small and to push the threshold voltage to positive values, and (v) the InAs channel thickness should be reduced to improve the short channel characteristics. By design optimizations we demonstrate a 20nm InAs HEMT design with S = 87 mV/dec, DIBL = 110 mV/V, I_{ON} =372 A/m, I_{ON}/I_{OFF} = 2.9×10^3, and $g_{m,max}$ = 1.7 mS/μm.

Approach

InAs HEMTs with gate lengths of 20, 30, 40, and 50 nm are considered (Fig.1, similar to Ref. [4]). The simulation domain is restricted to the gate contact region and an extension L_{side} of 50nm on each side to reduce the computational burden. The source and drain contacts are modeled via two series resistances R_S and R_D that dominate in the ON-state regime of the device [7]. The intrinsic device terminal voltages are reduced from the external voltages according to $\tilde{V}_{gs} = V_{gs} - R_S I_d$ and $\tilde{V}_{ds} = V_{ds} - (R_S + R_D) I_d$.

The device and the gate contact are treated as a single entity on a quantum mechanical level. Real-space Schrödinger and Poisson equations are solved self-consistently using the

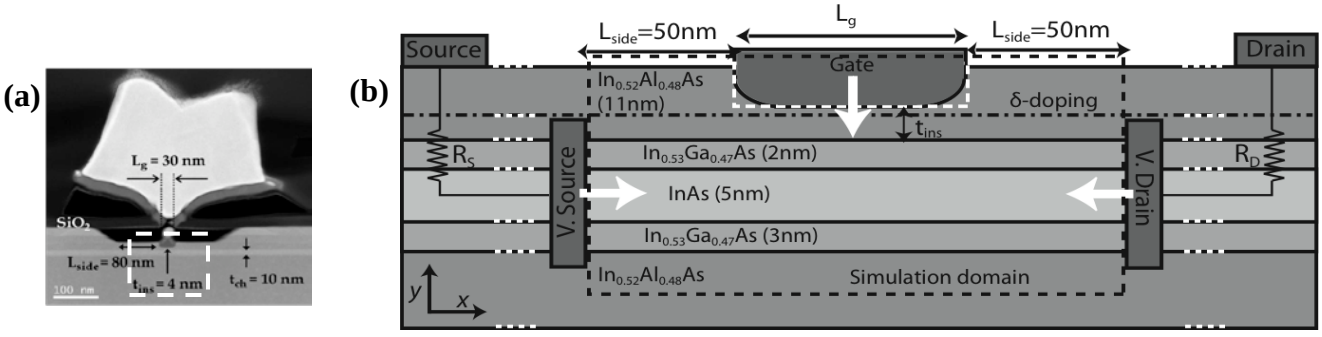

Fig. 1: (a) TEM image of InAs HEMT fabricated in Ref. [4]. The device region enclosed in the dotted white rectangle is included in the simulation domain. (b) Schematic view of the InAs HEMT. The channel region is composed of a 10 nm InGaAs/InAs/InGaAs Multi-Quantum-Well grown on a thick $In_{0.52}Al_{0.48}As$ layer lattice-matched to the InP substrate. $In_{0.52}Al_{0.48}As$ layer below the gate contact acts as an insulator. Black dash dotted line represents the δ-doped layer of concentration $N_D=3\times10^{12}cm^{-2}$. The dashed black rectangle encloses the simulation domain, which is restricted to the gate contact region and the extension of L_{side}=50nm on source/drain sides. The source/drain extensions beyond virtual contacts are modeled by two series resistances R_S and R_D, respectively. Two gate contact geometries curved (black) and flat (dashed white) are investigated.

