Impact of Effective Mass on the Scaling Behavior of the f_T and $f_{\max}$ of III–V High-Electron-Mobility Transistors

Sabbir Ahmed, Kyle David Holland, Navid Paydavosi, Christopher Martin Sinclair Rogers, Ahsan Ul Alam, Neophytos Neophytou, Diego Kienle, and Mani Vaidyanathan, *Member, IEEE*

***Abstract*—Among the contenders for applications at terahertz frequencies are III–V high-electron-mobility transistors (HEMTs). In this paper, we report on a tendency for III–V devices with low effective-mass channel materials to exhibit a saturation in their unity-current-gain and unity-power-gain cutoff frequencies (f_T and $f_{\max}$) with a downscaling of gate length. We focus on InGaAs and GaN HEMTs and examine gate lengths from 50 nm down to 10 nm. A self-consistent, quantum-mechanical solver based on the method of nonequilibrium Green's functions is used to quasistatically extract the f_T for intrinsic III–V devices. This model is then combined with the series resistances of the heterostructure stack and the parasitic resistances and capacitances of the metal contacts to develop a complete extrinsic model, and to extract the extrinsic f_T and $f_{\max}$. It is shown that the f_T and $f_{\max}$ of III–V devices will saturate, i.e., attain a maximum value that ceases to increase as the gate length is scaled down, and that the saturation is caused by the low effective mass of III–V materials. It is also shown that the InGaAs HEMTs have faster f_T at long gate lengths, but as a consequence of their lower effective mass, they experience a more rapid f_T saturation than the GaN HEMTs, such that the two devices have a comparable f_T at very short gate lengths ($\sim$10 nm). On the other hand, due to favorable parasitics, it is shown that the InGaAs HEMTs have a higher $f_{\max}$ at all the gate lengths considered in this paper.**

***Index Terms*—Barrier collapse, drain-induced barrier lowering (DIBL), equivalent circuit, GaN, high-electron-mobility transistor (HEMT), InGaAs, nonequilibrium Green's functions (NEGF), parasitic capacitance, parasitic resistance, subband.**

I. Introduction

HIGH-ELECTRON-MOBILITY transistors (HEMTs) fabricated with III–V materials are among the leading candidates for terahertz radio-frequency (RF) applications [1]. Of principle interest in assessing the ultimate potential of these devices is the behavior of their unity-current-gain and unity-power-gain cutoff frequencies (f_T and $f_{\max}$) as the device gate length is scaled down. To date, the scaling behavior has not been extensively studied or explained, with conflicting data and comments on the expected behavior of f_T and $f_{\max}$ with scaling, and on the relative performance of the different types of HEMTs.

Regarding the gate-length scaling, in a recent review of RF transistors for terahertz applications, and based on experimental data reported in the literature, Schwierz and Liou [1, Fig. 10] observed a tendency for the f_T of III–V HEMTs to saturate at very short gate lengths, i.e., to show no further increase with decreasing gate length once the gate length is sufficiently small; they also commented on the importance of considering the impact of a low density of states near the bottom of the conduction band in assessing the potential of III–V transistors, as originally discussed in [2] and [3] and as recently discussed in [4] and [5], but they did not connect the low density of states to the gate-length scaling behavior. On the other hand, in contrast to the observation of Schwierz and Liou, the simulations of Ayubi-Moak *et al*. [6] and Akis *et al*. [7] displayed no saturation in f_T versus gate length; their work was based on a semiclassical Monte Carlo approach that included electron scattering and that scaled InGaAs HEMTs down to 10 nm.

Regarding the relative performance of different III–V HEMTs, according to the 2009 International Technology Roadmap for Semiconductors (ITRS) [8], InGaAs and GaN devices are among the most important for RF and analog/mixed-signal technology. The InGaAs HEMTs have very high f_T and $f_{\max}$ [1], [9]–[11], an outcome of the very high electron mobility in the channel [1], [9], [10]. The GaN HEMTs have emerged as interesting candidates, because they offer not only high f_T and $f_{\max}$ [12]–[14] but also the ability to operate at high voltages and high powers, owing to the large bandgap and breakdown field in nitride-based materials [15], [16]. However, it is not clear how these HEMTs will perform in comparison with each other as the gate length is scaled down. Some researchers are confident that the InGaAs HEMTs are faster [1], [9], [10], [14] whereas the 2009 ITRS requires the GaN HEMTs to be as fast [8].

Manuscript received March 22, 2012; revised July 6, 2012; accepted August 4, 2012. Date of publication September 6, 2012; date of current version November 16, 2012. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, in part by the Queen Elizabeth II Graduate Scholarship, in part by Alberta Innovates, and in part by Alberta Advanced Education and Technology. The review of this paper was arranged by Associate Editor E. Tutuc.

S. Ahmed, K. D. Holland, C. M. S. Rogers, A. U. Alam, and M. Vaidyanathan are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: maniv@ece.ualberta.ca).

N. Paydavosi was with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada. He is now with the BSIM Group, University of California, Berkeley, CA 94720 USA.

N. Neophytou is with the Institute for Microelectronics, Technical University of Vienna, Vienna 1040, Austria.

D. Kienle is with the Theoretische Physik I, Universität Bayreuth, Bayreuth 95440, Germany.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNANO.2012.2217348

In this paper, we examine the gate-length scaling behavior of the f_T and $f_{\max}$ of III–V HEMTs, with the primary aim of clarifying the connection between the effective mass (and hence density of states) of the channel material and the scaling behavior. We focus on InGaAs and GaN HEMTs, and we use a fully quantum-mechanical approach based on the method of nonequilibrium Green's functions (NEGF). The primary outcome of our study is a demonstration that devices constructed using low effective-mass channel materials can indeed be expected to experience a saturation in their f_T and $f_{\max}$ with gate-length scaling, as observed by Schwierz and Liou [1, Fig. 10], along with a careful explanation of why the saturation occurs. Hence, we show that there is an inherent tradeoff in III–V devices: a low effective mass yields high electron velocity and mobility, but also leads to diminishing speed improvements with scaling at small gate lengths.

To date, the simulation work carried out for predicting the f_T and $f_{\max}$ in III–V HEMTs has utilized a semiclassical framework [6], [7], [14]. While semiclassical approaches are amenable to the inclusion of electron scattering, they require the use of an effective potential profile to model the quantum well defining the channel [17], [18]. However, as will be shown in this paper, a proper prediction of the scaling behavior requires a precise modeling of the quantum well and the associated subbands, which can only be accomplished by a fully quantum-mechanical approach; electron scattering is less important at short gate lengths ($< 50\,\mathrm{nm}$), as shown in [19] by the good agreement between ballistic simulations and experiment, such that a ballistic quantum-mechanical approach is sufficient for a first-order study.

This paper is organized as follows. Section II briefly explains the NEGF simulation approach and then presents results for the intrinsic f_T of III–V HEMTs. The bias dependence of the intrinsic f_T at various gate lengths is first examined, and these results are then used to help explain the gate-length scaling behavior of the *peak* f_T by investigating the scaling behavior of the corresponding transconductance g_m and the gate capacitance C_{gg}, where $f_T = g_m/(2\pi C_{\mathrm{gg}})$. Short gate lengths of 10, 20, 30, and 50 nm are considered and the insulator thickness is fixed at 3 nm. The insulator is not scaled with the gate length since it was observed in [19] that scaling the insulator from 3.8 to 3 nm has a negligible effect on device performance. Section III then discusses the construction of a complete device model by way of an equivalent circuit that adds external parasitics to the intrinsic model; the parasitics include the series resistances arising from the heterostructure stack and metal contacts, and the external pad capacitances, which are computed from COMSOL [20] using an open-device model. The complete circuit is used to extract the extrinsic f_T and $f_{\max}$ and to examine their behavior versus gate-length scaling. Section IV summarizes the conclusions.

II. Intrinsic f_T

A. Simulation Approach

To extract the intrinsic f_T of III–V HEMTs, we employ a 2-D NEGF-Poisson solver and use the same simulation structure as in [19], shown in Fig. 1. The HEMT has a channel-layer

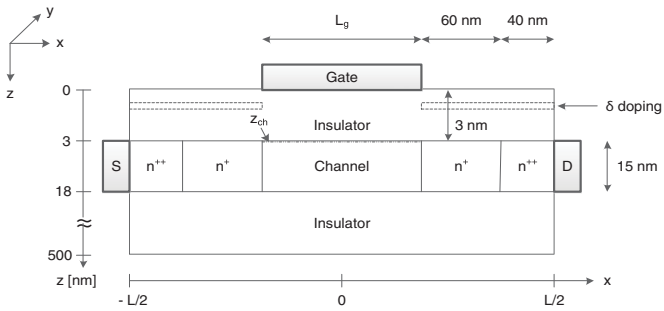


Fig. 1. III–V HEMT structure used in this paper.

thickness of 15 nm along the z-direction, and the upper and lower insulators (wide-bandgap layers) have thicknesses of 3 and 500 nm, respectively. As in [9], [12], [13], [19], and [21], InAlAs is the insulator for the InGaAs HEMTs, while AlGaN is used for the (wurtzite) GaN HEMT. The channel region is undoped, while the n^+ and n^{++} regions are electrostatically doped by means of a δ-layer having physical doping densities of $2 \times 10^{12}\ \mathrm{cm}^{-2}$ and $1 \times 10^{13}\ \mathrm{cm}^{-2}$, respectively, consistent with [19]. The value for the n^{++} regions is higher than that realized in practice, but it aids in numerical stability and otherwise does not affect the device [19].

The 2-D ballistic quantum transport in the channel, n^+, n^{++}, and upper insulator regions is described by the NEGF approach within the effective-mass approximation. The width W is assumed to be sufficiently large for the potential to be translationally invariant along the y-direction in Fig. 1. The effective-mass Hamiltonian of the device $h(x,z)$ is discretized in the 2-D (x,z) space, as outlined in [22], and the quantum open boundary conditions for transport at the source and drain terminals are modeled through the self-energies Σ_S and Σ_D, respectively [23]. The retarded Green's function [22] can then be written as $G\left[E(k_x,k_z)\right] = \left[E(k_x,k_z)I - h\left(x,z\right) - \Sigma_S - \Sigma_D\right]^{-1}$, where $E\left(k_x,k_z\right) \equiv E - E_{k_y}$ is the in-plane energy, E_{k_y} is the plane-wave energy along the width direction y, and E is the total electron energy. The Green's function $G\left[E(k_x,k_z)\right]$ can be used to evaluate the electron density $n(x,z)$ and ballistic current following the usual approach [22]–[25].

The electrostatic potential $V(x,z)$ is obtained by solving the 2-D Poisson equation $\nabla \cdot \left[\varepsilon_r(x,z)\nabla V(x,z)\right] = -\rho_V(x,z)/\varepsilon_0$, where ε_0 is the permittivity of free space, $\varepsilon_r(x,z)$ is the relative permittivity, and $\rho_V\left(x,z\right) = q\left[N_D\left(x,z\right) - n\left(x,z\right)\right]$ is the volume charge density, with $N_D(x,z)$ being the doping density and q being the magnitude of the electronic charge. Holes are presumed to be negligible in comparison to $N_D\left(x,z\right)$ and $n\left(x,z\right)$ and are hence ignored. The potential $V\left(x,z\right)$ is taken to be the vacuum potential $V\left(x,z\right) = -E_{\mathrm{vac}}(x,z)/q$, and the conduction-band edge is defined by $E_C\left(x,z\right) = E_{\mathrm{vac}}\left(x,z\right) - q\chi_{\mathrm{aff}}$, where χ_{aff} is the electron affinity. Dirichlet boundary conditions are used at the gate, whereas homogeneous Neumann boundary conditions are applied at all other boundary points.

