Received 10 June 2020; revised 21 July 2020; accepted 8 August 2020. Date of publication 18 August 2020; date of current version 15 September 2020.
The review of this article was arranged by Editor M. Chan.

Digital Object Identifier 10.1109/JEDS.2020.3017141

$In_{0.53}Ga_{0.47}As$/InAs Composite Channel MOS-HEMT Exhibiting 511 GHz f_τ and 256 GHz f_{max}

BRIAN MARKMAN (Member, IEEE), SIMONE TOMMASO ŠURAN BRUNELLI, ARANYA GOSWAMI, MATTHEW GUIDRY (Member, IEEE), AND MARK J. W. RODWELL (Fellow, IEEE)

Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA

CORRESPONDING AUTHOR: B. MARKMAN (e-mail: brianmarkman@ucsb.edu)

This work was supported in part by the Semiconductor Research Corporation (SRC) and DARPA through the JUMP Program, and in part by the UCSB HFM Lab with instrumentation supported by ONR (Dr. P. Maki) and ARO (Dr. J. Harvey) DURIP Grants.

ABSTRACT An $In_{0.53}Ga_{0.47}As$/InAs composite channel MOS-HEMT exhibiting peak $f_\tau = 511$ GHz and peak $f_{max} = 285$ GHz is demonstrated. Additionally, another device exhibiting peak $f_\tau = 286$ GHz and peak $f_{max} = 460$ GHz is reported. The devices have a 1 nm / 3 nm $Al_xO_yN_z$ interfacial layer and ZrO_2 gate dielectric on a 2 nm / 4 nm $In_{0.53}Ga_{0.47}As$ / InAs composite channel with a modulation doped $In_{0.52}Al_{0.48}As$ back barrier. To reduce parasitic gate-source and gate-drain capacitances, a modulation doped $In_{0.52}Al_{0.48}As$ / $In_{0.53}Ga_{0.47}As$ / InAs composite quantum well is included between the gate edges and the N+ source and drain. Compared to the work of Wu *et al.*, addition of an $In_{0.52}Al_{0.48}As$ back-barrier, scaling of S/D metal spacing, and scaling of channel thickness has enabled improved transconductance and increased f_τ. Short gate length devices f_{max} are limit by high R_G due to poor metal filling of the T-Gate stem and large C_{DS} due to a conductive etch stop layer. Long gate length devices exhibit better metal filling, reduced R_G, and balanced f_τ, f_{max}. Devices exhibit 10-15% DC-1 GHz $g_{m,e}$ suggesting that the high-k / semiconductor interface has low D_{it}.

INDEX TERMS InAs, MOSFETs, MOSHEMTs, RF, f_τ, f_{max}.

I. INTRODUCTION

InP-based transistors are of interest for future high-frequency communication systems [2]–[3]. High f_τ is important for future low noise, mm-wave communication systems [4]. State-of-the-art InP based HEMTs exhibit 610 GHz f_τ [2] and 703 GHz f_τ, [5]. Further scaling of f_τ requires higher $g_{m,e}$ which requires larger gate-channel capacitance C_{CH}. Minimum $In_{0.52}Al_{0.48}As$ gate-insulator thickness is limited by gate leakage current; high-k gate dielectrics truncate thermionic leakage current and reduce tunneling current at a given thickness while increasing the dielectric permittivity in the gated region providing a path forward. We report record $f_\tau = 511$ GHz for MOS-HEMT technology [1], [6]. While this technology has yet to surpass the maximum reported f_τ of standard InP-based HEMTs [7], improvements in the access region design, optimization of channel design [8], and further scaling of the gate dielectric can further increase $g_{m,e}$. Specifically, VLSI-optimized III-V MOSFETs have demonstrated extremely high $g_{m,e}$ of 3.0 mS/μm [9] and 3.45 mS/μm [10], even at the relatively small $V_{DS} = 0.5$ V and small $(V_{GS} - V_T)$ associated with VLSI operation; larger $g_{m,e}$ would be expected at larger voltages [1]. We report improvements to [1] achieved by including an $In_{0.52}Al_{0.48}As$ back-barrier (increase $g_{m,i}$) and reducing the source-drain metal spacing from 5 μm to 2 μm (decrease R_S). These improvements yielded an increase of DC $g_{m,e} = 1.5$ mS/μm to 2.2 mS/μm and an increase of $f_\tau = 357$ GHz to 511 GHz. Short-gate length devices f_{max} is limited by $R_G > 10\ \Omega$.