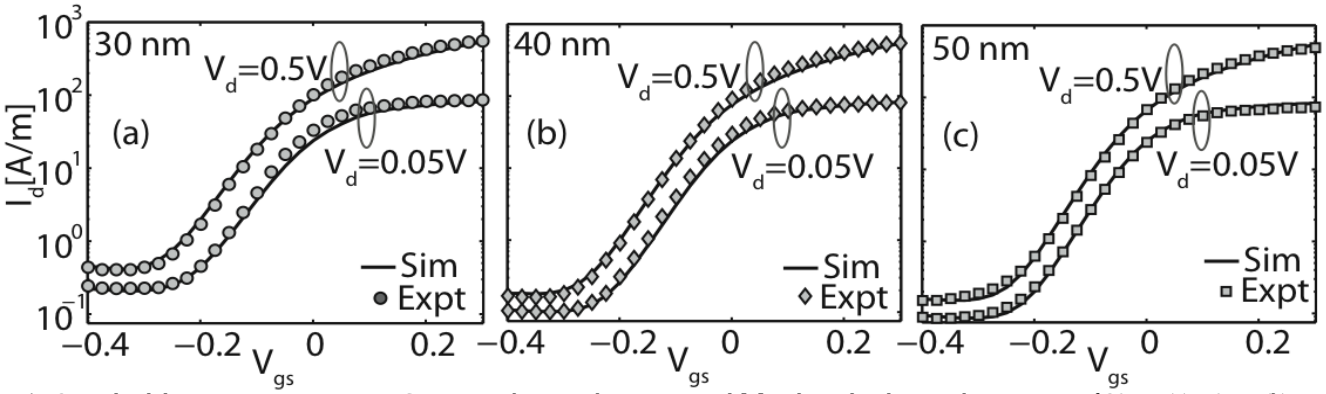

Fig. 2: Tool validation against experiments: Comparison between the experimental [4] and simulated I_d-V_{gs} characteristics of 30 nm (a), 40 nm (b), and 50 nm (c) gate length InAs HEMTs. Symbols refer to experimental data and solid curves refer to the simulations. Reported gate length (L_g) and insulator thickness (t_{ins}) of the experimental devices are (30 nm, 4 nm), (40 nm, 4 nm), and (50 nm, 4 nm). These values are slightly varied within experimental uncertainties to match the experimental data. (L_g, t_{ins}) values of the simulated devices are (34 nm, 3.6 nm), (42 nm, 3.8 nm) and (51 nm, 4.0 nm) respectively. Simulated devices have curved gate contacts, which resemble closely to the gate contact geometries resulting from the wet chemical etching step in the gate stack fabrication process. The source/drain series resistance values used in the simulation are R_S = 0.21 Ω·m and R_D = 0.23 Ω·m.

L_g [nm]		S [mV/dec]	DIBL [mV/V]	I_{ON}/I_{OFF}
30	Expt.	106.9	168.9	0.47×10^3
	Sim.	105.2	144.7	0.61×10^3
40	Expt.	90.9	126.0	1.38×10^3
	Sim.	89.4	99.3	1.86×10^3
50	Expt.	85.1	97.2	1.80×10^3
	Sim.	89.2	90.8	1.85×10^3

Fig. 3: Device parameters from simulated and experimental devices [4]. The V_T used in DIBL calculation is the V_{gs} that yields I_d=1 A/m.

effective mass approximation and a 2-D finite-difference grid [6]. Neumann boundary conditions are applied everywhere to the Poisson domain except at the gate contact where Dirichlet boundary conditions are applied. In the ballistic transport model used here, electrons are injected into the device at different wave-vector and energy values and the resulting contributions are summed up to give carrier and current densities.

The effective masses of the biaxially strained InAs channel are extracted from a tight-binding bandstructure calculation which includes InGaAs and InAlAs buffer layers Thus, important bandstructure effects due to band non-parabolicity, confinement, and biaxial strain in InAs channel are captured through modified effective masses.