The NEGF-Poisson system was solved self-consistently using well-established methods [22]–[25], and the solution was validated by comparing the current–voltage characteristics of

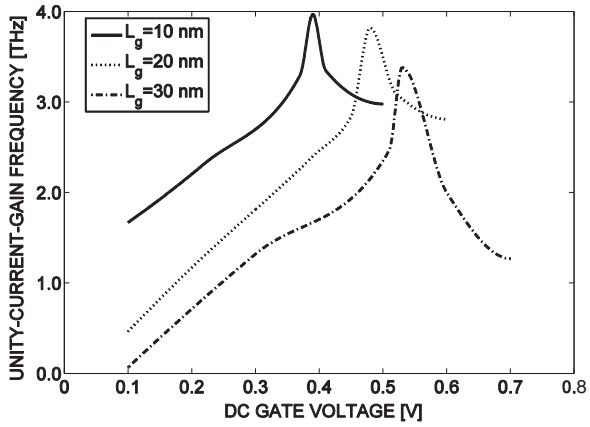


Fig. 2. Intrinsic unity-current-gain frequency f_T versus dc gate voltage V_G of an InGaAs HEMT having the structure shown in Fig. 1. The insulator thickness is $t_{\rm ins} = 3$ nm and results are shown for three different gate lengths L_g. The dc drain voltage V_D is held at 0.5 V.

an InGaAs HEMT having a gate length $L_g = 60$ nm with the results of [19].

B. Results

Results for the intrinsic f_T are discussed in this section. *For the convenience of the reader, where appropriate, we have summarized the key results from the detailed discussions; the reader may find the italicized statements near the ends of Sections II-B1(a), II-B4(e), II-B5, and II-B6 to be particularly useful.*

1) Bias Dependence of Intrinsic f_T: The intrinsic f_T can be obtained from the NEGF-Poisson results using the well-known expression

$$f_T = \frac{1}{2\pi}\frac{dI_D}{dQ_G} \tag{1}$$

where dI_D and dQ_G are the changes in the current and the magnitude of gate-electrode (or channel) charge, respectively, that result from a small change dV_G in gate voltage while the drain voltage is held fixed and the source is taken as the reference.

Fig. 2 shows the intrinsic f_T as a function of dc gate voltage V_G for an InGaAs HEMT at three different gate lengths equal to 10, 20, and 30 nm, and with the insulator thickness fixed at $t_{\rm ins} = 3$ nm. The drain bias V_D is held constant at 0.5 V, which is a typical bias voltage for HEMTs [26], [27, ch. 3], [28], [29]; HEMTs are required to operate at voltages substantially below 1 V in order to compete with Si CMOS technology [30] and to reduce the active power dissipation of the device [27, ch. 10].

Two important features of the bias dependence of the intrinsic f_T can be discerned from Fig. 2. First, for a fixed gate length L_g, i.e., for a given curve in Fig. 2, the f_T shows a significant variation with V_G, including a well-defined peak. Second, a comparison of the three curves in Fig. 2 reveals that the gate bias V_G at which the f_T peaks depends on the gate length L_g. We will now discuss each of these features in turn, and later refer back to the discussion (in Sections II-B2 and II-B5 below) to help explain the scaling behavior of the *peak* f_T.

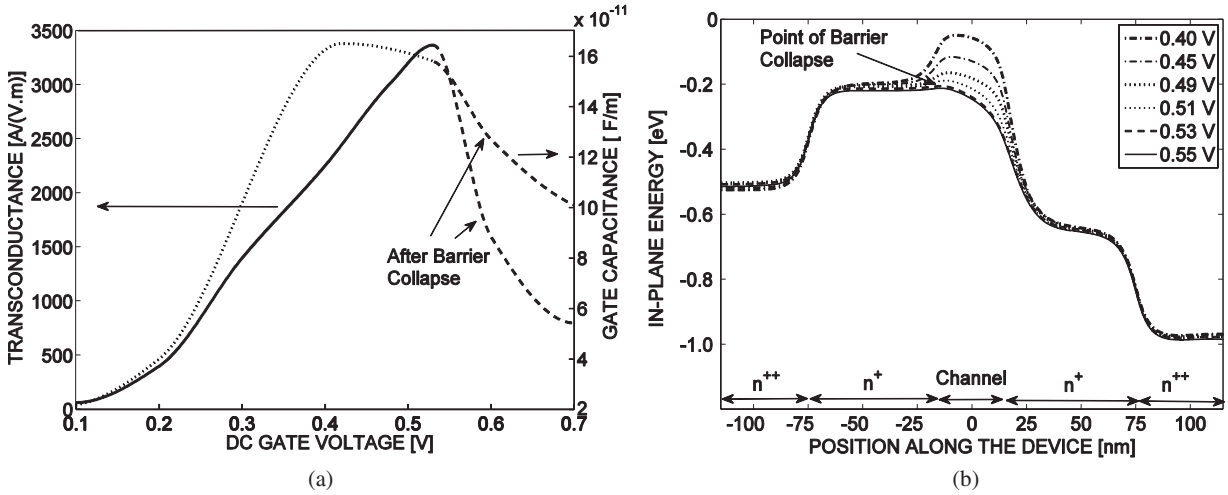
a) Variation of f_T with V_G: The variation of f_T with V_G (at a fixed L_g, i.e., for a given curve in Fig. 2) can be understood by first writing $f_T = g_m/(2\pi C_{\rm gg})$, where $g_m = dI_D/dV_G$ is the transconductance and $C_{\rm gg} = dQ_G/dV_G$ is the total intrinsic gate capacitance. Fig. 3(a) plots g_m and $C_{\rm gg}$ for the $L_g = 30$ nm case from Fig. 2. As shown, after reaching a maximum, both $C_{\rm gg}$ and g_m decrease with increasing V_G, but g_m degrades more rapidly, such that g_m controls the peaking in f_T. The trends in g_m and $C_{\rm gg}$ can be explained by analyzing the effect of gate voltage on the conduction-band edge in the channel, as depicted in Fig. 3(b) for the $L_g = 30$ nm case.

Fig. 3(b) shows the simulated conduction-band profile in the channel at different gate-bias voltages; here and elsewhere, the term "conduction-band profile in the channel" refers to $E_C(x, z_{\rm ch})$ versus x, where $z_{\rm ch}$ is a depth just below the insulator-channel interface, near the tip of the quantum well defining the channel, as marked in Fig. 1 [and Figs. 10(a) and 12 further below]. Initially, changes in the gate bias V_G are effective in pushing down the barrier at the n^+-channel junction, and correspondingly, incremental changes dV_G in gate voltage are effective in introducing more electrons into the channel. However, once the gate bias has pushed the conduction-band edge to the point of "barrier collapse," i.e., to the point where the band edge in the channel reaches the same level as that in the n^+ region near the source, as shown by the dashed curve ($V_G = 0.53$ V) in Fig. 3(b), an incremental change in gate voltage dV_G can only weakly modulate the $n^{++} - n^+$ junction [19]; the incremental change dV_G thus loses the ability to introduce new electrons into the channel. The resulting increments in channel charge dQ_G and channel current dI_D arising from dV_G are hence diminished, leading to values of $g_m = dI_D/dV_G$ and $C_{\rm gg} = dQ_G/dV_G$ that decrease beyond the point of barrier collapse.

Overall, the results in Fig. 3(a) and (b) establish that, for a fixed gate length L_g, *peak f_T occurs at the gate bias corresponding to the onset of barrier collapse*.

b) Variation in V_G for Peak f_T: The variation in the gate bias V_G at which the f_T peaks (as L_g changes, i.e., between curves in Fig. 2) can be understood to be a result of drain-induced barrier lowering (DIBL). As the gate length is scaled down, it is well known that the effect of the drain potential on the barrier gets stronger, acting as an additional source of barrier lowering [26], [27, ch. 10], [31, ch. 6]. Thus, for the same gate bias V_G, the devices with shorter gate lengths will have lower barriers, as shown by the simulation results in Fig. 4. Therefore, the scenario of barrier collapse leading to peak f_T is achieved with smaller gate bias voltages at shorter gate lengths, which explains why the locations of the peaks in Fig. 2 shift to the left as L_g is reduced.

2) Gate-Length Scaling of Intrinsic f_T: To study the scaling behavior, the intrinsic *peak* f_T is plotted as a function of the gate length in Fig. 5. It is evident that the peak f_T of III–V HEMTs shows a signature saturation as the gate length L_g is scaled down, an outcome that can also be discerned from the experimental results collected by Schwierz and Liou [1, Fig. 10]. The saturation can be explained by analyzing the scaling behavior of the intrinsic g_m and $C_{\rm gg}$ [since $f_T = g_m/(2\pi C_{\rm gg})$], where the relevant values of g_m and $C_{\rm gg}$ are those at the gate bias corresponding to the onset of barrier



(a) (b)

Fig. 3. (a) Transconductance g_m and gate capacitance $C_{\rm gg}$ versus gate bias V_G. (b) Conduction-band profile in the channel, i.e., $E_C(x, z_{\rm ch})$ versus x, at different gate-bias voltages for the InGaAs HEMT with $L_g = 30$ nm considered in Fig. 2. Dashed lines are used in part (a) for gate voltages above the value causing barrier collapse.

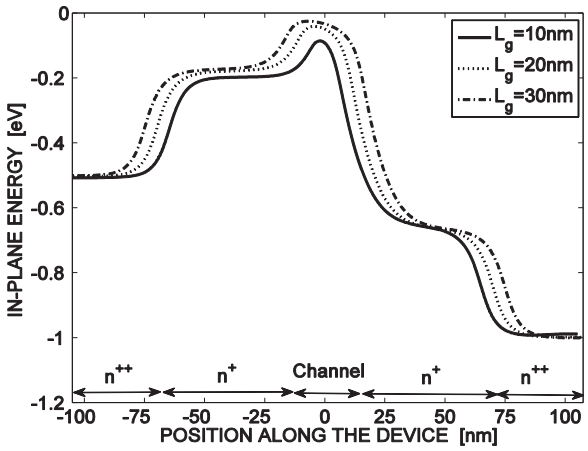


Fig. 4. Conduction-band profile, i.e., $E_C(x, z_{\rm ch})$ versus x, of the InGaAs HEMTs considered in Fig. 2 at $V_G = 0.3$ V, illustrating the impact of DIBL.

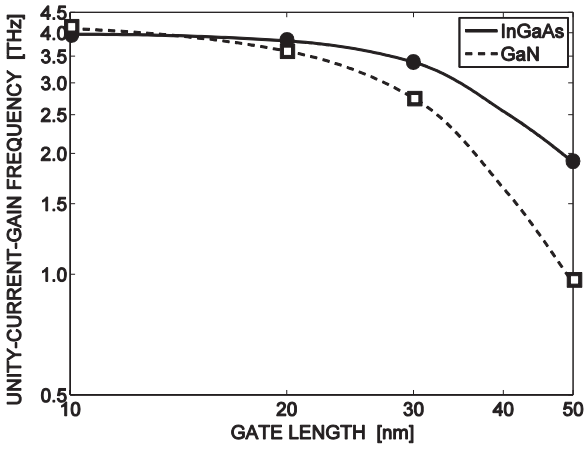


Fig. 5. Intrinsic *peak* f_T versus gate length L_g for the III–V HEMTs considered in this paper. The lines have been drawn as guides for the eye.

collapse, i.e., those leading to the peak f_T at each L_g, as discussed in the previous section. In what follows, we first establish the capacitive input equivalent circuit seen looking into the gate, and then examine the scaling behavior of g_m and $C_{\rm gg}$, referring to the circuit as an aid when appropriate; the observations are then used to explain the relative scaling behavior of InGaAs and GaN HEMTs, based on the difference in the effective mass of these materials.