II. DEVICE FABRICATION

Devices were fabricated on a (100) semi-insulating InP substrate. The starting epitaxial structure, shown in Figure 1, was purchased from Intelligent Epitaxy. From substrate to

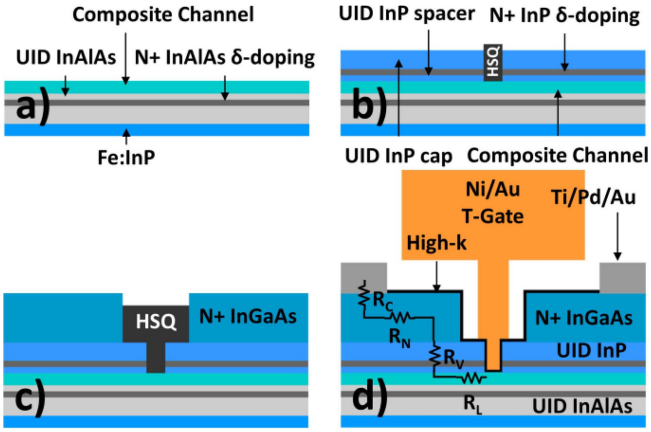

FIGURE 1. Schematic of the RF-MOSFET process flow: a) Buffer and channel epitaxy b) Dummy gate and modulation doped access region regrowth c) Second dummy gate and N+ source/drain contacts d) High-K deposition, T-gate metallization, and source/drain metallization.

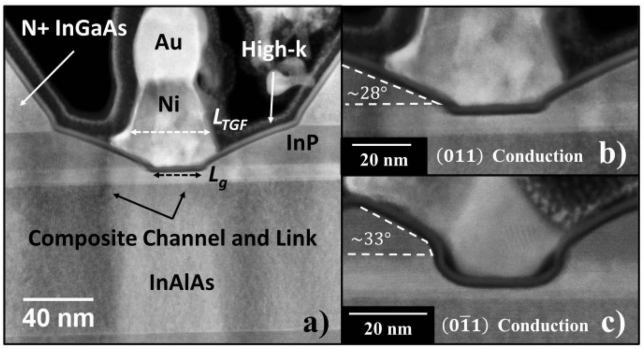

FIGURE 2. TEM cross-section of a) $L_g = 22$ nm device and faceting exhibited for b) (011) conduction device c) (0$\bar{1}$1) conduction device (separate sample).

air the grown layers are: 200 nm UID- $In_{0.52}Al_{0.48}As$ buffer, 20 nm UID-InP etch stop, 100 nm UID-$In_{0.52}Al_{0.48}As$, 3 nm modulation doped Si:$In_{0.52}Al_{0.48}As$ ($1 \times 10^{19} cm^{-3}$), 3 nm UID-$In_{0.52}Al_{0.48}As$ spacer, 3 nm / 4 nm / 5 nm UID-$In_{0.53}Ga_{0.47}As$ / strained UID-InAs / UID-$In_{0.53}Ga_{0.47}As$ composite channel. The $In_{0.52}Al_{0.48}As$ has a larger conduction band-offset (CBO) than InP to the $In_{0.53}Ga_{0.47}As$ / InAs. Larger CBO enables more ($E_F - E_1$) beneath the gate, corresponding to higher ($V_{GS} - V_{TH}$) and increased ballistic $g_{m,i}$ [11].