Results

Benchmarking at Large Gate Lengths: The experimental transfer characteristics I_d-V_{gs} at V_{ds}=0.05 V and V_{ds}=0.5 V of InAs HEMTs with gate lengths of 30, 40, and 50 nm are compared to simulation results in Fig. 2. To obtain a good agreement between simulation and measurement, the device dimensions such as gate length, insulator thickness, and metalwork function are slightly modified within experimental uncertainties. The performance parameters calculated from

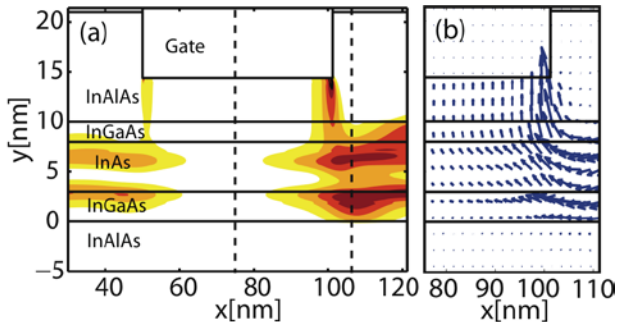

Fig. 4: Flat gate-insulator interface is formed when anisotropic etching or metal gate sinking is used to thin down the insulator. Fig. (a) shows the contour plot of the magnitude of the current in OFF-state (V_{gs}=-0.4V and V_{ds}=0.5V) where tunneling current is dominated by edge-leakage. Fig. (b) shows the direction of the current flow.

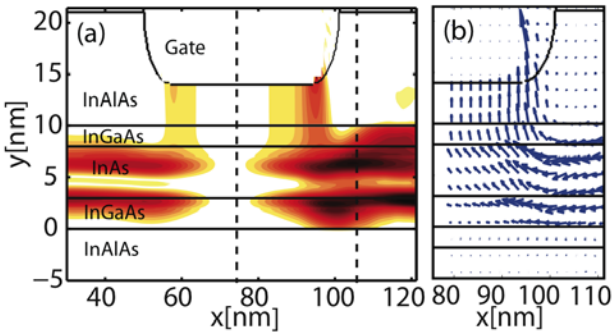

Fig. 5: Curved gate-insulator interface is formed when isotropic wet etching is used to thin down the insulator. Figs. (a) and (b) show the contour plot of the magnitude of the current and the direction of the current flow, respectively in the OFF-state (V_{gs}=-0.4V and V_{ds}=0.5V).

the experimental and simulated I_d-V_{gs} show a reasonable agreement as shown in the table in Fig. 3. An accurate description of the gate contact shape strongly influences gate leakage current distributions (Fig. 4 and 5) and is critical to reproduce the experimental I_d-V_{gs}, especially in the

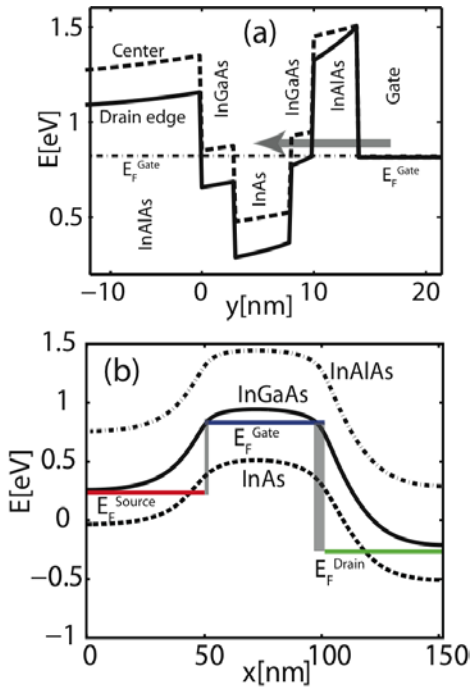

Fig. 6: Edge leakage mechanism: (a) Band diagram along the vertical lines through the center and the drain side edge of the gate contact of a HEMT in Fig. 4. As identified by the gray arrow, the gate leakage current arises due to the electron tunneling through the InAlAs insulator and InGaAs barrier layers between the gate contact and the InAs channel. Electrons at the edges of the gate contact encounter only the InAlAs tunnel barrier while the electrons at the center encounter an additional InGaAs tunnel barrier. Due to this gate-leakage current is concentrated at the edges of the gate contact. The high gate leakage paths are identified by the gray shaded rectangles in the band diagram along the channel direction in Fig. (b).