3) Input Equivalent Circuit: Consider first the input equivalent circuit seen from the gate under the conditions of a perturbation in gate voltage dV_G, as sketched in the different parts of Fig. 6, where the top terminal is the gate and the bottom terminal is the shorted source and drain combination, taken here as the reference, since both terminals are at ac ground under the conditions needed for an f_T extraction. In its simplest form, the input circuit is just the total input capacitance $C_{\rm gg} = dQ_G/dV_G$, as shown in Fig. 6(a). However, it is well known that $C_{\rm gg}$ can be modeled as a series combination of insulator capacitance $C_{\rm ins}$ and the so-called inversion-layer capacitance[1] $C_{\rm inv}$ [27, ch. 3], [32], with $C_{\rm ins}$ and $C_{\rm inv}$ each being found as an integrated value of a change in charge with respect to potential along the channel:

$$C_{\rm ins} = \int_{-L/2}^{L/2} \frac{dQ_G(x)}{[dV_G + (1/q)\, dE_C(x, z_{\rm ch})]} dx \qquad (2)$$

and

$$C_{\rm inv} = \int_{-L/2}^{L/2} \frac{dQ_G(x)}{-(1/q)\, dE_C(x, z_{\rm ch})} dx \qquad (3)$$

where $dQ_G(x)$ and $dE_C(x, z_{\rm ch})$ represent the changes (due to dV_G) in gate-electrode charge and conduction-band edge, respectively, at each point x, and L is the total length of the

[1]It is worth noting that the term "inversion-layer capacitance" used in the context of HEMTs is equivalent to the term "quantum capacitance" used in the context of emerging transistors [35, ch. 7], [53], [54]. With HEMTs, the term "quantum capacitance" is used in a different way [32].

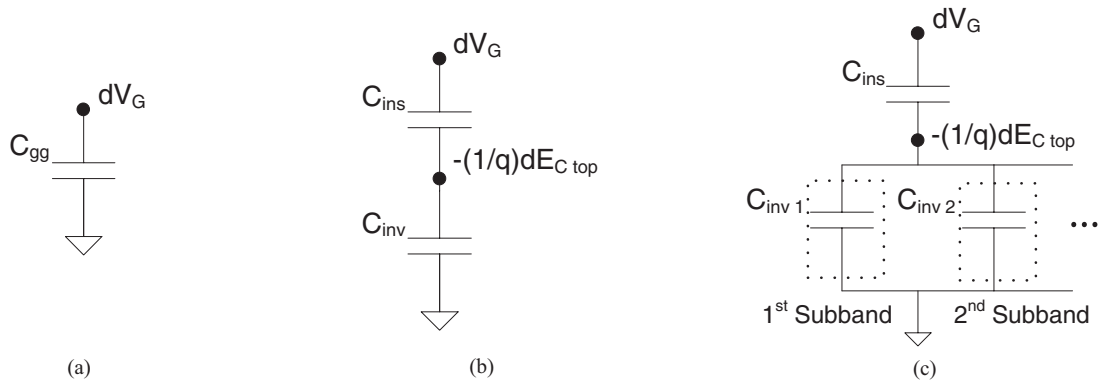


Fig. 6. Input equivalent circuit of the HEMTs. (a) Overall input circuit, which is just the input capacitance $C_{\rm gg}$. (b) Separation of $C_{\rm gg}$ into the series combination of insulator and inversion-layer capacitances, $C_{\rm ins}$ and $C_{\rm inv}$. (c) Further subdivision of $C_{\rm inv}$ into contributions arising from each subband, where $C_{{\rm inv}i}$ is the contribution from the ith subband.

device from the source to the drain, as depicted in Fig. 1. A simple representation of $C_{\rm gg}$ is then given by the circuit in Fig. 6(b), where $dE_{C\rm top}$ is the change in the conduction-band edge at the *top* of the n^+-channel barrier [33], located at a point $(x,z)=(x_{\rm top},z_{\rm ch})$; this circuit applies in a lumped model of a ballistic device even when the conduction-band edge is not flat across the channel [34]. An alternative representation is provided in Fig. 6(c), where $C_{\rm inv}$ is further subdivided into a parallel combination of capacitances arising from the occupied subbands, with $C_{{\rm inv}i}$ representing the inversion-layer capacitance from the ith subband, and where the subbands themselves arise due to vertical confinement within the channel (i.e., in the z-direction of Fig. 1).

4) Gate-Length Scaling of g_m:

a) Expression for g_m: Consider now the scaling behavior of the transconductance g_m at peak f_T. From its definition, and utilizing the expression for current within the NEGF formalism [22], the intrinsic g_m can be written as

$$g_m=\frac{dI_D}{dV_G}=\frac{d}{dV_G}\left(\frac{q}{\hbar^2}\sqrt{\frac{m^*k_BT}{2\pi^3}}\int_{-\infty}^{\infty}T\left[E(k_x,k_z)\right]\times\left\{F_{-(1/2)}\left[\mu_S-E(k_x,k_z)\right]-F_{-(1/2)}\left[\mu_D-E(k_x,k_z)\right]\right\}dE(k_x,k_z)\right)\tag{4}$$

where m^* is the electron effective mass, $T\left[E(k_x,k_z)\right]$ is the total transmission function at an in-plane energy $E(k_x,k_z)$, μ_S is the source Fermi level, μ_D is the drain Fermi level, $F_{-(1/2)}$ is the Fermi–Dirac integral of order $-1/2$,

$$F_{-(1/2)}(\theta)=\int_0^{\infty}\frac{\eta^{-1/2}}{1+\exp[\eta-(\theta/k_BT)]}d\eta\tag{5}$$

and it is to be understood that the right side of (4) and all subsequent expressions in this discussion should be evaluated at the gate bias corresponding to peak f_T.

For the purpose of examining the scaling behavior, the circuit in Fig. 6(b) can be exploited to substitute $dV_G=\left[\left(C_{\rm ins}+C_{\rm inv}\right)/C_{\rm ins}\right](-1/q)\,dE_{C\rm top}$ into (4), yielding

$$g_m=\frac{C_{\rm ins}}{C_{\rm ins}+C_{\rm inv}}\times\frac{d}{(-1/q)dE_{C\rm top}}\left(\frac{q}{\hbar^2}\sqrt{\frac{m^*k_BT}{2\pi^3}}\int_{-\infty}^{\infty}T\left[E\left(k_x,k_z\right)\right]\times\left\{F_{-(1/2)}\left[\mu_S-E\left(k_x,k_z\right)\right]-F_{-(1/2)}\left[\mu_D-E\left(k_x,k_z\right)\right]\right\}dE\left(k_x,k_z\right)\right).\tag{6}$$

Equation (6) then reveals that the scaling behavior of the transconductance depends on the scaling behavior of two quantities. The first is the capacitance ratio $C_{\rm ins}/\left(C_{\rm ins}+C_{\rm inv}\right)$, which reflects (through voltage division) the ability of a perturbation in gate voltage dV_G to move the conduction-band edge at the top of the barrier by an amount $dE_{C\rm top}$. The second is the change in the integrated electron current for a given change $dE_{C\rm top}$, as specified by the remaining factor, i.e., the derivative in (6); the key elements in this derivative are the transmission function $T\left[E\left(k_x,k_z\right)\right]$, which reflects the likelihood that an electron incident from the source with an in-plane energy $E\left(k_x,k_z\right)$ will be able to reach the drain, and the difference in Fermi–Dirac integrals, which reflects the "difference in agenda" [35, ch. 1] between the source and drain contacts at each in-plane energy $E\left(k_x,k_z\right)$.

To gain further insight from (6), we note that under ballistic transport, the transmission function can be approximated as a sum of unit-step functions, with the steps occurring at those energies corresponding to the subband edges at the top of the barrier:

$$T\left[E\left(k_x,k_z\right)\right]=\sum_i u\left[E\left(k_x,k_z\right)-E_{C\rm top}-\Delta_i\right]\tag{7}$$

where Δ_i is the bottom edge of subband i with respect to $E_{C\rm top}$. The use of (7) in (6) then leads to the following result for g_m:

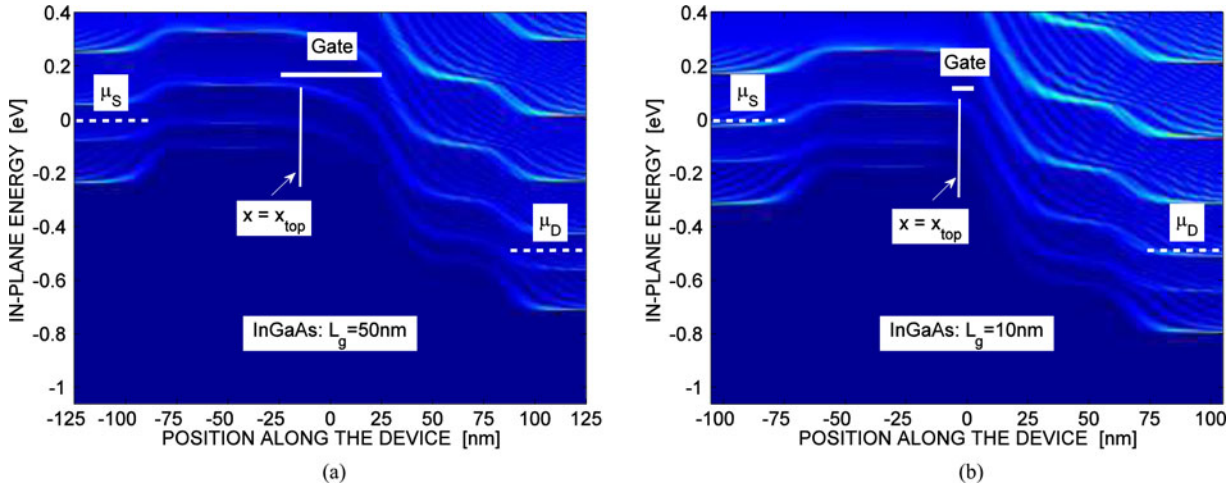


Fig. 7. Total spectral function $A_S\left[x, z_{\rm ch}, E\left(k_x, k_z\right)\right] + A_D\left[x, z_{\rm ch}, E\left(k_x, k_z\right)\right]$ displayed as intensity versus $E\left(k_x, k_z\right)$ and x, along with the positions of the Fermi levels μ_S and μ_D, for the InGaAs HEMT with (a) $L_g = 50$ nm and (b) $L_g = 10$ nm, showing that at the shorter gate length, the subband edges at peak f_T are lower in position with respect to the source Fermi level.

$$g_m = \frac{q^2}{\hbar^2}\sqrt{\frac{m^* k_B T}{2\pi^3}}\frac{C_{\rm ins}}{C_{\rm ins}+C_{\rm inv}}\sum_i \left\{F_{-(1/2)}\left[\mu_S - E_{C\,\rm top} - \Delta_i\right] - F_{-(1/2)}\left[\mu_D - E_{C\,\rm top} - \Delta_i\right]\right\}. \quad (8)$$

According to (8), the scaling behavior of g_m thus ultimately depends on the scaling behavior of the capacitance ratio $C_{\rm ins}/\left(C_{\rm ins}+C_{\rm inv}\right)$, which describes the ability of the gate to modulate the top of the barrier, and on the difference in Fermi–Dirac integrals *evaluated at each subband edge* $\varepsilon_i \equiv E_{C\rm top} + \Delta_i$, which describes the "difference in agenda" of the source and drain contacts. As we now discuss, the capacitance ratio and the difference in Fermi–Dirac integrals at peak f_T change only weakly with scaling and in opposition to each other, such that the corresponding g_m remains relatively insensitive to scaling.

b) Position of subband edges: To describe this outcome, it is necessary to follow the relative positions of the subband edges at peak f_T as the gate length is scaled down. While a detailed explanation of the phenomenon will be provided in Section II-B5, for the present discussion, it suffices to note that the edges of the first few subbands at peak f_T will be located *further below* the source Fermi level as the gate length is scaled down. This result can be discerned from the plots of the spectral functions in Fig. 7. The figure shows the total spectral function in the channel versus in-plane energy and position, i.e., $A_S\left[x, z_{\rm ch}, E\left(k_x, k_z\right)\right] + A_D\left[x, z_{\rm ch}, E\left(k_x, k_z\right)\right]$ displayed as an intensity versus $E\left(k_x, k_z\right)$ and x, for the two extreme gate-length InGaAs devices, i.e., the 10- and 50-nm devices, at the gate bias corresponding to peak f_T. An inspection of the plots reveals that the subband edges (indicated by the brightly shaded regions) under the gate do indeed move down with respect to μ_S as the gate length is scaled from 50 to 10 nm.

c) Impact on Fermi–Dirac integrals: The impact of a change in the relative positions of the subband edges on the Fermi–Dirac integrals in (8) is shown in Fig. 8(a), where the subband edges at the top of the barrier and at the gate bias corresponding to peak f_T are superimposed on a sketch of the difference $\{F_{-(1/2)}\left[\mu_S - E(k_x, k_z)\right] - F_{-(1/2)}\left[\mu_D - E(k_x, k_z)\right]\}$. The shift in the positions of the subbands with downscaling causes an enhanced contribution to the difference in the Fermi–Dirac integrals evaluated at each subband edge, i.e., there is a greater difference in agenda between the source and drain contacts at peak f_T for each subband as the gate length is scaled down, and this will tend to increase g_m.

d) Impact on capacitance ratio: By contrast, as the gate length is scaled down, the lower position of the subband edges causes the capacitance ratio $C_{\rm ins}/\left(C_{\rm ins}+C_{\rm inv}\right)$ to *decrease*, which will tend to *decrease* g_m. The decrease in the capacitance ratio can be understood by noting that $C_{\rm ins}$ is primarily determined by the insulator thickness, which is fixed in this paper, causing $C_{\rm ins}$ to scale linearly with gate length, while $C_{\rm inv}$ is larger at each gate length than would be expected from a purely linear dependence on L_g.