The gate recess was defined using an electron beam lithography (EBL) exposure of hydrogen silsesquioxane (HSQ). The lithographically defined gate recess defines the minimum metallurgical gate length, as illustrated in Figure 2. Gate lengths of devices to be discussed are defined with respect to the lithographically defined gate length, drawn in HSQ, as well as the direction of conduction: Device 1 is $L_g = 8$ nm (0$\bar{1}$1) conduction, Device 2 is $L_g = 22$ nm (011) conduction, and Device 3 is $L_g = 90$ nm (011) conduction.

After developing HSQ, one cycle of digital etching in HCl:H_2O 1:10 was done immediately before loading into the metal organic chemical vapor deposition (MOCVD) chamber. A modulation doped InP link region was then grown at 600°C: 3 nm UID-InP spacer, 2 nm Si:InP ($1 \times 10^{19} cm^{-3}$) modulation doping, 15 nm UID-InP cap. Hall measurements were performed on simultaneously processed samples without dummy gates and yielded electron sheet carrier density and mobility of $n_L = 2.5 \times 10^{12}$ cm^{-2} and $\mu_L = 11,000$ cm^2/Vs. The high mobility suggests that the strained InAs is not relaxed. The sheet carrier concentration is lower than that reported in [1] but the mobility doubling results in similar link region sheet resistivity (ρ_L). The first dummy gate is then stripped in BHF and the second dummy gate defined by the same HSQ EBL process. The second dummy gate is made to be 100 nm larger than the first, providing symmetric 50 nm gate-source and gate-drain recesses. A single digital etch in HCl:H_2O 1:10 was again performed immediately prior to re-loading into the MOCVD chamber and re-growing 60 nm N+ $In_{0.53}Ga_{0.47}As$ (target 4×10^{19} cm^{-3}) source-drain regions.

The growth steps define the source-drain recess and gate recess. Post-growth steps define the mesa width (W_g), the T-Gate footprint, and source-drain metal spacing. Devices are mesa isolated by selective wet etching. The $In_{0.53}Ga_{0.47}As$ channel cap was then thinned by 3 cycles of digital etching in dilute HCl. A BHF dip was done immediately prior to loading into the ALD where the channel surface was passivated using 9 cycles of N_2-plasma and trimethylaluminum (TMAl) followed by 40 cycles of H_2O and tetrakis(ethylmethylamido) zirconium(IV) (TEMAZ). The sample was then annealed at 400℃ for 15 mins in forming gas to passivate dangling bonds at the semiconductor / high-k interface. A two-step T-Gate EBL exposure is used to realize sub-100 nm T-Gate footprints. CSAR 62:Anisole 1:1 was spun and exposed at high dose and developed in amyl acetate, defining the T-Gate foot. Samples were then coated in UV6, exposed at low dose, and develop in AZ300-MIF to define the T-Gate head. Ni / Au 30 / 300 nm T-Gates were then thermally evaporated and lifted off. The T-Gate foot-width (L_{TGF}) is larger than L_g to allow for re-alignment tolerance. Source-drain vias were then exposed and the high-k removed using BHF. Finally, Ti / Pd / Au 20 / 20 / 100 nm pad metal was electron beam evaporated and lifted-off.

Figure 2 shows the TEM cross-section of device 2. The T-Gate and dummy gates are aligned very well with minimal parasitic gate metal overlap outside of the intrinsic gate recess. The faceting observed at the gate edges depends on the device orientation. Devices with conduction in (0$\bar{1}$1) generally have steeper facets resulting in lower extrinsic gate source fringe capacitance ($C_{GS,f}$) and higher f_τ. This likely due to the formation of {011} facets at the dummy gate edge of (0$\bar{1}$1) conduction devices during regrowth. Channel thinning by isotropic digital etching prior to high-k deposition causes the facet edge to round, as seen in Figure 2c. The ZrO_2 and $Al_xO_yN_z$ interfacial layer thicknesses were determined by cross-sectional TEM to be $t_{ox} = 2.09$ nm $\pm$ 0.26 nm and $t_{int} = 0.80$ nm $\pm$ 0.11 nm. The