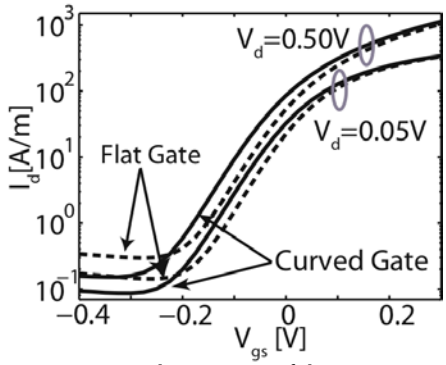

Fig. 7: Intrinsic I_d-V_{gs} characteristics of the L_g=51nm InAs HEMTs with flat (dotted lines) and curved (solid lines) gate-insulator interface. Both devices perform similarly in the ON-state regime, however, flat gate device exhibits higher gate-leakage in the subthreshold regime.

subthreshold regime. The edge-leakage mechanism of current crowding at the edges of the gate contact is explained in Fig. 6. A square or a curved gate contact geometry acts differently on the current magnitude as illustrated in Fig. 6. The fits in Fig. 2 have been enabled by the consideration of a curved contact geometry, characterized by a thicker dielectric layer at

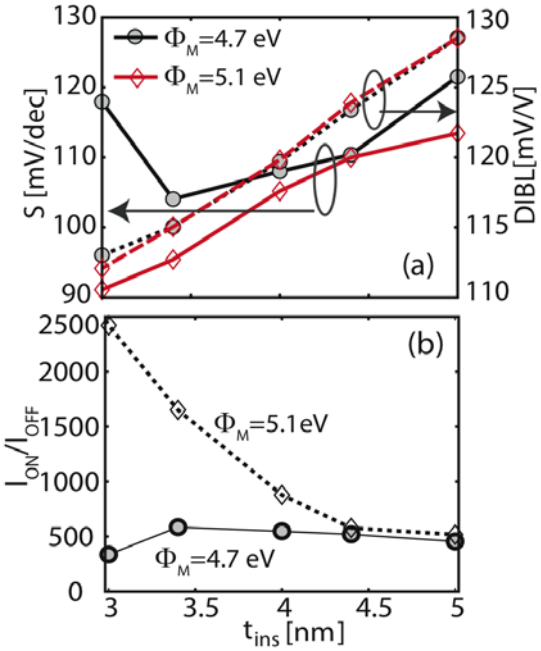

Fig. 8: InAlAs insulator thickness scaling: For metal work function Φ_M = 4.7 eV, the subthreshold slope (S) and I_{ON}/I_{OFF} ratio improves as t_{ins} decreases till 3.4 nm. Devices with thinner insulator suffer from excessive gate leakage which leads to high S and low I_{ON}/I_{OFF}. This degradation can be controlled by increasing the Φ_M to 5.1 eV. DIBL improves as t_{ins} is reduced. Effects of Φ_M and t_{ins} on the device performance are highly correlated. Device dimensions are same as in Fig. 1 with L_g = 20 nm and t_{ins} = 4 nm.