To understand the behavior of $C_{\rm inv}$, we note that the charge at node (x, z) from the NEGF formalism [22] can be represented as an appropriate integral (over energy) of $F_{-(1/2)}$ times the source and drain spectral functions (local densities of states):

$$\begin{aligned} q \times n\left(x,z\right) &= \frac{q}{ab}\sqrt{\frac{m^* k_B T}{2\pi^3\hbar^2}}\int_{-\infty}^{\infty}\left\{F_{-(1/2)}\left[\mu_S - E\left(k_x,k_z\right)\right]\right. \\ &\quad \times A_S\left[x,z,E\left(k_x,k_z\right)\right] + F_{-(1/2)}\left[\mu_D - E\left(k_x,k_z\right)\right] \\ &\quad \left.\times\ A_D\left[x,z,E\left(k_x,k_z\right)\right]\right\} dE\left(k_x,k_z\right) \quad (9a) \\ &\approx \frac{q}{ab}\sqrt{\frac{m^* k_B T}{2\pi^3\hbar^2}}\int_{-\infty}^{\infty} F_{-(1/2)}\left[\mu_S - E\left(k_x,k_z\right)\right] \\ &\quad \times A_S\left[x,z,E\left(k_x,k_z\right)\right] dE\left(k_x,k_z\right) \quad (9b) \end{aligned}$$

where a and b are the grid sizes along the x- and z-directions, respectively, A_S and A_D are the spectral functions due to the

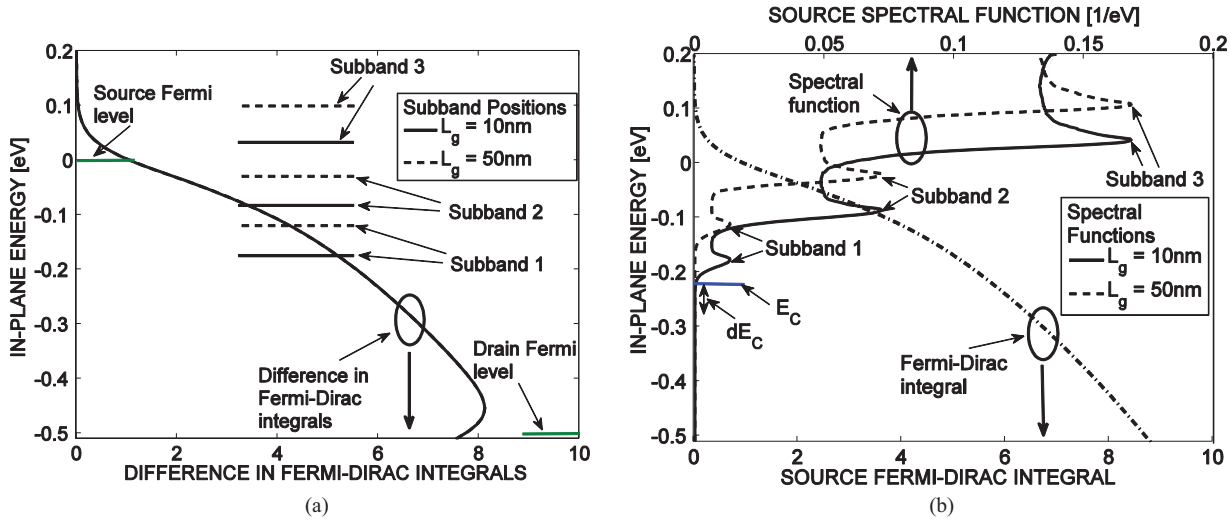


Fig. 8. (a) Difference in the Fermi–Dirac integrals $\left\{F_{-(1/2)}\left[\mu_S - E(k_x,k_z)\right] - F_{-(1/2)}\left[\mu_D - E(k_x,k_z)\right]\right\}$, along with the positions of the occupied subband edges at the top of the barrier, $(x,z) = (x_{\rm top}, z_{\rm ch})$. (b) Source Fermi–Dirac integral $F_{-(1/2)}\left[\mu_S - E\left(k_x,k_z\right)\right]$ and source spectral function $A_S\left[x_{\rm top}, z_{\rm ch}, E\left(k_x,k_z\right)\right]$ versus in-plane energy $E(k_x,k_z)$ for the InGaAs HEMT with gate lengths $L_g = 10$ and 50 nm; the source Fermi level is $\mu_S \equiv 0$, and the conduction-band edge E_C is marked.

source and drain contacts, respectively, and μ_D is assumed to lie sufficiently below μ_S for the second term in the integrand of (9a) to be neglected in comparison with the first at all points x within the channel to yield (9b) (which will be true at typical drain bias voltages). The inversion capacitance $C_{\rm inv}$ is then given by the prescription in (3), where $dQ_G(x)$ is equal to the right side of (9b) integrated over z in the channel. Rather than integrate over z, it is more instructive to examine the trends in the charge located at the single point corresponding to the top of the barrier, i.e., the point $(x,z) = (x_{\rm top}, z_{\rm ch})$, which will be representative of the trends in $dQ_G(x)$. Fig. 8(b) shows plots of the key factors of (9b) found in this way, at peak f_T and for the same InGaAs device and gate lengths considered in Fig. 7.

Plotted in Fig. 8(b) are hence the Fermi–Dirac integral $F_{-(1/2)}\left[\mu_S - E\left(k_x,k_z\right)\right]$ and the spectral function $A_S\left[x_{\rm top}, z_{\rm ch}, E\left(k_x,k_z\right)\right]$; the Fermi–Dirac integral looks the same at both gate lengths, with $\mu_S \equiv 0$ taken as the reference, so that the only difference is in the spectral function. As expected, the spectral function shows a sequence of $1/\sqrt{E\left(k_x,k_z\right)}$ dependencies that are consistent with the existence of a sequence of 1-D subbands arising from confinement in the z-direction, i.e., consistent with a dispersion relation of the form $E\left(k_x,k_z\right) = \varepsilon_n + \frac{\hbar^2}{2m^*}k_x^2$; the subband edges are marked in the figure. In each case, the charge $q \times n\left(x_{\rm top}, z_{\rm ch}\right)$ is given by the area under the *overlap* of the Fermi–Dirac integral and the spectral function; more importantly, under dynamic conditions, i.e., under a modulation dE_C in E_C, which can be visualized as shifting the spectral functions, the *incremental* overlap and hence the *incremental* charge is greater in the shorter gate-length device. This result is not immediately obvious, but it can be discerned from the figure for two reasons: 1) the incremental charge will be greater in the first two subbands of the shorter gate-length device, since they are further below the source Fermi level and hence experience a greater population modulation (since the Fermi–Dirac integral is greater); and 2) the third subband will participate in the shorter gate-length device, whereas in the longer gate-length device, it is too far above the source Fermi level to hold any charge or to participate in charge modulation, i.e., the associated states are always empty. As a result of the greater charge modulation with a modulation dE_C in E_C, the capacitance $C_{\rm inv}$ is larger in the smaller device than would be expected from a purely linear dependence on gate length, and the capacitance ratio $C_{\rm ins}/\left(C_{\rm ins} + C_{\rm inv}\right)$ thus *decreases* with a downscaling of gate length.

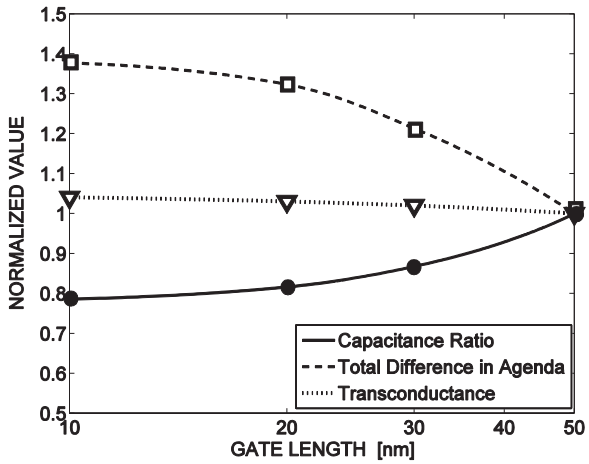


Fig. 9. Capacitance ratio $C_{\rm ins}/(C_{\rm ins} + C_{\rm inv})$, total difference in Fermi–Dirac integrals ("difference in agenda"), i.e., the result of the summation in (8), and the transconductance g_m (at peak f_T) versus gate length L_g for the InGaAs HEMT. The actual values have been normalized with respect to the value of each component at the gate length $L_g = 50$ nm in order to illustrate the scaling behavior, and the lines have been drawn as guides for the eye.

e) Overall scaling of g_m: Overall, the capacitance ratio and the difference in Fermi–Dirac integrals in (8) thus act in opposition to each other as the gate length is scaled down, as shown in Fig. 9, leading to a g_m (at peak f_T) that does not scale significantly with gate length, as also shown in Fig. 9. *In*

