The Impact of Width Downscaling on the High-Frequency Characteristics of InGaAs Nanowire FETs

P. HUANG1, Q. H. LUC2, A. SIBAJA-MORENO3, H. L. KO1, J. Y. WU1, N. A. TRAN2, N. COLLAERT3, AND E. Y. CHANG4 (Life Fellow, IEEE)

1 International College of Semiconductor Technology, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan
2 Department of Materials Science and Engineering, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan
3 IMEC, 3001 Leuven, Belgium
4 Department of Materials Science and Engineering, Department of Electronics Engineering, and the International College of Semiconductor Technology, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan

CORRESPONDING AUTHOR: E. Y. CHANG (e-mail: edc@mail.nctu.edu.tw)

This work was supported in part by the “Center for the Semiconductor Technology Research” from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE), Taiwan; in part by the Ministry of Science and Technology, Taiwan, under Grant NSTC-110-2622-8-A49-008-SB, Grant NSTC-111-2622-8-A49-018-SB, and Grant NSTC-111-2634-F-A49-008; and in part by the National Chung-Shan Institute of Science and Technology under Grant NCSIST-404-V407(111).

ABSTRACT This work demonstrates the high-frequency characteristics of In0.53Ga0.47As nanowire with scaled wire width by implementing TCAD simulations. The physical models and correlated parameters have been calibrated to the experiments (Ko et al., 2022). As the width of the nanowire is scaled to 10 nm, the electron density peaks are no longer located close to the oxide/semiconductor interface. Instead, the peaks merge and volume inversion effects appear due to the strong quantum confinement. The volume inversion effects lead to higher cut-off frequency due to the reduced total transport delay time. To have a better understanding of this phenomena, the high-frequency properties of nanowire were quantified with the assistance of small-signal analysis and delay time analysis using TCAD. It is found that the channel charging delay increases with narrower wire width due to the raise of source/drain resistance. Regarding the extrinsic and intrinsic delay, they increase with smaller wire width and drop at width of 10 nm due to the volume inversion effects. Electron distribution which aims to clarify the above-mentioned observation is also plotted.

INDEX TERMS III-V semiconductor materials, MOSFET, high-k gate dielectrics, cutoff frequency, TCAD.

I. INTRODUCTION

The development of high-frequency transistors based on III-V materials, which possess outstanding properties such as high electron mobility and high injection velocity [1], [2], [3], has made significant advances in the past 20 years. From sub-THz GaAs-based [4] to THz domain InP-based [5] high electron mobility transistors (HEMTs), the III-V devices have been extensively used in the high-frequency applications. However, the complexity of the epi structure and fabrication process makes the III-V HEMTs difficult to be integrated with commercial Si CMOS technology for system on a chip (SOC) applications. Therefore, a surface channel device with CMOS compatible process emerges to be necessary. For the surface channel transistors, the high-frequency properties have been greatly improved by scaling down the effective channel length ($L_{eff}$) due to the reduced spacing from the source edge to the drain [6] and [7]. Unfortunately, the benefits are deteriorated by the short channel effects (SCEs) such as large subthreshold swing (SS), drain induced barrier lowering (DIBL) and threshold voltage roll-off. To keep scaling the device channel length while maintaining gate control, multi-gate field-effect transistors (MuFETs) such as FinFETs, nanowire and nanosheet with high-$k$ dielectric are implemented [8], [9], [10], [11] due to their higher area-to-volume ratio and thus good immunity to the SCEs. The gate control can be further enhanced by scaling down the width of the transistors [12] and [13]. An additional benefit of
scaling down the width is the reduction of power consumption. However, the high-frequency properties can be dramatically sacrificed in devices with smaller width [14]. So far, there are several research works demonstrating the high-frequency properties of InGaAs MuFETs [15], [16], [17], [18] from the results of experiments and the circuit simulations. Nevertheless, there are less research work discussing about the impact of device width on the high-frequency properties of the InGaAs transistors especially for widths down to merely 10 nm. In this work, we demonstrate the impact of scaled width on the high-frequency properties of In$_{0.53}$Ga$_{0.47}$As nanowire with the assistance of TCAD simulations. The DC properties have been calibrated to the experiments carried out on devices [19] with a good agreement while the high-frequency properties are simulated with the help of small-signal analysis of the device admittance (Y). Our attention is focus on the device cut-off frequency ($f_T$) since it is directly related to the intrinsic transistor properties. The strong quantum effects in very small device dimension are also taken into consideration. It is found that the nanowire width scaling is necessary to suppress the SCEs and that the overall transport delay can be improved when the device width is scaled to 10 nm. Therefore, the high-frequency properties can be less sacrificed.