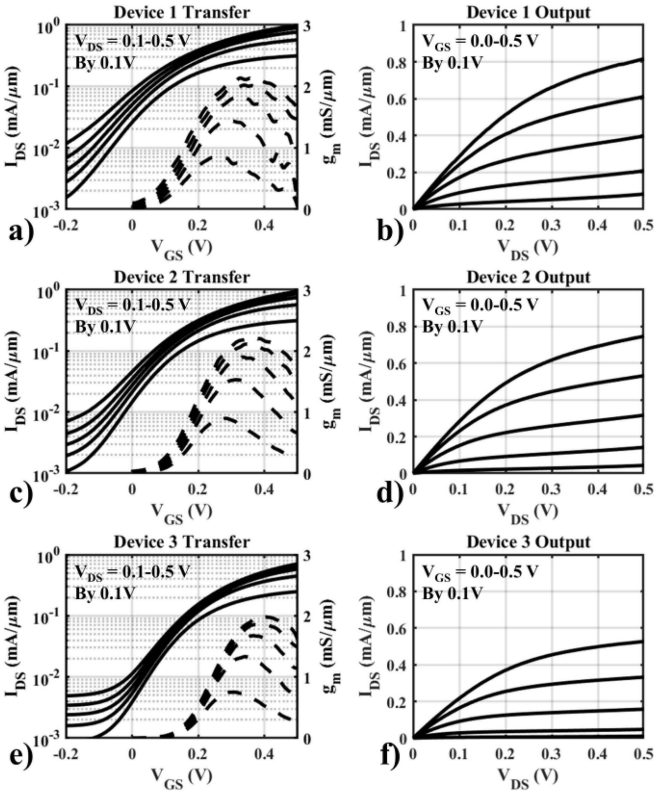

FIGURE 3. Transfer and output characteristics for representative devices.

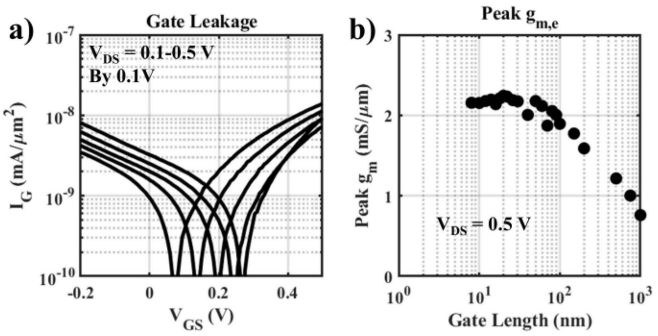

FIGURE 4. a) gate leakage and b) gate length series of peak DC transconductance.

dielectric permittivity's are estimated to be $\varepsilon_{r,int} = 10$ and $\varepsilon_{r,ox} = 22$ from [12] corresponding to an EOT = 0.69 nm.

III. DC PERFORMANCE

Figure 3 shows the transfer and output characteristics of three representative devices: Device 1 (peak f_τ), Device 2 (peak $g_{m,e}$), and Device 3 (peak f_{max}).