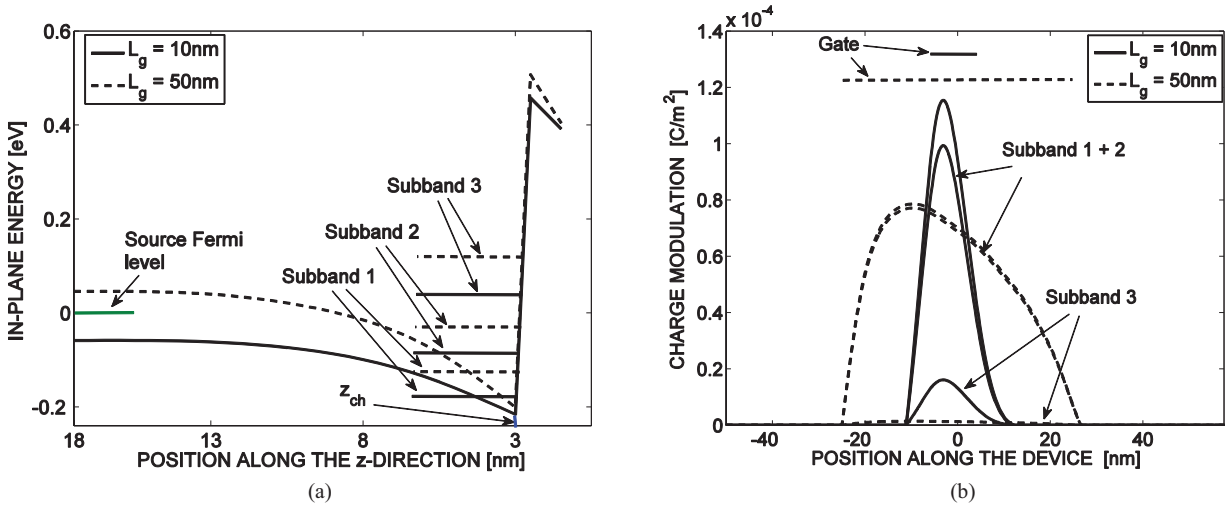


Fig. 10. (a) Conduction-band profiles at peak f_T along the depth of the channel, i.e., $E_C\,(x_{\rm top}, z)$ versus z, for the InGaAs HEMTs with $L_g = 10$ and 50 nm; also shown are the positions of the occupied subband edges at $(x, z) = (x_{\rm top}, z_{\rm ch})$, illustrating that the subbands are lower in position with respect to the source Fermi level for the shorter gate-length device. (b) Charge modulation along the transport direction (i.e., the x-direction) for the InGaAs HEMTs with $L_g = 10$ and 50 nm; the results show a relatively greater charge modulation (under the gate) in the first three subbands of the shorter gate-length device.

essence, as the gate length is scaled down, the subband positions at peak f_T move lower in energy with respect to the source Fermi level, causing the difference in agenda between the source and drain contacts to increase for each subband (reflected by the greater difference in Fermi–Dirac integrals) but with this increase being offset by a weaker control of the channel barrier by the gate (reflected by the decreased capacitance ratio). It is worth mentioning that a similar trend in g_m was experimentally observed in [36, Fig. 6] for an InAs HEMT with $t_{\rm ins} =$ 4 nm, where scaling the gate length below 90 nm resulted in an insignificant improvement (by only a few percent) in g_m.

Two points are worth adding regarding the insensitivity of g_m to gate-length scaling. First, we found the same outcome even when the insulator thickness was scaled (from 3 nm down to 2 nm) with gate length, where the g_m improved only slightly (by a few percent) and where the same tradeoffs occurred. Second, while we have used the InGaAs devices to illustrate the result, a similar trend in g_m can be expected irrespective of the material, i.e., irrespective of the precise value of the effective mass; this follows because the compensating increase and decrease leading to an insensitivity of g_m to gate length will always occur, with only the extent of the increase and decrease varying between materials.

Since the g_m is relatively constant with gate length, the scaling behavior of $f_T = g_m/(2\pi C_{\rm gg})$ is determined by the scaling behavior of $C_{\rm gg}$.

5) Gate-Length Scaling of C_{gg}: In contrast to the behavior of g_m, the total gate capacitance $C_{\rm gg}$ at peak f_T is significantly affected by the scaling of gate length.

As previously illustrated in Fig. 6(b), $C_{\rm gg}$ can be modeled as a series combination of insulator and inversion capacitance:

$$C_{\rm gg} = \frac{C_{\rm ins}C_{\rm inv}}{C_{\rm ins} + C_{\rm inv}}. \tag{10}$$

Since III–V materials have a relatively low density of states (due to a relatively small electron effective mass) and are fabricated with very thin, high-k insulators, then $C_{\rm inv}$ is significantly smaller than $C_{\rm ins}$; for example, the ratio $C_{\rm inv}/C_{\rm ins}$ ranged from 0.05 to 0.26 for the InGaAs and 0.25 to 0.39 for the GaN device when the gate length was scaled from 50 to 10 nm. Hence, to a first approximation, $C_{\rm gg} \sim C_{\rm inv}$, and $C_{\rm inv}$ controls the scaling behavior of $C_{\rm gg}$.

As explained earlier, DIBL causes the peak f_T to occur at a smaller gate bias V_G for the shorter gate-length devices. As a consequence, the channel charge is less tightly held near the insulator interface by V_G, or equivalently, the quantum well defining the channel is *less sharp*; this can be observed from the plot of conduction-band profiles along the depth of the channel (along the z-direction) at $x = x_{\rm top}$, i.e., from a plot of $E_C(x_{\rm top}, z)$ versus z, as illustrated in Fig. 10(a), which shows results for the same InGaAs device and gate lengths considered in Fig. 7. For the shorter gate length, the triangular shape of the well is less pronounced, i.e., there is reduced quantum confinement, and hence the subbands lie closer to the conduction-band edge at the tip of the well [37, ch. 1], i.e., closer to $E_C\,(x_{\rm top}, z_{\rm ch})$. Since the position of $E_C\,(x_{\rm top}, z_{\rm ch})$ relative to μ_S at peak f_T is fixed, which follows because the latter always occurs at the onset of barrier collapse, i.e., when E_C in the channel aligns with that in the n^+ region near the source [see Fig. 3(b)], it then follows that the subbands at peak f_T are lower in position with respect to μ_S at shorter gate lengths. This result, already mentioned previously in Section II-B4, is readily seen from the data in Fig. 10(a). Since the subband positions are lower with respect to μ_S at shorter gate lengths, then as already discussed in conjunction with Fig. 8(b), a modulation dE_C in E_C leads to a relatively larger charge modulation (at each point x under the gate) in the first three subbands. This outcome is illustrated in Fig. 10(b) for the same InGaAs devices considered in Fig. 7.

Therefore, at shorter gate lengths, there is a relatively larger overall charge modulation than would be expected from a purely linear dependence on gate length, causing $C_{\rm inv}$ to be larger

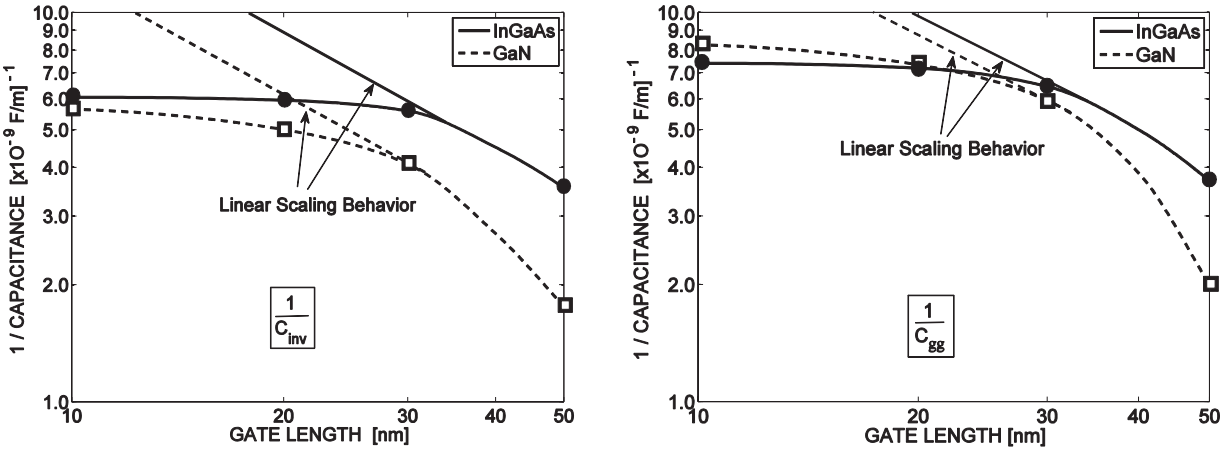


Fig. 11. (a) Reciprocal of inversion-layer capacitance $1/C_{\rm inv}$ and (b) reciprocal of total gate capacitance $1/C_{\rm gg}$ versus gate length L_g for the HEMTs. The lines have been drawn as guides for the eye, and to emphasize the saturating behavior at short gate lengths, additional lines are shown to illustrate the expected behavior based on linear scaling with L_g, extrapolated to lower gate lengths from $L_g \sim 35$ nm.

than would be expected from a purely linear dependence on gate length. This behavior of $C_{\rm inv}$ with scaling is illustrated in Fig. 11(a), where we have plotted $1/C_{\rm inv}$; as shown, the larger than expected values of $C_{\rm inv}$ cause the behavior of $1/C_{\rm inv}$ to saturate at short gate lengths. In turn, $1/C_{\rm gg} \sim 1/C_{\rm inv}$ behaves in a similar manner, as shown in Fig. 11(b).

Since $1/C_{\rm gg}$ saturates, while the corresponding g_m remains almost constant (Section II-B4), *the peak $f_T = g_m/(2\pi C_{\rm gg})$ in III–V materials will exhibit a signature saturation with a downscaling of gate length, as we depicted in Fig. 5 and as originally observed by Schwierz and Liou* [1, Fig. 10].

6) Impact of Channel-Material Effective Mass: Among the HEMT channel materials considered in this paper, InGaAs has the smaller effective mass in comparison to wurtzite GaN ($0.048 \times m_0$ [19] versus $0.2 \times m_0$ [38], [39], where m_0 is the electron rest mass). While many factors can impact the relative positions of the subbands at the point of barrier collapse and hence at peak f_T, we note that *when comparing materials*, the most important consideration is the requirement that E_C in the channel be aligned with that in the n^+ region near the source; this requirement then implies that the charge at the top of the barrier (integrated over the extent of the channel in the z-direction) be equal (to first order) to that in the n^+ region near the source [19], where the latter is determined by the doping concentration and is hence the same for the two HEMTs under consideration.

To accommodate the required charge at the top of the barrier, the subband edges in InGaAs—which has the lower effective mass and hence the lower density of states—must move further below the source Fermi level than those in GaN; this outcome is illustrated in the plot of conduction-band profiles along the z-direction at $x = x_{\rm top}$ in Fig. 12 for the InGaAs and GaN HEMTs having a gate length $L_g = 10$ nm. For reasons already discussed previously in Sections II-B4 and II-B5, the inversion capacitance $C_{\rm inv}$ in InGaAs, and hence the total gate capacitance $C_{\rm gg} \sim C_{\rm inv}$ in InGaAs, will thus exhibit a greater deviation from purely linear gate-length scaling than in GaN; equivalently, the reciprocal capacitances $1/C_{\rm inv}$ and $1/C_{\rm gg}$ will experience a more pronounced saturation at shorter gate lengths in InGaAs versus GaN, as depicted in the two parts of Fig. 11. Since the saturation of $1/C_{\rm gg}$ is more pronounced in the lighter mass material, the peak $f_T = g_m/(2\pi C_{\rm gg})$ experiences a more pronounced saturation, as shown in Fig. 5. More generally, these results illustrate that *devices with a small effective mass will suffer from a more rapid saturation in peak f_T with a downscaling of gate length in comparison with devices having a heavy effective mass.*

III. Complete Device

To examine the scaling behavior of the *extrinsic* RF metrics, the NEGF-Poisson solver was utilized to find the components of a small-signal equivalent circuit for the intrinsic device, which was then combined with the parasitic elements to develop a complete (extrinsic) device model. This model was then used to extract the device y-parameters and subsequently the extrinsic f_T and $f_{\rm max}$.

A. Small-Signal Equivalent Circuit

The small-signal equivalent circuit for an intrinsic III–V HEMT is the classical circuit usually used for field-effect transistors [31, Fig. 9.5], shown here in Fig. 13(a), with the definitions of the elements (which have their usual meanings) provided in Fig. 13(b). The source/drain charge partitioning factor χ is not critical to the results of interest in this paper; for completeness, we have chosen $\chi = 0.4$, which is the value suggested in [40] and a value that is also commonly used for MOSFETs. The values of all the other parameters, for both the InGaAs and GaN HEMTs, are provided in Table I; values are listed for the $L_g = 30$ nm case as a representative example.

