f_T = 688 GHz and f_{max} = 800 GHz in L_g = 40 nm $In_{0.7}Ga_{0.3}As$ MHEMTs with g_{m_max} > 2.7 mS/μm

Dae-Hyun Kim, Berinder Brar and *Jesús A. del Alamo

Teledyne Scientific Company (TSC), Thousand Oaks, CA 91360, USA, E-mail: vtsrc3@gmail.com

*Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, U.S.A

Abstract

We have demonstrated 40-nm $In_{0.7}Ga_{0.3}As$ Metamorphic HEMTs (MHEMTs) with a record value in f_T. The devices feature a Pt gate sinking process to effectively thin down the $In_{0.52}Al_{0.48}As$ barrier layer, together with dual Si d-doping in the barrier to lower the potential barrier in the S/D access region. The fabricated device with L_g = 40-nm exhibits V_T = 0.05 V, $g_{m,max}$ = 2.7 mS/μm, f_T = 688 GHz and f_{max} = 800 GHz. In addition, we have developed an analytical model of f_T in a III-V HEMT based on a small-signal equivalent circuit, which provides an excellent agreement with measured f_T. This in turns guides a realistic way to further improve f_T beyond THz.

Introduction

The last several years have witnessed an explosion of interest in devices suitable for ultra high frequency applications, such as in the Tera-Hz regime. In particular, InGaAs-based high-electron mobility transistors (HEMTs) and heterojunction bipolar transistors (HBTs) have shown great success in microwave and millimeter-wave applications, and are now considered the best candidates for future THz electronics [1-4]. These remarkable results have been achieved through the combination of downscaling of minimum feature size, parasitics reduction, and the use of a channel material with very high electron velocity.

In this paper, we report on E-mode L_g = 40 nm $In_{0.7}Ga_{0.3}As$ metamorphic-HEMTs (MHEMTs) on GaAs substrate with a record f_T = 688 GHz, which to the knowledge of the authors, is the highest ever reported in any FET on any material system. In addition, we have constructed a simple and analytical model for f_T based on a small-signal equivalent circuit and shown that very high intrinsic transconductance (g_{mi}) in excess of 4 mS/μm in our L_g = 40 nm device is effective in mitigating a parasitic charging delay associated with the extrinsic gate capacitances. This greatly contributes to the record value of f_T in our L_g = 40 nm MHEMTs.

Process Technology

Fig. 1 shows cross section and TEM images of the fabricated InGaAs MHEMTs on GaAs substrate. From top to bottom, the epitaxial layer structure consists of a heavily doped multi-layer cap ($In_{0.7}Ga_{0.3}As/In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$), 6-nm InP etch-stopper, 2-nm $In_{0.52}Al_{0.48}As$ barrier, upper Si δ-doping, 6-nm $In_{0.52}Al_{0.48}As$ barrier, lower Si δ-doping, 2-nm $In_{0.52}Al_{0.48}As$ spacer, 1-nm $In_{0.7}Al_{0.3}As$ spacer, 10-nm $In_{0.7}Ga_{0.3}As$ channel, 300-nm $In_{0.52}Al_{0.48}As$ buffer and 0.3 μm graded metamorphic buffer on GaAs substrate. A dual Si δ-doping and an $In_{0.7}Al_{0.3}As$ spacer were utilized to lower the potential barrier in the access regions and reduce the parasitic resistance. The device fabrication is almost identical to [1]. After a two-step recess process that exposes an InAlAs barrier, a Pt/Ti/Pt/Au gate was created. Subsequently, the devices were annealed at 250 °C for 2 minutes to drive the Pt into the InAlAs barrier. In this way, a gate-to-channel distance (t_{ins}) of about 4-nm was achieved. The gate stem height was increased to 250 nm to mitigate the fringing capacitance of the T-gate. From the TEM images, the physical L_g, and side-recess-spacing (L_{side}) were 40 nm and 100 nm, respectively.