II. DEVICE STRUCTURE AND SIMULATIONS SET-UP

The device structure for fabrication and simulation is shown in Fig. 1 and the details of the fabrication process can be referenced elsewhere [11], [20]. The 3-nm HfO$_2$ gate dielectric was deposited using atomic layer deposition (ALD) technique and followed by post remote-plasma (NH$_3$/N$_2$) treatment [21], [22], [23]. A thin ALD-TiN and physical vapor deposition (PVD) TiN were deposited as gate metal. To investigate the high-frequency properties, the DC simulations were carefully fitted to the experiments in the literature [19]. The fitting results based on the DC measurements of the transfer characteristics of the device biased at drain to source voltage ($V_{DS}$) at linear (0.05 V) and saturation (0.5 V) regions and the transmission electron microscope (TEM) image of the nanowire are shown in Fig. 2(a) and (b). The device channel length ($L_{ch}$) is 50 nm while the top width ($W_N$) and thickness ($T_N$) are 60 nm and 40 nm, respectively. The simulated results show excellent agreement with measurements reported in reference [19]. As for the physical models, Fermi-Dirac statistics is applied. Density gradient is used for the quantum corrections. Impurities scattering, phonon scattering and surface roughness scattering are considered by applying the Inversion and Accumulation Layer (IAL) mobility model. The Coulomb scattering induced by interface traps is also considered. The conduction band non-parabolicity and multi-valley models are also taken into consideration. The subthreshold and off-state properties are captured by Shockley-Read-Hall, Auger recombination and Hurkx band-to-band tunneling models. The doping profile and the mobility models have been carefully calibrated to the planar configuration as shown in [24]. Though the device configuration is different, with a slight modification of the mobility models we can achieve at good fitting results. Critical parameters used for calibration are shown in Table 1. It is worth noting that default parameters are used if not reported in the Table. The number of mid-gap traps is relatively high due to non-optimized interfacial treatment. The interface quality can be further improved by changing the species of the plasma or the insertion of interfacial layers.

Fig. 3 shows the accuracy of the applied physical models with scaled width for (a) minimum SS versus $W_N$
TABLE 1. Parameters for TCAD Calibration.

| Symbol | Quantity                                      | Value               |
|--------|-----------------------------------------------|---------------------|
| μn ep | impurities scattering mobility                | 7581 cm²/V·s at low doping¹ |
|       |                                               | 4041 cm²/V·s at medium doping² |
|       |                                               | 1815 cm²/V·s at high doping³ |
| β     | surface roughness scattering mobility         | 1 × 10⁻⁴ cm²/V·s    |
| Vsat  | saturation velocity                           | 3.2 × 10⁹ cm/s      |
| ρ <sup>0</sup> | contact resistivity                         | 5 × 10⁻⁵ Ω·cm²     |
| γ     | quantum correction                            | 1.776               |
| D0    | mid-gap traps                                 | 5.88 × 10⁻¹⁵ cm²·eV⁻¹ |
| ϕ <sup>∞</sup> | metal work function                         | 4.88 eV             |

¹1 × 10⁻¹⁷ cm², ²1 × 10⁻¹⁸ cm², ³1 × 10⁻¹⁹ cm²

and (b) peak Gₘ per wire versus Wₙ. The measurement data of three device is taken for each width. From the plots, the good agreement between the experiments carried out on devices and simulation results is observed. Based on the successful prediction of scaling tendency, we can extend to the high-frequency domain with the assistance of small-signal AC simulations.