Device 2 exhibits a peak DC transconductance ($g_{m,e}$) of 2.23 mS/μm at $V_{DS} = 0.5$ V and $V_{GS} = 0.3$ V. All devices exhibited large $I_{off} \approx 1\mu A/\mu m$, independent of gate length, due to parallel leakage through the 20 nm UID-InP etch stop layer. The minimum subthreshold slope is convoluted with the leakage and cannot be used to infer dielectric quality. Instead, comparison of DC and RF $g_{m,e}$ must be used. At $V_{DS} = 0.5$ V and $V_{GS} = 0.3$ V, device 2 exhibits $g_{m,e}$ of 2.53 mS/μm at 1 GHz. This small deviation suggests low D_{it} and a high quality high-k gate dielectric. The off-state leakage can be reduced by eliminating or thinning the UID-InP etch stop. Excellent gate leakage $I_G < 10^{-7}$ mA/μm^2 is observed, shown in Figure 4a. Peak transconductance is observed at $V_{GS} \approx 0.4$ V for long gate length devices, shifting to 0.3 V for short gate length devices due to short-channel effects. Peak DC $g_{m,e}$ is observed to saturate for $L_g < 50$ nm and only moderately decreases for $L_g < 20$ nm suggesting excellent electrostatics, shown in Figure 4b.

Two sets of TLMs are used to estimate source resistance (R_S): The first measures the contact resistance of metal to N+ InGaAs (R_C) and the film resistance of the N+ InGaAs (R_N), while the second measures the vertical access resistance from the N+ InGaAs, through the modulation doped InP to the composite channel (R_A) and the link quantum well resistance (R_L). The film resistances, normalized to gate width, are $R_C = 8\Omega \cdot \mu m$, $R_N = 21\Omega \cdot \mu m$, $R_A = 69\Omega \cdot \mu m$, $R_L = 13\Omega \cdot \mu m$ for a total $R_S = 111\Omega \cdot \mu m$. R_N was reduced from [1] by reducing the source-drain metal spacing from 5 μm to 2 μm. It is unclear why R_A is significantly lower than in our previous work [1] given that the N+ sheet resistance and R_C are comparable suggesting similar doping in the source-drain regions.

IV. RF PERFORMANCE

S-parameters were measured from 250 MHz to 67 GHz using on-wafer probing and −25 dBm port power. Prior to measurement, off-wafer load-reflect-reflect-match calibration was done. On-wafer open and short-circuit pad parasitics were de-embedded from the transistor measurements. f_τ and f_{max} are determined by fitting the −20 dB/dec roll-off of H_{21} and U from 10 GHz to 30 GHz and 40 GHz to 50 GHz respectively. The order of pad extraction (open first vs. short first) only minimally changes the transistors extracted 2-port parameters [13], with the extracted f_τ and f_{max} changed by less than 6% and 12% respectively. The paper quotes the results of the more pessimistic extraction. Substrate dielectric mode coupling at $f > 30$ GHz creates artifacts in U. The measured unilateral power gain (U), maximum stable and available gain (MSG/MAG), and current gain (H_{21}) are shown in Figure 5 for the representative devices biased for peak f_τ, f_{max}. Contour plots of the figure of merit (FOM) of interest for each device are also shown. Devices are 2 finger, $W_g = 20\mu m$ for a total gate periphery of 40 μm.

Small signal equivalent circuit parameters for the intrinsic devices were extracted by fitting Y- and Z-parameters, as reported in [14]. Figure 6 shows the extracted FOMs, capacitances, gate-resistance, transconductance, and gate conductance. From Figure 6a, $(0\bar{1}1)$ conduction devices generally exhibit larger than f_τ than (011) conduction devices. Figure 6d shows that $g_{m,e}$ is independent of orientation while Figure 6b suggests that C_{GS} and C_{GD} are moderately smaller for $(0\bar{1}1)$ devices, explaining the moderate improvement in f_τ. Power-gain cut-off frequency (f_{max}) increases as the gate