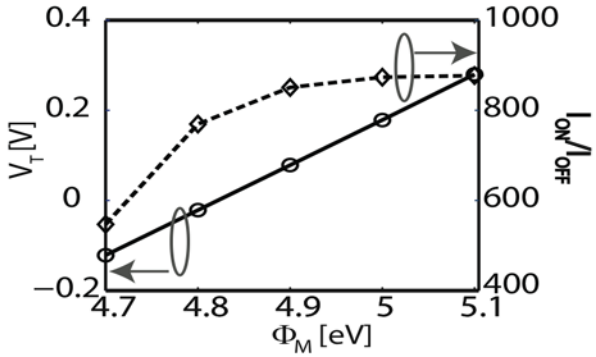

Fig. 9: Gate metal work-function (Φ_M) engineering: Higher Φ_M results in higher threshold voltage and high I_{ON}/I_{OFF} ratio. Variation in I_{ON} with Φ_M is negligible. The high I_{ON}/I_{OFF} ratio is achieved mainly because I_{OFF} is significantly suppressed in the devices with higher Φ_M because of higher gate to channel tunneling barriers. Device dimensions are same as in Fig. 1 with L_g = 20 nm and t_{ins} = 4 nm.

its edges and smaller leakage currents. The curved geometry is a result of the isotropic wet chemical etching process used to thin down the InAlAs insulator layer. Not only do we match experiments quantitatively but we provide device metrology [8] through simulation.

Design exploration of a 20nm InAs HEMT: After benchmarking our simulator to the "large" experimental structures we explore the performance of devices with a 20

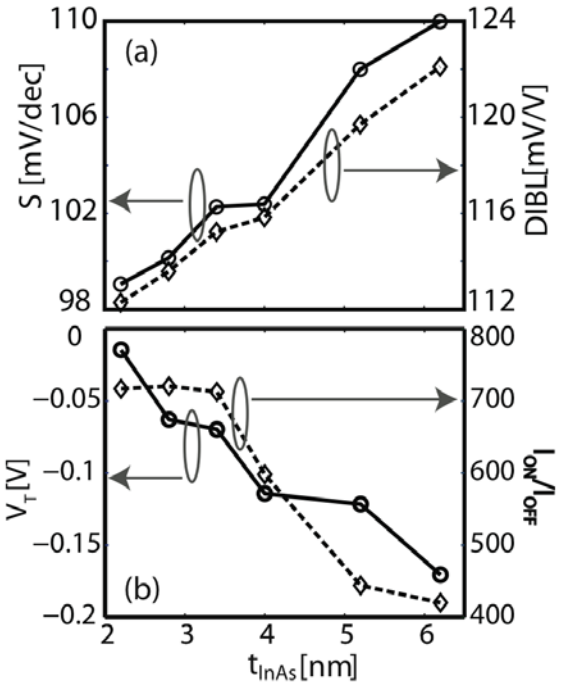

Fig. 10: InAs quantum well (QW) thickness scaling: (a) Subthreshold slope and DIBL improve as the thickness of the InAs QW in the channel is reduced. The 2D electron gas in the channel is located closer to the gate in thinner QW devices which results in stronger gate control and improved short channel characteristics. (b) Devices with thinner QW have higher V_T as a result of the stronger quantum confinement in the channel. High I_{ON}/I_{OFF} is ratio is achieved as a result of suppressed I_{OFF} due to higher gate electron tunneling barriers in thin QW devices. Device dimensions except InAs channel thickness are same as in Fig. 1 with L_g = 20 nm and t_{ins} = 4 nm. Metal work function is Φ_M = 4.7 eV.

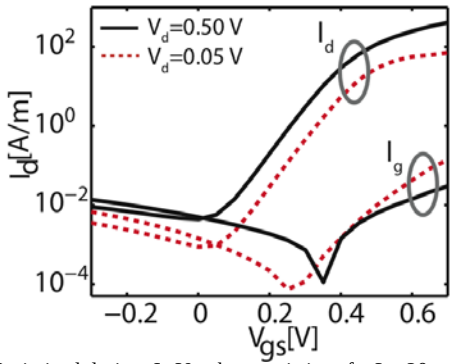

Fig. 11: Optimized design: I_d-V_{gs} characteristics of a L_g=20 nm InAs HEMT with t_{ins}=3 nm, Φ_M = 5.1 eV, and t_{InAs} = 3.4 nm. Performance parameters of this device are S = 87 mV/dec, DIBL = 110 mV/V, $I_{ON}/I_{OFF} = 2.9\times10^3$, and $g_{m,max}$ = 1.7 mS/μm.