III. RESULTS AND DISCUSSIONS

Fig. 4 presents the device cut-off frequency (fₜ) as a function of the VGS, calculated from the small-signal admittance (Y) at L_ch = 50 nm and V_DS = 0.5 V with various wire widths. fₜ is the frequency where the device current gain degrades to 0 dB. To avoid the high-frequency distortion that might affect our extracted fₜ, a -20 dB/dec extraction methodology is implemented. The peak fₜ of the device is shown to be 229 GHz with Wₙ = 60 nm. From Fig. 4 we see that fₜ degrades from Wₙ = 60 nm to Wₙ = 20 nm, while from Wₙ = 20 nm to Wₙ = 10 nm, the fₜ raises to an even higher value than Wₙ = 30 nm. Inset shows the peak fₜ versus various Wₙ. More data points from the simulations have been provided to show a clear tendency of fₜ as a function of wire width.

To further understand the mechanisms, we analyzed the delay time of the nanowire based on [25] and [26]. Fig. 5(a) shows the total delay time (τ_d) versus the reciprocal of the drain current density (W_g/|IDS|) with different wire widths. W_g is the perimeter of the gate, defined as 2(Wₙ + T_N).

The formula of τ_d is as shown in [26]:

\[ \tau_d = \frac{1}{2 \times \pi \times f_t} = \tau_{int} + \tau_{ext} + \tau_{ch}. \]  

(1)

where τ_int, τ_ext and τ_ch are the intrinsic, extrinsic and channel charging delay, respectively. The sum of intrinsic and extrinsic delay (τ_int + τ_ext) can be extracted by extrapolating the linear curve, which passes through the minimum total delay (τ_d_min), to the intercept of infinite drain current density as shown in Fig. 5(b). τ_ch is the difference between τ_d_min and τ_int + τ_ext. The total delay time is high in lower drain current density because of the low conductance in the channel. In high drain current density, the total delay time increases due to the degradation of transconductance and the extraction method is no longer applicable.

The value of extrinsic delay is extracted from the linear extrapolation of τ_int + τ_ext according to the channel length as shown in Fig. 6. When the channel length is reduced to 0 nm, the intrinsic delay no longer exists and there remains only the extrinsic delay.

The extracted delay time components are shown in Fig. 7, in which we notice that the intrinsic delay and extrinsic delay increase with reduced wire width but drop at 10 nm. On the other hand, the channel charging delay increases monotonously with narrower wire width. Paying attention to delay components of widths 10 nm and 20 nm in Fig. 7,
where \( g_{m,i} \) is the intrinsic transconductance, \( C_{gs,e} \) and \( C_{gd,e} \) are the extrinsic gate-to-source and gate-to-drain capacitances. \( C_{gs} \) and \( C_{gd} \) are the overall gate-to-source and gate-to-drain capacitances, respectively. \( g_{sd} \) is the output conductance and \( C_{gg,i} \) is the intrinsic gate capacitance. \( C_{gs,i} \), \( C_{gd,i} \), \( g_{m,i} \) and \( R_{ch} \) can be obtained by de-embedding the parasitic components of the transistors (including extrinsic capacitance, \( R_s \) and \( R_d \)). These components are extracted when the device is in the cold-FET state (\( V_{DS} = 0 \) V). The extraction methodology can be referred to [27] while the de-embedding process can be referred to [28].

For the intrinsic delay, the values of \( C_{gg,i} \) are shown in Fig. 9(a). It is very clear that the intrinsic gate capacitance at 100 nm gate perimeter does not follow the trendline (pure scaling) proportional to the scaling of capacitance due to total gate perimeter. The gap between the pure scaling line and the \( C_{gg,i} \) is broadened with narrower \( W_g \). There is a 36% decrease of \( C_{gg,i} \) compared to the pure scaling line at \( W_g = 100 \) nm (\( W_N = 10 \) nm) due to the volume inversion effects [29]. Fig. 9(b) shows the comparison of extrinsic transconductance (\( g_{m,e} \)) and the intrinsic transconductance where their relationship is shown in eq. (11) from reference [30].

\[
g_{m,i} = g_{m,e}\left[1 - g_{m,e} \times R_s - g_{sd} \times (R_s + R_d)\right]
\]

The intrinsic transconductance show nearly twice of the extrinsic transconductance due to the exclusion of the source/drain resistance. The source/drain resistance are extracted from the voltage drop in the source/drain region. It is very interesting to notice that the intrinsic transconductance degrades at a slower rate than gate perimeter scaling at smaller \( W_g \), which is opposite to the behavior of extrinsic transconductance. The phenomenon is due to the reduced carrier scattering in the inversion channel and the improvement of the effective mobility [29], [31] and [32] from volume inversion effects. However, these benefits are hidden when the source/drain resistance are considered.