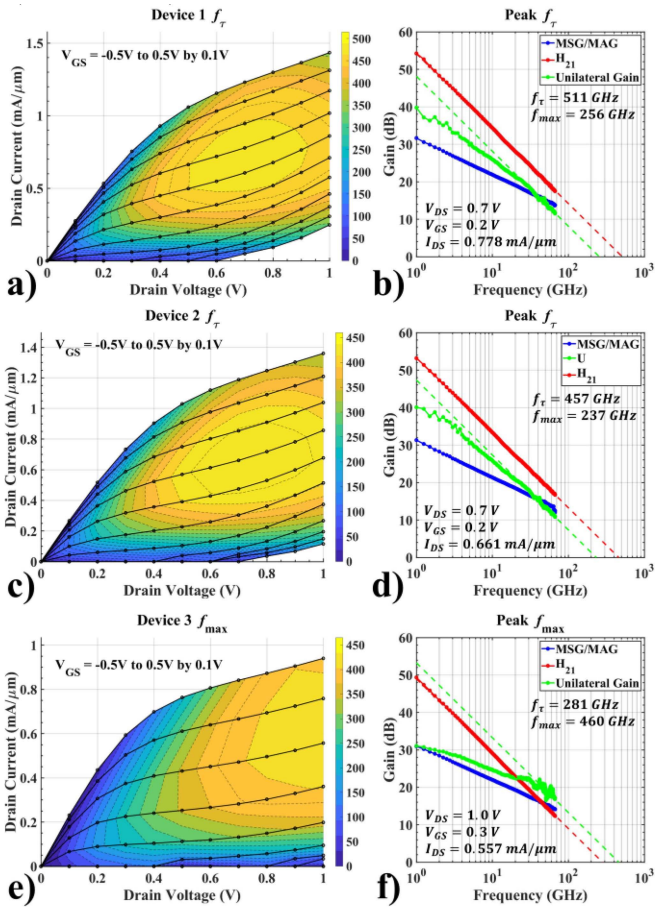

FIGURE 5. Output and FOM contour with fitted FOM of interest for representative devices.

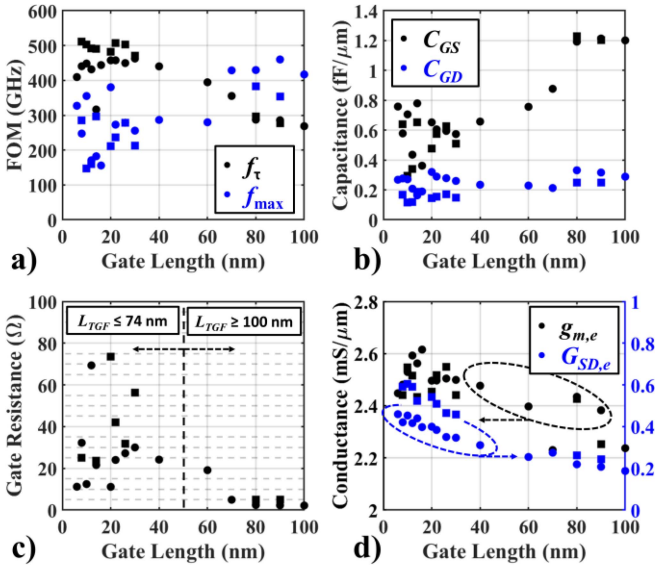

FIGURE 6. Gate length series of extracted a) high-frequency FOMs (f_τ , f_{max}) b) C_{GS} and C_{GD} c) R_G and d) $g_{m,e}$ and $G_{DS,e}$ all at peak f_τ bias where squares are $(0\bar{1}1)$ and circles are (011).

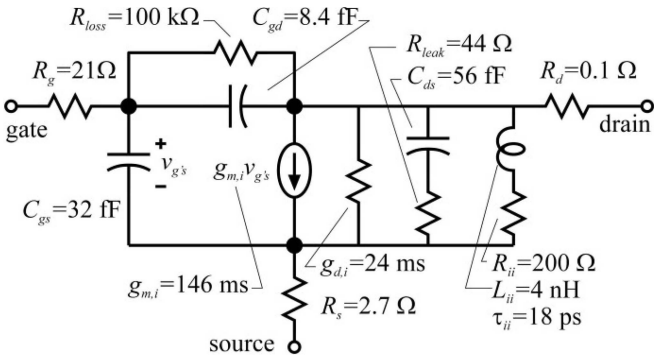

FIGURE 7. Equivalent circuit used to fit S-parameters of L_g = 8 nm $(0\bar{1}1)$ conduction device at peak f_τ bias V_{GS} = 0.20 V, V_{DS} = 0.70 V.