B. Modeling of Parasitics

Fig. 14 shows the device structure used to extract the parasitic resistances and capacitances; the structure is consistent with those reported in [19], [21], [41], and [42]. The source and

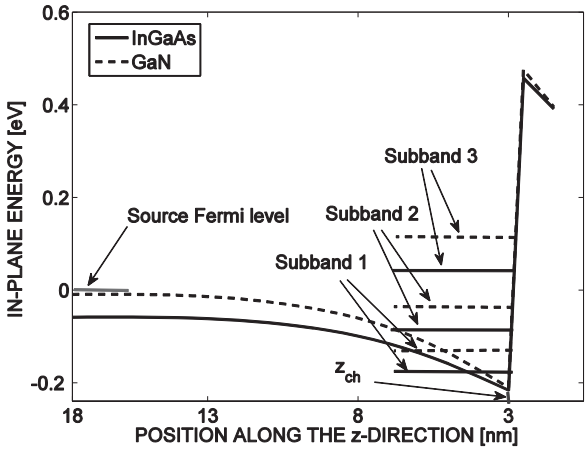


Fig. 12. Conduction-band profiles at peak f_T along the depth of the channel, i.e., $E_C(x_{\text{top}}, z)$ versus z, for the InGaAs and GaN HEMTs with $L_g = 10$ nm; the subbands for the InGaAs device are lower in position with respect to the source Fermi level.

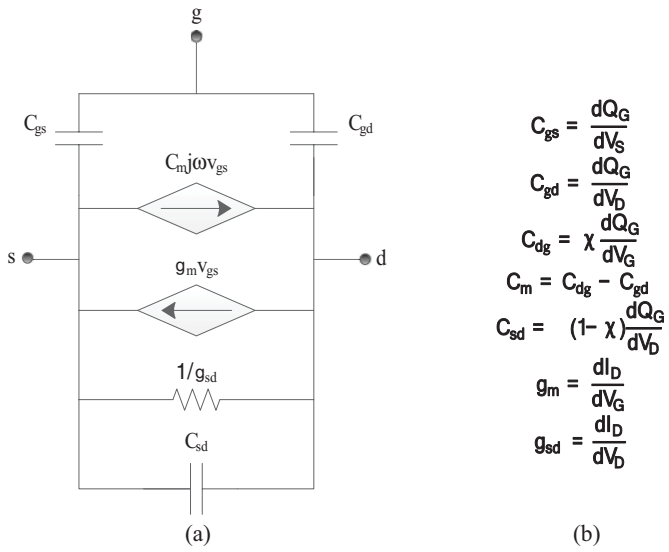


Fig. 13. (a) Small-signal equivalent circuit of an intrinsic III–V HEMT from [31, Fig. 9.5]. (b) Circuit elements, to be evaluated at the dc operating point; the symbols bear their usual meanings.

TABLE I
ELEMENT VALUES FOR THE CIRCUIT OF FIG. 13 WHEN $L_g = 30$ nm

	C_{gs} [F/m]	C_{gd} [F/m]	C_m [F/m]	C_{sd} [F/m]	g_m [S/m]	g_{sd} [S/m]
InGaAs HEMT	1.08×10^{-10}	0.46×10^{-10}	0.15×10^{-10}	0.27×10^{-10}	3.30×10^{3}	1.20×10^{2}
GaN HEMT	1.14×10^{-10}	0.47×10^{-10}	0.17×10^{-10}	0.28×10^{-10}	2.78×10^{3}	1.25×10^{2}

drain metal contacts are 50 nm in both length and height, and they are placed 1 μm apart on InGaAs/InAlAs and GaN/InGaN heterostructure stacks for the InGaAs and GaN HEMTs, respectively, with each stack having a thickness of 15 nm. The height of the gate metal is 150 nm, with the lower and upper parts having lengths of L_g and L_g+300 nm, respectively, where L_g varies

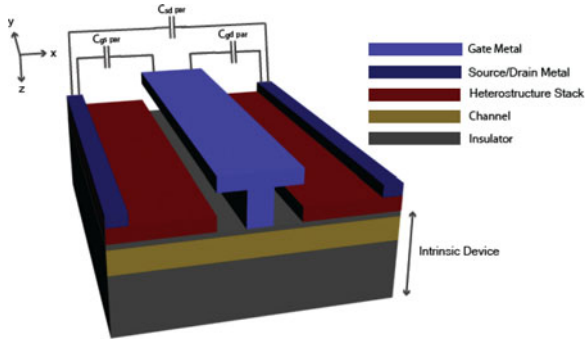


Fig. 14. Full device structure of the HEMTs.

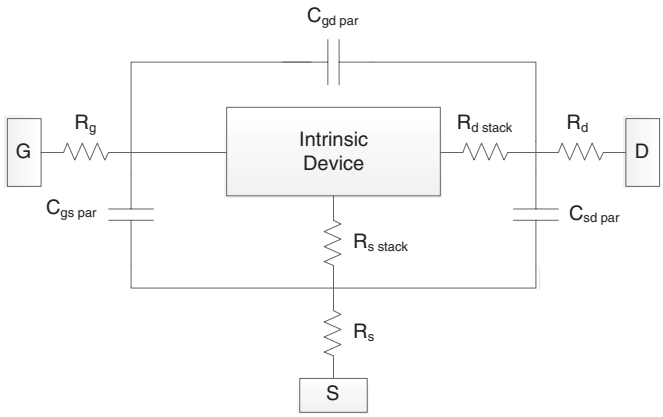


Fig. 15. Complete device model of the III–V HEMTs. The intrinsic device is represented by the small-signal circuit of Fig. 13.

from 50 to 10 nm with scaling. The extent of the intrinsic device is indicated in the figure, and the width W in the y-direction is considered to be 1 μm in this study.

The gate resistance R_g arising from the gate metal is considered to be distributed in nature, similar to MOSFETs, and it is modeled by a single lumped and effective resistance in series with the gate lead, as in [43]. Ni/Au and Ti/Pt/Au are assumed to be the gate metals for the InGaAs and GaN HEMTs, as in [44] and [45], respectively. Although very small, the source and drain metal resistances, R_s and R_d, respectively, are also considered in the complete device; Ni/Ge/Au and Ti/Al are used as the source and drain contact metals, as in [21] and [45], for the InGaAs and the GaN HEMTs, respectively. In each case, instead of modeling the multilayer nature of the contact metals, only the lowermost metal layer is considered to represent the entire contact; this assumption can be justified by the fact that the lowermost metal layer is the material that sets the work function and thus controls the overall behavior of the contact.

The parasitic capacitances between the contact metals are extracted by designing an "open-device" model [46] for the HEMTs, where the structure is exactly like the actual device but with zero charge in the intrinsic portion, and then using COMSOL as outlined in [43]. These parasitic capacitances are represented as $C_{\text{gs par}}$, $C_{\text{gd par}}$, and $C_{\text{sd par}}$ in the complete device model, the topology of which is based on [31, Fig. 8.30] and which is shown in Fig. 15.

TABLE II
PARASITIC ELEMENT VALUES OF THE HEMTS (FOR $W = 1\,\mu$m)

	$C_{\text{gs par}}$ [F]	$C_{\text{gd par}}$ [F]	$C_{\text{sd par}}$ [F]	R_g [for L_g = 30 nm] [Ω]	R_s [Ω]	R_d [Ω]	$R_{s\,\text{stack}}$ [Ω]	$R_{d\,\text{stack}}$ [Ω]
InGaAs HEMT	3.41×10^{-17}	3.37×10^{-17}	9.29×10^{-18}	3.69	0.072	0.072	200	200
GaN HEMT	2.56×10^{-17}	2.53×10^{-17}	6.27×10^{-18}	20.51	0.4	0.4	400	400

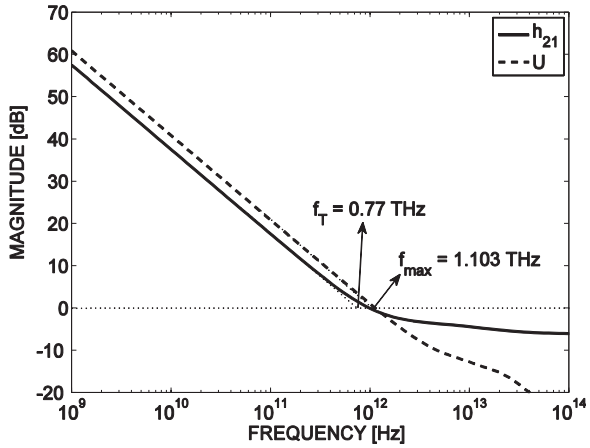


Fig. 16. Current gain $|h_{21}|$, unilateral power gain U, and extrapolated f_T and $f_{\max}$ of the InGaAs HEMT at $L_g = 30$ nm.

Apart from the aforementioned parasitic elements, the heterostructure stacks at the source and drain can also impact the performance of the HEMTs. This effect can be modeled simply by two series resistances, $R_{s\,\text{stack}}$ and $R_{d\,\text{stack}}$, connected at the two ends of the intrinsic model [19], [26], [27, ch. 3]. These are considered to be $200\,\Omega \cdot \mu\text{m}/W$ for the InGaAs HEMTs [19], [27, ch. 3] and $400\,\Omega \cdot \mu\text{m}/W$ for the GaN HEMTs [47].

The values of the parasitic elements are listed in Table II for both the InGaAs and GaN HEMTs; the values do not scale with gate length except for R_g, which is listed for the $L_g = 30$ nm case as a representative example. With these values, the model of Fig. 15 was utilized to generate the overall y-parameters and hence to find the extrinsic f_T and $f_{\max}$.

C. Results

We focus on the scaling behavior of the extrinsic f_T and $f_{\max}$, where for each gate length, the bias point was chosen to be that for peak f_T of the intrinsic device. The current gain $|h_{21}| = |y_{21}/y_{11}|$ [48] and the unilateral power gain U [49], calculated from the overall y-parameters, are extrapolated to obtain the extrinsic f_T and $f_{\max}$, respectively; sample plots of $|h_{21}|$ and U are provided in Fig. 16.

Fig. 17, which plots the extrinsic f_T versus gate length L_g, shows that the extrinsic f_T exhibits the same saturation at short gate lengths that was discussed for the intrinsic f_T. Moreover, as expected, the InGaAs HEMT exhibits a more pronounced saturation, given its lower effective mass. As a result of the rapid saturation in the f_T of the InGaAs HEMTs, the GaN HEMTs have comparable values of extrinsic f_T at short gate lengths. We have also shown experimental data for the f_T of InGaAs HEMTs in Fig. 17, as collected in [1, Fig. 10], and they exhibit the same trend as our simulations.

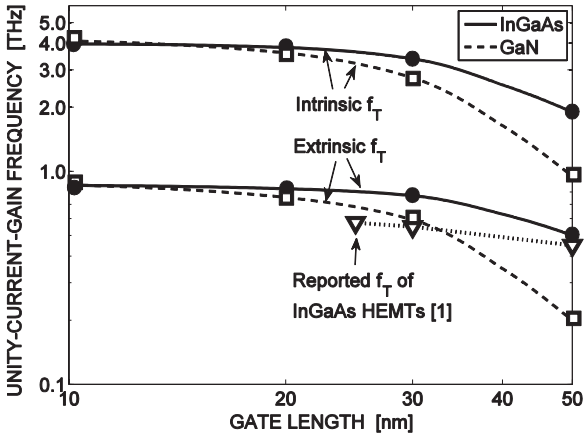


Fig. 17. Extrinsic and intrinsic f_T of the HEMTs considered in this paper, and the reported f_T of InGaAs HEMTs [1, Fig. 10], versus L_g. The lines have been drawn as guides for the eye.