This phenomenon can be investigated through the electron distribution in Fig. 10. The 2D cross-section view shown in Fig. 10 (a) is a cut plane in the middle of the channel and the direction is perpendicular to the current flow. The 1D electron density distribution cut in the A-A’ direction is shown in Fig. 10 (b). The electron density has been normalized to the peak value of \( W_N = 10 \) nm. Due to
the enhancement of quantum effects with scaled width, the electrons are pushed away from the oxide/semiconductor interface and finally merge at the center, this is referred as the volume inversion effect. This effect results in the increase of capacitance equivalent thickness (CET) and thus a drop of C_{gg,i} [33]. Fig. 10 (c) shows the electron mobility (μ_e) distribution cut in the A-A’ direction. Obviously, the electron mobility at W_N = 10 nm shows the highest value. Since the channel carriers are pushed away from the interface, the surface roughness scattering effects and the interface traps scattering effects are mitigated. These can be used to describe the tendency of intrinsic transconductance that we observed in Fig. 9(b).

As for the extrinsic delay, Fig. 7 shows that the extrinsic delay component follows the same trend as the intrinsic delay. This behavior is related to same mechanism of volume inversion effects previously described for the intrinsic delay. Since there is an overlap region in the nanowire (Fig. 8b), it also suffers from the quantum confinement effects from the gate dielectric. The volume inversion also occurs at these overlap regions.

For the channel charging delay, equation (4), the value of R_s + R_d is shown in Fig. 11(a) while the term including capacitive components [C_{gd} + (C_{gs} + C_{gd}) × (g_{sd}/g_{m,i})] is shown in Fig. 11(b). The device suffers from huge S/D series resistance when the dimension is scaled. The R_s + R_d and [C_{gd} + (C_{gs} + C_{gd}) × (g_{sd}/g_{m,i})]^{-1} for each width are normalized corresponding to their value at 60 nm for clearer comparison as shown in Fig. 11(c). Although the overall capacitance reduces with smaller width, the huge R_s and R_d still lead to the monotonous increase of channel charging delay.

IV. CONCLUSION

The high-frequency InGaAs nanowire FETs with scaled wire width have been systematically investigated with the assistance of TCAD simulation tools. It is found that the width scaling is effective in suppressing the device short channel effect without greatly sacrifice the cut-off frequency. Due to the strong quantum and volume inversion effects in smaller dimensions, the transport delay of the transistors can be improved with respect to the expected RF performance when only traditional width scaling is considered. These could be used as a potential knob for future RF optimizations of high-frequency III-V electronic devices.