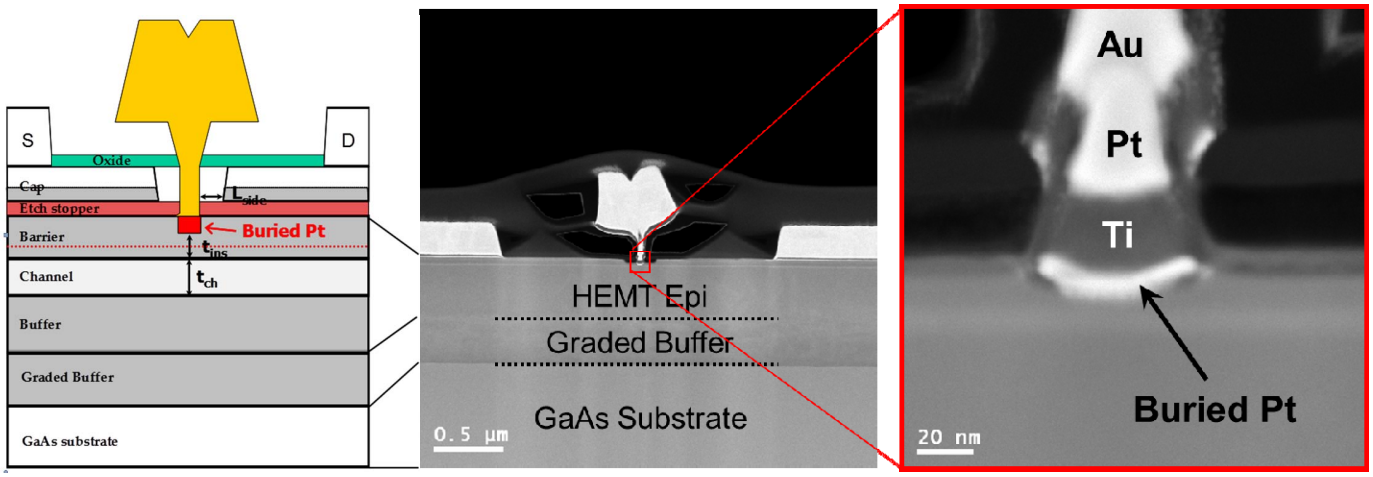

Fig. 1 Schematic of InGaAs MHEMTs and TEM images of the fabricated L_g = 40 nm device. It features 0.3 μm graded metamorphic buffer on GaAs substrate. The device brings unique aspects to mitigate parasitics and short-channel effects. The physical L_g and side-recess spacing (L_{side}) were 40 nm and 100 nm, respectively.

DC & Microwave Characteristics

Figs. 2 and **3** show output and transconductance (g_m) characteristics of an L_g = 40 nm InGaAs MHEMT. The device exhibits excellent pinch-off and drain current saturation behavior up to V_{DS} = 0.8 V. The device is enhancement-mode with V_T = 0.05 V. A very small value of R_{ON} = 280 Ω-μm is obtained, mainly due to the dual Si δ-doping and the InAs-rich $In_{0.7}Al_{0.3}As$ spacer. As a consequence, the device has outstanding g_{m_max} = 2.75 mS/μm at V_{DS} = 0.8 V, and in excess of 2 mS/μm even at V_{DS} = 0.3 V, both of which make our device technology attractive for high-performance and very low-power applications.