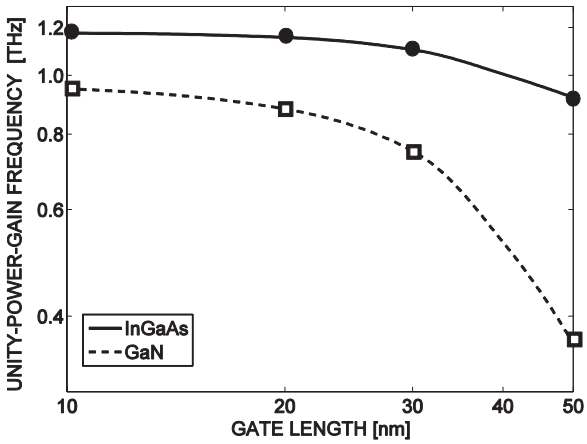


Fig. 18. Extrinsic $f_{\max}$ of the HEMTs versus L_g. The lines have been drawn as guides for the eye.

Regarding the $f_{\max}$, it is well known that, to a first approximation, $f_{\max} \propto \sqrt{f_T/(RC)_{\text{eff}}}$ [50]–[52] for RF transistors, where $(RC)_{\text{eff}}$ refers to an effective charging time. While the scaling behavior of the $f_{\max}$ based on such a relationship can be involved due to the involved nature of $(RC)_{\text{eff}}$, it is clear that the saturating behavior of the f_T at short gate lengths will also tend to saturate the $f_{\max}$, as depicted in Fig. 18. In addition, among the HEMTs under consideration, the GaN devices suffer from substantially larger parasitic resistances in the gate, source, and drain leads, as indicated by the higher values of the associated resistances in Table II, and this will tend to degrade the $f_{\max}$ through a larger value of $(RC)_{\text{eff}}$; thus, although the two HEMTs have comparable extrinsic f_T at short gate lengths, the GaN HEMTs have a lower $f_{\max}$ at all gate lengths, as shown in Fig. 18.

It should be mentioned that apart from the effective mass, other effects can contribute to the diminishing enhancement of the extrinsic f_T and $f_{\max}$ at short gate lengths. For example, while not an issue for the devices studied in this paper, parasitic resistances and capacitances that do not scale with gate length at the same rate as the elements of the intrinsic device can exacerbate the saturating behavior of the extrinsic figures of merit.

IV. Conclusion

The following conclusions can be drawn from this study of the impact of effective mass on the gate-length scaling behavior of the f_T and $f_{\max}$ of III–V HEMTs.

1) The intrinsic peak f_T occurs at gate voltages corresponding to the point of barrier collapse, beyond which the f_T degrades significantly.
2) At shorter gate lengths, DIBL causes the barrier to collapse at lower gate bias voltages, such that the intrinsic peak f_T occurs at lower gate voltages.
3) For a given channel material, with a downscaling of gate length, the transconductance g_m at the gate bias corresponding to peak f_T remains relatively insensitive to scaling. On the other hand, the low effective mass causes the intrinsic gate capacitance $C_{\rm gg}$ to roughly equal the inversion capacitance $C_{\rm inv}$, which scales slower than would be expected from a purely linear dependence on gate length; the slower scaling is an outcome of the peak f_T occurring at lower gate biases at shorter gate lengths (due to DIBL), which reduces the sharpness of the potential well defining the channel and thereby lowers the position of the conduction subbands with respect to the source Fermi level, leading to a larger than expected charge modulation and hence a larger than expected capacitance. The intrinsic peak $f_T = g_m/(2\pi C_{\rm gg})$ thus exhibits a saturating behavior when the gate length is scaled down, i.e., it shows no further increase with decreasing gate length once the gate length is sufficiently small; our results (see Figs. 11 and 17) show this to occur for gate lengths below approximately 30 nm.
4) In comparing channel materials, the material with lower effective mass will exhibit a more pronounced saturation in its peak f_T as the gate length is scaled down; this occurs because the subbands in the lighter mass material must move further below the source Fermi level to accommodate the required charge at the top of the barrier at peak f_T (as set by the doping in the n^+ region), and the lower positioning of the subbands accentuates the larger than expected gate capacitance as the gate length is scaled down.
5) The extrinsic peak f_T of HEMTs reflects the saturating behavior of the intrinsic f_T. In comparing HEMTs, the InGaAs HEMTs have a more pronounced saturation in their f_T due to their lower effective mass, such that the f_T of the two HEMTs are comparable at short gate lengths.
6) The saturating behavior of the f_T of III–V HEMTs at short gate lengths will contribute to the saturating behavior of the $f_{\max}$. In comparing HEMTs, the larger parasitic resistances of GaN HEMTs cause them to have a lower $f_{\max}$ than the InGaAs HEMTs at all gate lengths.

Overall, the most important outcome of this paper is the connection between the effective mass and the scaling behavior of RF performance. While a low effective mass is desirable for high mobility and potentially high-speed operation, it leads to diminishing improvements in the peak f_T and $f_{\max}$ as the gate length is scaled down.

References

[1] F. Schwierz and J. J. Liou, "RF transistors: Recent developments and roadmap toward terahertz applications," *Solid State Electron.*, vol. 51, no. 8, pp. 1079–1091, 2007.

[2] M. V. Fischetti and S. E. Laux, "Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures—Part II: Submicrometer MOSFET's," *IEEE Trans. Electron Devices*, vol. 38, no. 3, pp. 650–660, Mar. 1991.

[3] P. M. Solomon and S. E. Laux, "The ballistic FET: Design, capacitance and speed limit," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2001, pp. 95–98.

[4] A. Rahman, G. Klimeck, and M. Lundstrom, "Novel channel materials for ballistic nanoscale MOSFETs-bandstructure effects," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2005, pp. 601–604.

[5] K. C. Saraswat, C. O. Chui, D. Kim, T. Krishnamohan, and A. Pethe, "High mobility materials and novel device structures for high performance nanoscale MOSFETs," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2006, pp. 1–4.

[6] J. Ayubi-Moak, R. Akis, D. K. Ferry, S. M. Goodnick, N. Faralli, and M. Saraniti, "Towards the global modeling of InGaAs-based pseudomorphic HEMTs," *J. Comput. Electron.*, vol. 7, no. 3, pp. 187–191, 2008.

[7] R. Akis, J. Ayubi-Moak, D. K. Ferry, S. M. Goodnick, N. Faralli, and M. Saraniti, "Full-band cellular Monte Carlo simulations of terahertz high electron mobility transistors," *J. Phys.: Condens. Matter*, vol. 20, no. 38, pp. 384201-1–9, 2008.

[8] International Technology Roadmap for Semiconductors. (2009). "Radio frequency and analog/mixed-signal technologies for wireless communications," 2009 ed. [Online]. Available:http://www.itrs.net/Links/2009ITRS/2009Chapters_2009Tables/2009_Wireless.pdf

[9] Y. Yamashita, A. Endoh, K. Shinohara, K. Hikosaka, T. Matsui, S. Hiyamizu, and T. Mimura, "Pseudomorphic In0.52Al0.48As/In0.7Ga0.3As HEMTs with an ultrahigh fT of 562 GHz," *IEEE Electron Device Lett.*, vol. 23, no. 10, pp. 573–575, Oct. 2002.

[10] B. O. Lim, K. L. Mun, J. B. Tae, M. Han, C. K. Sung, and J. K. Rhee, "50-nm T-gate InAlAs/InGaAs metamorphic HEMTs with low noise and high fT characteristics," *IEEE Electron Device Lett.*, vol. 28, no. 7, pp. 546–548, Jul. 2007.

[11] W. R. Deal, "Solid-state amplifiers for terahertz electronics," in *Proc. IEEE Microwave Symp. Digest*, May 2010, pp. 1122–1125.

[12] T. Palacios, Y. Dora, A. Chakraborty, C. Sanabria, S. Keller, S. P. DenBaars, and U. K. Mishra, "Optimization of AlGaN/GaN HEMTs for high frequency operation," *Phys. Stat. Sol. (a)*, vol. 203, no. 7, pp. 1845–1850, 2006.

[13] A. Endoh, Y. Yamashita, K. Ikeda, M. Higashiwaki, K. Hikosaka, T. Matsui, S. Hiyamizu, and T. Mimura, "Fabrication of sub-50-nm-gate i-AlGaN/GaN HEMTs on sapphire," *Phys. Stat. Sol. (c)*, vol. 0, no. 7, pp. 2368–2371, 2003.

[14] D. Guerra, R. Akis, F. A. Marino, D. K. Ferry, S. M. Goodnick, and M. Saraniti, "Aspect ratio impact on RF and DC performance of state-of-the-art short-channel GaN and InGaAs HEMTs," *IEEE Electron Device Lett.*, vol. 31, no. 11, pp. 1217–1219, Nov. 2010.

[15] R. Tulek, A. Ilgaz, S. Gokden, A. Teke, M. K. Ozturk, M. Kasap, S. Ozgelik, E. Arslan, and E. Ozbay, "Comparison of the transport properties of high quality AlGaN/AlN/GaN and AlInN/AlN/GaN two-dimensional electron gas heterostructures," *J. Appl. Phys.*, vol. 105, no. 1, pp. 013707-1–013707-6, 2009.

[16] Y. Wang, L. Ma, Z. Yu, and L. Tian, "Optimization of two-dimensional electron gases and I-V characteristics for AlGaN/GaN HEMT devices," *Superlattices Microstruct.*, vol. 36, no. 4–6, pp. 869–875, 2004.