length increases despite decreasing f_τ. Drain conductance ($G_{DS,e}$) increases with decreasing gate length due to degradation of electrostatics. Devices oriented in $(0\bar{1}1)$ exhibit larger $G_{DS,e}$ due to the increased separation of overlapping T-Gate metal to link 2DEG. The overlapping T-Gate metal likely modulates the 2DEG to a lesser degree than the intrinsic channel. $G_{DS,e}$ degrades from 0.2 mS/μm at $L_g = 80$ nm to 0.6 mS/μm at $L_g < 20$ nm. Extraction of $G_{DS,e}$ is difficult due to the parallel source-drain leakage which manifests itself as a parallel resistor. Figure 6c shows that short gate length devices exhibit $R_G > 10$ Ω, often > 20 Ω, severely limiting $f_{\max}$.

The small signal equivalent circuit fit to the S-parameters measured from device 1 is illustrated in Figure 7. The leaky etch stop layer requires a large series C_{DS} and R_{Leak} to fit imag(Y_{22}). Because R_{Leak} is comparable to $1/G_{DS}$, it is difficult to accurately determine the specific value of either. A series L-R network is needed to fit the low frequency Y_{22} source-drain conductance. This is necessary at large V_{DS} where impact ionization or band-to-band tunneling at the gate-drain edge results in low-frequency dispersion [1]. The fitted L/R time constant is 18 ps which is half that reported in [1].

Compared to [1], f_τ is improved due to the larger $g_{m,i}$ while $f_{\max}$ of short channel devices is limited by large R_G. Better T-Gate metal filling on longer gate length devices enables balanced f_τ, $f_{\max}$. Transconductance is nearly constant for $L_g \leq 80$ nm because of the thick channel and associated long electron mean free path. $C_{GD} \approx 0.2$ fF/μm is consistent with [1] and [10] while C_{GS} is higher than [10] due to the larger insulator capacitance and thinner channel. C_{GS} is lower than [1] because of the thicker channel. Because the T-Gate footprint (L_{TGF}) is larger than the channel recess, and because L_{TGF} is not scaled with each L_g, it is non-trivial to extract the intrinsic and fringe capacitances.

V. CONCLUSION

Record f_τ = 511 GHz for composite-channel metal-oxide-semiconductor high-electron-mobility-transistors (MOS-HEMT) has been demonstrated. Further scaling of our previously reported devices resulted in improved R_S and $g_{m,e}$ yielding increased f_τ though poor T-Gate

metal filling resulted in large R_G that limited power gain in short L_g devices. The improved performance suggests that MOS-HEMTs are promising candidates for mm-wave applications. Further improvement in the transistor DC and RF FOMs can be realized by further scaling the high-k dielectric thickness, reducing the source-drain metal spacing, optimizing the channel thickness, and improving the T-Gate metallization process. Reducing the gate resistance of the short-L_g devices should substantially increase f_{max} from 285 GHz.

REFERENCES

[1] J. Wu, Y. Fang, B. Markman, H.-Y. Tseng, and M. J. W. Rodwell, "$L_g = 30$nm InAs channel MOSFETs exhibiting $f_{max} = 410$ GHz and $f_\tau = 357$ GHz," *IEEE Electron Device Lett.*, vol. 39, no. 4, pp. 472–475, Apr. 2018, doi: 10.1109/LED.2018.2803786.

[2] X. Mei *et al.*, "First demonstration of amplification at 1 THz using 25-nm InP high electron mobility transistor process," *IEEE Electron Device Lett.*, vol. 36, no. 4, pp. 327–329, Apr. 2015, doi: 10.1109/LED.2015.2407193.