REFERENCES

[1] J. A. del Alamo, “Nanometre-scale electronics with III-V compound semiconductors.” Nature, vol. 479, no. 7373, pp. 317–323, Nov. 2011, doi: 10.1038/nature10677.
[2] S. Takagi et al., “Carrier-transport-enhanced channel CMOS for improved power consumption and performance,” IEEE Trans. Electron Devices, vol. 55, no. 1, pp. 21–39, Jan. 2008, doi: 10.1109/TED.2007.911034.
[3] D.-H. Kim, J. del Alamo, D. A. Antoniadis, and B. Brar, “Extraction of virtual source injection velocity in sub-100 nm III–V HFETs,” in Proc. IEEE Int. Electron Devices Meeting (IEDM), 2009, pp. 1–4, doi: 10.1109/IEDM.2009.5424268.
[4] T. Takahashi et al., “Enhancement of fmax to 910 GHz by adopting asymmetric gate recess and double-side-doped structure in 75-nm-gate InAlAs/InGaAs HEMTs,” IEEE Trans. Electron Devices, vol. 64, no. 1, pp. 89–95, Jan. 2017, doi: 10.1109/TED.2016.2624899.
[5] 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.
[6] S. Lee et al., “Record RF performance of 45-nm SOI CMOS technology,” in IEDM Tech. Dig., Dec. 2007, pp. 255–258, doi: 10.1109/IEDM.2007.4418916.
[7] H. Li et al., “Technology scaling and device design for 350 GHz RF performance in a 45nm bulk CMOS process,” in Proc. IEEE Symp. VLSI Technol., 2007, pp. 56–57, doi: 10.1109/VLSIT.2007.4339725.
[8] J. J. Gu, Y. Q. Liu, Y. Q. Wu, R. Colby, R. G. Gordon, and P. D. Ye, “First experimental demonstration of gate-all-around III–V MOSFETs by top-down approach,” in IEDM Tech. Dig., Washington, DC, USA, Dec. 2011, pp. 33.2.1–33.2.4, doi: 10.1109/IEDM.2011.6131662.
[9] N. Waldron et al., “InGaAs gate-all-around nanowire devices on 300 mm Si substrates,” IEEE Electron Device Lett., vol. 35, no. 11, pp. 1097–1099, Nov. 2014, doi: 10.1109/LED.2014.2359579.
[10] K. Mistry et al., “A 45nm logic technology with high-k metal gate transistors, strained silicon, 9 Cu interconnect layers, 193nm dry patterning, and 100% Pb-free packaging,” in Proc. IEEE Int. Electron Devices Meeting, 2007, pp. 247–250, doi: 10.1109/IEDM.2007.4418914.
[11] Q. H. Luc et al., “In0.53Ga0.47 As FinFET and GAA-FET with remote-plasma treatment,” in IEEE Electron Device Lett., vol. 39, no. 3, pp. 339–342, Mar. 2018, doi: 10.1109/LED.2018.2798589.
[12] J. Zhang, M. Si, X. B. Lou, W. Wu, R. G. Gordon, and P. D. Ye, “InGaAs 3D MOSFETs with drastically different shapes formed by anisotropic wet etching,” in IEDM Tech. Dig., Dec. 2015, pp. 15.2.1–15.2.4, doi: 10.1109/IEDM.2015.7409702.
M. Hellenbrand, E. Lind, O.-P. Kilpi, and L.-E. Wernersson, “Effects of traps in the gate stack on the small-signal RF response of III-V nanowire MOSFETs,” Solid-State Electron., vol. 171, Sep. 2020, Art. no. 107840, doi: 10.1016/j.sse.2020.107840.

C. B. Zota, G. Roll, L.-E. Wernersson, and E. Lind, “Radio-frequency characterization of selectively regrown InGaAs lateral nanowire MOSFETs,” IEEE Trans. Electron Devices, vol. 61, no. 12, pp. 4078–4083, Dec. 2014, doi: 10.1109/TED.2014.2363732.

C. Zota, L.-E. Wernersson, and E. Lind, “In0.53Ga0.47As multiple-gate field-effect transistors with selectively regrown channels,” IEEE Electron Device Lett., vol. 35, no. 3, pp. 342–344, Mar. 2014, doi: 10.1109/LED.2014.2301843.

S. Johansson, E. Memisevic, L.-E. Wernersson, and E. Lind, “High-frequency gate-all-around vertical InAs nanowire MOSFETs on Si substrates,” IEEE Electron Device Lett., vol. 35, no. 5, pp. 518–520, May 2014, doi: 10.1109/LED.2014.2310119.

H.-L. Ko et al., “Electrical characteristics of In0.53Ga0.47As gate-all-around MOSFETs with different nanowire shapes,” IEEE Trans. Electron Devices, vol. 69, no. 8, pp. 4183–4187, Aug. 2022, doi: 10.1109/TED.2022.3181572.

H.-L. Ko et al., “Nitrogen-passivated (010) In0.53Ga0.47As FinFETs with high peak gm and reduced leakage current,” IEEE Trans. Electron Devices, vol. 69, no. 2, pp. 495–499, Feb. 2022, doi: 10.1109/TED.2021.3133222.