[17] D. K. Ferry, "The onset of quantization in ultra-submicron semiconductor devices," *Superlattices Microstruct.*, vol. 27, no. 2-3, pp. 61–66, 2002.
[18] R. Akis, L. Shifren, D. K. Ferry, and D. Vasileska, "The effective potential and its use in simulation," *Phys. Stat. Sol. (b)*, vol. 226, no. 1, pp. 1–8, 2001.
[19] N. Neophytou, T. Rakshit, and M. S. Lundstrom, "Performance analysis of 60-nm gate-length III-V InGaAs HEMTs: Simulations versus experiments," *IEEE Trans. Electron Devices*, vol. 56, no. 7, pp. 1377–1387, Jul. 2009.
[20] *COMSOL Multiphysics Version 3.5a*. COMSOL, Inc. Stockholm, Sweden, 2008.
[21] D. Kim and J. A. del Alamo, "Lateral and vertical scaling of In0.7Ga0.3As HEMTs for post-Si-CMOS logic applications," *IEEE Trans. Electron Devices*, vol. 55, no. 10, pp. 2546–2553, Oct. 2008.
[22] R. Venugopal, Z. Ren, S. Datta, M. S. Lundstrom, and D. Jovanovic, "Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches," *J. Appl. Phys.*, vol. 92, no. 7, pp. 3730–3739, 2002.
[23] S. Datta, "Nanoscale device modeling: the Green's function method," *Superlattices. Microstruct.*, vol. 28, no. 4, pp. 253–278, 2000.
[24] M. P. L. Sancho, J. M. L. Sancho, and J. Rubio, "Highly convergent schemes for the calculation of bulk and surface Green functions," *J. Phys. F: Met. Phys.*, vol. 15, no. 4, pp. 851–858, 1985.
[25] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal, "Two-dimensional quantum mechanical modeling of nanotransistors," *J. Appl. Phys.*, vol. 91, no. 4, pp. 2343–2354, 2002.
[26] Y. Liu and M. S. Lundstrom, "Simulation of III-V HEMTs for high-speed low-power logic applications," in *ECS Trans.*, 2009, vol. 19, pp. 331–342.
[27] S. Oktyabrsky and P. D. Ye, Eds., Fundamentals of III-V Semiconductor MOSFETs, 1st ed.. New York: Springer, 2010.
[28] S. P. Kumar, A. Agrawal, R. Chaujar, S. Kabra, M. Gupta, and R. S. Gupta, "3-Dimensional analytical modeling and simulation of fully depleted AlGaN/GaN modulation doped field effect transistor," in *Proc. Int. Workshop Phys. Semicond. Devices*, Dec. 2007, pp. 373–376.
[29] G. Steinhoff, B. Baur, G. Wrobel, S. Ingebrandt, A. Offenhausser, A. Dadgar, A. Krost, M. Stutzmann, and M. Eickhoff, "Recording of cell action potentials with AlGaN/GaN field-effect transistors," *Appl. Phys. Lett.*, vol. 86, no. 3, pp. 33901-1–33901-3, 2005.
[30] J. A. del Alamo, "Is nanometer-scale III–V CMOS cool enough to rejuvenate Moore's Law?" *Compound Semicond.*, vol. 16, no. 5, pp. 19–22, Jul. 2010.
[31] Y. Tsividis, *Operation and Modeling of the MOS Transistor*, 2nd ed. New York: McGraw-Hill, 1999.
[32] D. Jin, D. Kim, T. Kim, and J. A. del Alamo, "Quantum capacitance in scaled down III-V FETs," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2009, pp. 1–4.
[33] A. Rahman, J. Guo, S. Datta, and M. S. Lundstrom, "Theory of ballistic nanotransistors," *IEEE Trans. Electron Devices*, vol. 50, no. 9, pp. 1853–1864, Sep. 2003.
[34] S. Hasan, S. Salahuddin, M. Vaidyanathan, and M. A. Alam, "High-frequency performance projections for ballistic carbon-nanotube transistors," *IEEE Trans. Nanotechnol.*, vol. 5, no. 1, pp. 14–22, Jan. 2006.
[35] S. Datta, *Quantum Transport: Atom to Transistor*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[36] D. Kim and J. A. del Alamo, "Scalability of sub-100 nm InAs HEMTs on InP substrate for future logic applications," *IEEE Trans. Electron Devices*, vol. 57, no. 7, pp. 1504–1511, Jul. 2010.
[37] B. V. V. Zeghbroeck, *Principles of Semiconductor Devices and Heterojunctions*, 1st ed. Upper Saddle River, NJ: Prentice-Hall, 2009.
[38] A. S. Barker, Jr. and M. Ilegems, "Infrared lattice vibrations and free-electron dispersion in GaN," *Phys. Rev. B*, vol. 7, no. 2, pp. 743–750, 1973.
[39] B. K. Meyer, D. Volm, A. Graber, H. C. Alt, T. Detchprohm, K. Amano, and I. Akasaki, "Shallow donors in GaN - The binding energy and the electron effective mass," *Solid State Commun.*, vol. 95, no. 9, pp. 597–600, 1995.
[40] D. L. John, F. Allerstam, T. Rodle, S. K. Murad, and G. D. J. Smit, "A surface-potential based model for GaN HEMTs in RF power amplifier applications," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2010, pp. 8.3.1–8.3.4.
[41] J. M. Lopez, T. Gonzalez, D. Pardo, S. Bollaert, T. Parenty, and A. Cappy, "Design optimization of AlInAs-GaInAs HEMTs for high-frequency applications," *IEEE Trans. Electron Devices*, vol. 51, no. 4, pp. 521–528, Apr. 2004.
[42] W. D. Hu, X. S. Chen, Z. J. Quan, X. M. Zhang, Y. Huang, C. S. Xia, W. Lu, and P. D. Ye, "Simulation and optimization of GaN-based metal-oxide-semiconductor high-electron-mobility-transistor using field-dependent drift velocity model," *J. Appl. Phys.*, vol. 102, no. 3, pp. 034502-1–034502-7, 2007.
[43] N. Paydavosi, J. P. Rebstock, K. D. Holland, S. Ahmed, A. U. Alam, and M. Vaidyanathan, "RF performance potential of array-based carbon-nanotube transistors.—Part II: Extrinsic results," *IEEE Trans. Electron Devices*, vol. 58, no. 7, pp. 1941–1951, Jul. 2011.
[44] Y. Q. Wu, W. K. Wang, O. Koybasi, D. N. Zakharov, E. A. Stach, S. Nakahara, J. C. M. Hwang, and P. D. Ye, "0.8-V supply voltage deep-submicrometer inversion-mode In0.75 Ga0.25As MOSFET," *IEEE Electron Device Lett.*, vol. 30, no. 7, pp. 700–702, Jul. 2009.
[45] Y. Yamashita, A. Endoh, K. Ikeda, K. Hikosaka, T. Mimura, M. Higashiwaki, T. Matsui, and S. Hiyamizu, "Effect of thermal annealing on 120-nm-T-shaped-TiPtAu-gate AlGaNGaN high electron mobility transistors," *J. Vac. Sci. Technol. B*, vol. 23, no. 3, pp. 895–899, 2005.
[46] C. Kocabas, S. Dunham, Q. Cao, K. Cimino, X. Ho, H. Kim, D. Dawson, J. Payne, M. Stuenkel, H. Zhang, T. Banks, M. Feng, S. V. Rotkin, and J. A. Rogers, "High-frequency performance of submicrometer transistors that use aligned arrays of single-walled carbon nanotubes," *Nano Lett.*, vol. 9, no. 5, pp. 1937–1943, 2009.
[47] H. Yu, L. McCarthy, S. Rajan, S. Keller, S. Denbaars, J. Speck, and U. Mishra, "Ion implanted AlGaN-GaN HEMTs with nonalloyed ohmic contacts," *IEEE Electron Device Lett.*, vol. 26, no. 5, pp. 283–285, May 2005.
[48] N. Paydavosi, K. D. Holland, M. M. Zargham, and M. Vaidyanathan, "Understanding the frequency- and time-dependent behavior of ballistic carbon-nanotube transistors," *IEEE Trans. Nanotechnol.*, vol. 8, no. 2, pp. 234–244, Mar. 2009.
[49] M. S. Gupta, "Power gain in feedback amplifiers, a classic revisited," *IEEE Trans. Microwave Theory Tech.*, vol. 40, no. 5, pp. 864–879, May 1992.
[50] L. C. Castro and D. L. Pulfrey, "Extrapolated fmax for carbon nanotube field-effect transistors," *Nanotechnology*, vol. 17, no. 1, pp. 300–304, 2006.
[51] M. Vaidyanathan and D. L. Pulfrey, "Extrapolated f_{max} of heterojunction bipolar transistors," *IEEE Trans. Electron Devices*, vol. 46, no. 2, pp. 301–309, Feb. 1999.
[52] C. L. Tao and G. A. Armstrong, "The impact of the intrinsic and extrinsic resistances of double gate SOI on RF performance," *Solid State Electron.*, vol. 50, no. 5, pp. 774–783, 2006.
[53] S. Luryi, "Quantum capacitance devices," *Appl. Phys. Lett.*, vol. 52, no. 6, pp. 501–503, 1988.
[54] N. Paydavosi, M. M. Zargham, K. D. Holland, C. M. Dublanko, and M. Vaidyanathan, "Non-quasi-static effects and the role of kinetic inductance in ballistic carbon-nanotube transistors," *IEEE Trans. Nanotechnol.*, vol. 9, no. 4, pp. 449–463, Jul. 2010.

Sabbir Ahmed received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 2005 and 2007, respectively. Since 2008, he has been working toward the Ph.D. degree at the University of Alberta, Edmonton, AB, Canada.

He was a Lecturer in the Department of Electrical and Electronic Engineering, BUET, from 2005 to 2008. His research interests include theory, modeling, and simulation of nanoscale electronic devices, with an emphasis on the high-frequency and circuit-level performance of III–V high-electron-mobility transistors, carbon-based transistors, and solar cell devices.

Mr. Ahmed received the F. S. Chia Doctoral Scholarship in 2008 and 2009, and the Queen Elizabeth II Graduate Scholarship in 2010, 2011, and 2012 at the University of Alberta.

Kyle David Holland received the B.Sc. degree in engineering physics (nanoengineering option) from the University of Alberta, Edmonton, AB, Canada, in 2009, where he is currently working toward the Ph.D. degree in electrical engineering.

His research interests include quantum simulation of carbon-based nanoelectronics, with an emphasis on modeling the high-frequency performance of graphene devices.

Mr. Holland currently holds the Natural Sciences and Engineering Research Council of Canada Alexander Graham Bell Canada Graduate Scholarship and the Alberta Innovates Graduate Student Scholarship. He has also received the Ralph Steinhauer Award of Distinction.

Navid Paydavosi received the B.A.Sc. degree in electrical engineering from Shahid Beheshti University, Tehran, Iran, in 2005, and the Ph.D. degree in electrical engineering from the University of Alberta, Edmonton, AB, Canada, in 2011.

He is currently a Postdoctoral Scholar at the BSIM Group, University of California, Berkeley. His research interests include theory and modeling of future alternatives to ordinary silicon transistors, including carbon-based and III–V high-electron-mobility devices, with an emphasis on the high-frequency characteristics relevant for RF applications, such as the extrinsic cutoff frequency, the attainable power gain, the unity-power-gain frequency, and linearity.

Dr. Paydavosi received the Queen Elizabeth II Graduate Scholarship for September 2009 and January 2010, and two Tuition Supplement Awards in September 2006 and April 2007 from the University of Alberta.

Christopher Martin Sinclair Rogers received the B.Sc. degree in engineering physics (nanoengineering option) from the University of Alberta, Edmonton, AB, Canada, in 2012. In fall 2012, he will start working toward the Ph.D. degree in electrical engineering at Stanford University, Stanford, CA.

His current research interests include physics and modeling of nanoscale electronic devices, with an emphasis on radio-frequency characteristics and 2-D materials.

Mr. Rogers received the Natural Sciences and Engineering Research Council of Canada Undergraduate Student Research Award at the University of Alberta during the summers of 2011 and 2012. During his B.Sc. degree, he received the President's Citation from the University of Alberta. He has also received the Governor General's Silver Medal and The Rt. Hon. C.D. Howe Memorial Fellowship.

Ahsan Ul Alam received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 2005 and 2007, respectively. He is currently working toward the Ph.D. degree at the University of Alberta, Edmonton, AB, Canada.

From 2005 to 2008, he was a Lecturer at BUET, teaching courses and performing research in the area of semiconductor devices. His research interests include theory and modeling of semiconductor devices, where he has worked on topics ranging from the study of spin transport in carbon-based devices to the modeling of high-frequency distortion for wireless applications.

Mr. Alam received the Kintar-ul Huq Lashkar Gold Medal from BUET for securing the top position in his undergraduate studies. He also received the F. S. Chia Ph.D. Scholarship in the first two years and the Queen Elizabeth II Graduate Scholarship in the third, fourth, and fifth years of his Ph.D. studies at the University of Alberta.

Neophytos Neophytou received the Ph.D. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, in 2008.

He is currently a Postdoctoral Researcher at the Institute of Microelectronics, Technical University of Vienna, Vienna, Austria. His area of specialization is the theory, computational modeling, and simulation of transport in nanoelectronic devices. His current research interests include thermoelectric transport in nanostructured devices for applications in energy conversion and generation.

Diego Kienle received the B.S. (Vordiplom) and M.S. (Diplom) degrees from the University of Bayreuth, Bayreuth, Germany, and the Ph.D. (Dr.rer.nat.) degree from the Research Center Jülich and the University of Saarland, Saarland, Germany, all in theoretical physics.

After postdoctoral appointments with the Electrical and Computer Engineering Department, Purdue University, West Lafayette, IN, and the Material Science Department, Sandia National Laboratories, Livermore, CA, he is currently at the Institute of Theoretical Physics, University of Bayreuth. His research interests include formal theory, modeling, and simulation of ac quantum electronic transport in nanoscale materials and devices with a focus on the understanding of the quantum dynamic processes in low-dimensional materials and their potential application in solid-state terahertz devices. His past research interests include theory and modeling of complex fluids by means of Brownian dynamics with a focus on many-body hydrodynamic interaction effects in diluted polymer solutions.

Mani Vaidyanathan (M'99) received the Ph.D. degree in electrical engineering from the University of British Columbia, Vancouver, BC, Canada, in 1999.

He is currently an Associate Professor in the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. His research interests include modeling, simulation, and understanding of electronic devices for future technologies.

Dr. Vaidyanathan received the University of Alberta's Provost's Award and the University of Alberta's Alexander Rutherford Award, both for excellence in teaching.