[3] J. C. Rhode *et al.*, "Indium phosphide heterobipolar transistor technology beyond 1-THz bandwidth," *IEEE Trans. Electron Devices*, vol. 62, no. 9, pp. 2779–2785, Sep. 2015, doi: 10.1109/TED.2015.2455231.

[4] M. J. W. Rodwell *et al.*, "100-340GHz systems: Transistors and applications," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, Dec. 2018, pp. 1–4, doi: 10.1109/IEDM.2018.8614537.

[5] H.-B. Jo *et al.*, "L_g = 25nm InGaAs/InAlAs high-electron mobility transistors with both f_T and f_{max} in excess of 700 GHz," *App. Phys. Exp.*, vol. 12, May 2019, Art. no. 54006, doi: 10.7567/1882-0786/ab1943.

[6] A. Tessman *et al.*, "20-nm In0.8$Ga_{0.2}$As MOSHEMT MMIC technology on Si," *IEEE J. Solid-State Circuits*, vol. 54, no. 9, pp. 2411–2418, Sep. 2019, doi: 10.1109/JSSC.2019.2915161.

[7] E. Chang, C.-I. Kuo, H.-T. Hsu, C.-Y. Chiang, and Y. Miyamoto, "InAs thin-channel high-electron-mobility transistors with very high current-gain cutoff frequency for emerging submillimeter-wave applications," *Appl. Phys. Exp.*, vol. 6, no. 3, Feb. 2013, Art. no. 034001, doi: 10.7567/APEX.6.034001.

[8] J. Lin, D. A. Antoniadis, and J. A. del Alamo, "Impact of intrinsic channel scaling on InGaAs quantum-well MOSFETs," *IEEE Trans. Electron Devices*, vol. 62, no. 11, pp. 3470–3476, Nov. 2015, doi: 10.1109/TED.2015.2444835.

[9] S. Lee *et al.*, "Highly scalable raised source/drain InAs quantum well MOSFETs exhibiting I_{on} = 482 at I_{off} = 100 nA/and V_{DD} = 0.5 V," *IEEE Electron Device Lett.*, vol. 35, no. 6, pp. 621–623, Jun. 2014, doi: 10.1109/LED.2014.2317146.

[10] J. Lin, X. Cai, Y. Wu, D. A. Antoniadis, and J. A. del Alamo, "Record maximum transconductance of 3.45 mS/μm*for* III-V FETs," *IEEE Electron Device Lett.*, vol. 37, no. 4, pp. 381–384, Apr. 2016, doi: 10.1109/LED.2016.2529653.

[11] M. Lundstrom and J. Guo, "The ballistic nanotransistors," in *Nanoscale Transistors: Device Physics, Modeling, and Simulation*. New York, NY, USA: Springer, 2006, ch. 3, sec. 4, pp. 91–103.

[12] P. J. Harrop and J. N. Wanklyn, "The dielectric constant of zirconia," *Brit. J. Appl. Phys.*, vol. 18, no. 6, pp. 739–740, 1967, doi: 10.1103/PhysRev.36.1232.

[13] M. C. A. M. Koolen, J. A. M. Geelen, and M. P. J. G. Versleijen, "An improved de-embedding technique for on-wafer high-frequency characterization," in *Proc. Bipolar Circuits Technol. Meeting*, Minneapolis, MN, USA, 1991, pp. 188–191, doi: 10.1109/BIPOL.1991.160985.

[14] S. Lee, H. K. Yu, C. S. Kim, J. G. Koo, and K. S. Nam, "A novel approach to extracting small-signal model parameters of silicon MOSFET's," *IEEE Microw. Guid. Wave Lett.*, vol. 7, no. 3, pp. 75–77, Mar. 1997, doi: 10.1109/75.556037.