nm gate length. A flat gate contact geometry is used because it providests a better gate control as compared to the curved contact and it can be realized with slight fabrication modifications [1]. The threshold voltage (V_T) is determined from linear extrapolation of I_d-V_{gs} at the peak transconductance to zero I_d (maximum-g_m method). The ON-state is defined to be $V_g=V_T+2V_{DD}/3$, $V_d=V_{DD}$, $V_s=0$ while the OFF-state is defined to be $V_g=V_T-V_{DD}/3$, $V_d=V_{DD}$, $V_s=0$, where V_{DD}=0.5 V. The effect of InAlAs insulator thickness (t_{ins}) on the HEMT performance is depicted in Fig. 8. For devices with a metal work function (Φ_M) of 4.7 eV the subthreshold slope (S) degrades as the InAlAs layer becomes thinner than 3.4 nm due to an increase of the gate leakage current. This degradation can be controlled by increasing Φ_M to 5.1 eV, which increases the tunneling barrier heights between the gate and the InAs channel, reduces the gate leakage, and therefore increases the I_{ON}/I_{OFF} ratio, even with an InAlAs layer scaled down to 3 nm (Fig. 8 and 9). Enhancement-mode devices can be realized by increasing Φ_M to shift V_T to a positive value which is essential in CMOS logic applications (Fig. 9). The effect of t_{ins} and Φ_M on the device performance are highly correlated and both these parameters should be optimized together to improve device performance. The InAs quantum well thickness also plays a crucial role to improve the short channel behavior because it affects confinement and gate control of carriers in the channel. The subthreshold slope, Drain-Induced-Barrier-Lowering (DIBL), and I_{ON}/I_{OFF} ratio improve as the InAs channel thickness is reduced (Fig. 10). Finally, Fig. 11 demonstrates that a very good short channel performance and enhancement-mode operation can be achieved if a thin InAlAs insulator (t_{ins}=3nm) is combined with a thin InAs channel (t_{InAs}=3.4nm), and a high metal work function (Φ_M=5.1eV).

Conclusion

We have developed simulation tools and methodologies for the analysis, metrology, and optimization of ultra-scaled III-V HEMTs. Gate tunneling is found to be critical in the device analysis. The accurate description of the shape of the gate contact is crucial to replicate the experimental results. The scaling study on 20 nm InAs HEMT suggests that best performance can be achieved in thin InAs channel devices by reducing the insulator thickness to improve the gate control while increasing the gate work function to suppress the gate leakage. We released the simulation tool *OMEN_FET* that generates these results on nanoHUB.org.

Acknowledgements

This work was partially supported by Semiconductor Research Corporation (SRC) and MARCO Focus Center on Materials, Structures, and Devices. Computational support was provided by NSF through Network for Computational Nanotechnology (NCN) on nanoHUB.org and by the National Institute for Computational Sciences (NICS).

References

[1] D. H. Kim et al., *IEDM 2008*, p. 719. [2] Chien-I Kuo et *al*., *Elec. Dev. Lett.* 29, 290 (2008). [3] D. H. Kim et *al*., *IEDM 2007*, p. 629. [4] D. H. Kim et *al*., *Elec. Dev. Lett.* 29, 830 (2008). [5] M. K. Hudait et al. *IEDM 2007*, p. 627. [6] M. Luisier et al., *IEEE Trans. on Elec. Dev.* 55, 1494, 2008. [7] M. Luisier et *al*., *IEDM 2008*, p. 887. [8] M. Lundstrom et *al*., AIP Conf. Proc. 931, 563 (2007). [9] N Kharche, et *al*. "OMEN_FET," http://nanohub.org/contribtool/omenhfet.