Q. H. Luc et al., “Electrical characteristics of n-p-In0.53Ga0.47As MOSCAPs with in situ PEALD-AlN interfacial passivation layer,” IEEE Trans. Electron Devices, vol. 61, no. 8, pp. 2774–2778, Aug. 2014, doi: 10.1109/TED.2014.2329479.

Q. H. Luc, H. B. Do, M. T. H. Ha, C. C. Hu, Y. C. Lin, and E. Y. Chang, “Plasma enhanced atomic layer deposition passivated HfO2/AlN/In0.53Ga0.47As MOSCAPs with sub-nanometer equivalent oxide thickness and low interface trap density,” IEEE Electron Device Lett., vol. 36, no. 12, pp. 1277–1280, Dec. 2015, doi: 10.1109/LED.2015.2486771.

Q. H. Luc et al., “Effects of In-Situ plasma-enhanced atomic layer deposition treatment on the performance of HfO2/In0.53Ga0.47As metal-oxide-semiconductor field-effect transistors,” IEEE Electron Device Lett., vol. 37, no. 8, pp. 974–977, Aug. 2016, doi: 10.1109/LED.2016.2581175.

P. Huang et al., “Investigation of device transport characteristics enhancement of In0.53Ga0.47As MOSFET through in situ NH3/N2 remote-plasma treatment,” AIP Adv., vol. 11, no. 1, 2021, Art. no. 15050, doi: 10.1063/5.0037378.

N. Moll, M. R. Hueschen, and A. Fischer-Colbrie, “Pulse-doped AlGaAs/InGaAs pseudomorphic MODFETs,” IEEE Trans. Electron Devices, vol. 35, no. 7, pp. 879–886, Jul. 1988, doi: 10.1109/16.3339.

W. Xing et al., “InAlN/GaN HEMTs on Si with high fT of 250 GHz,” IEEE Electron Device Lett., vol. 39, no. 1, pp. 75–78, Jan. 2018, doi: 10.1109/LED.2017.2773054.

S. Cho, K. R. Kim, B.-G. Park, and I. M. Kang, “RF performance and small-signal parameter extraction of junctionless silicon nanowire MOSFETs,” IEEE Trans. Electron Devices, vol. 58, no. 5, pp. 1388–1396, May 2011, doi: 10.1109/TED.2011.2109724.

I. M. Kang and H. Shin, “Non-quasi-static small-signal modeling and analytical parameter extraction of SOI FinFETs,” IEEE Trans. Nanotechnol., vol. 5, no. 3, pp. 205–210, May 2006, doi: 10.1109/TNANO.2006.869946.

S. D. Suk et al., “Investigation of nanowire size dependency on TSNWFE,” in Proc. IEEE Int. Electron Devices Meeting, 2007, pp. 891–894, doi: 10.1109/IEDM.2007.4419093.

J. Chung, M.-C. Jeng, G. May, P. Ko, and C. Hu, “Intrinsic transconductance extraction for deep-submicrometer MOSFETs,” IEEE Trans. Electron Devices, vol. 56, no. 1, pp. 140–142, Jan. 1989, doi: 10.1109/TED.1989.58516.

L. Ge and J. G. Fossum, “Analytical modeling of quantization and volume inversion in thin Si-film double-gate MOSFETs,” IEEE Trans. Electron Devices, vol. 49, no. 2, pp. 287–294, Feb. 2002, doi: 10.1109/16.981219.

E. G. Marin, F. G. Ruiz, I. M. Tienda-Luna, A. Godoy, P. Sánchez-Moreno, and F. Gámiz, “Analytic potential and charge model for III–V surrounding gate MOSFETs,” J. Appl. Phys., vol. 112, no. 8, Oct. 2012, Art. no. 084512, doi: 10.1063/1.4759275.

G. Curatola and S. Nuttinck, “The role of volume inversion on the intrinsic RF performance of double-gate FinFETs,” IEEE Trans. Electron Devices, vol. 54, no. 1, pp. 141–150, Jan. 2007, doi: 10.1109/TED.2006.887224.