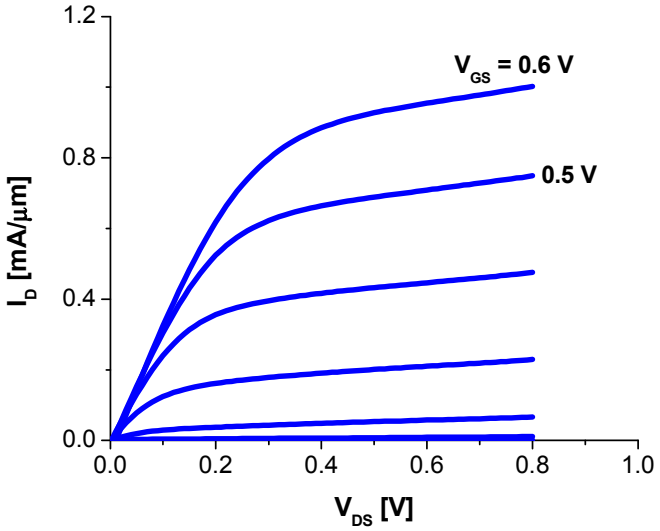

Fig. 2 Output characteristics of L_g = 40 nm $In_{0.7}Ga_{0.3}As$ MHEMTs on GaAs substrate.

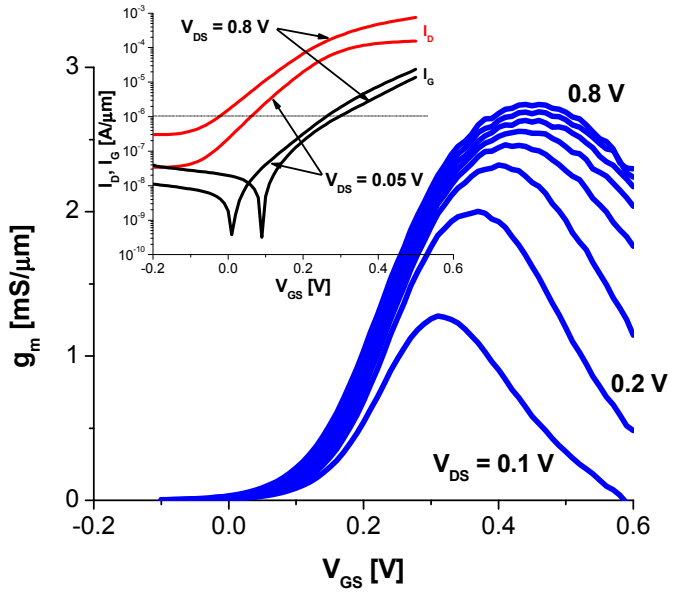

Fig. 3 transconductance (g_m) characersitics of L_g = 40 nm devices. Inset is subthreshold and gate leakage characteristics.

Microwave performance was characterized using two different network analyzers with an off-wafer standard LRRM calibration: 1) 1 - 50 GHz by HP-8510C, and 2) 1 - 67 GHz by Agilent-PNA. On-wafer open and short patterns were used to subtract pad capacitances and inductances from the measured device S parameters. Using the de-embedded S-parameters, we constructed a small-signal model. **Fig. 4** plots $|h_{21}|^2$, MSG, U_g and stability factor (k) against frequency of the 40 nm MHEMT with $W_g = 2 \times 20$ μm at V_{DS} = 0.6 V and I_D = 0.4 mA/μm. In this particular measurement, a value of f_T = 688 GHz was obtained by extrapolating $|h_{21}|^2$ with a slope of -20 dB/decade from both measurement systems. The value of f_T in our device was also verified through Gummel's approach [5] (inset) and small-signal modeling, yielding f_T = 690 and 680 GHz, respectively. To the knowledge of the authors, this is the highest f_T ever reported in any FET on any material system. Regarding f_{max}, it is hard to directly extract it from experimental measurements of U_g, especially as V_{DS} increases. Instead, we estimated it from the small-signal model, as in [1]. The modeled U_g nicely predicts the measured one, yielding f_{max} = 800 GHz. Our small-signal model reveals a very high value of g_{mi} = 4.4 mS/□m, and excellent combinations of g_{mi}/g_{oi} = 5 and C_{gs}/C_{gd} = 10. **Fig. 5** shows f_{max} against f_T for our InGaAs MHEMT, together with recent reports on HEMTs and HBTs in the literature. Our 40 nm device constitutes the first demonstration of both f_T and f_{max} above 680 GHz at the same bias point in any technology.

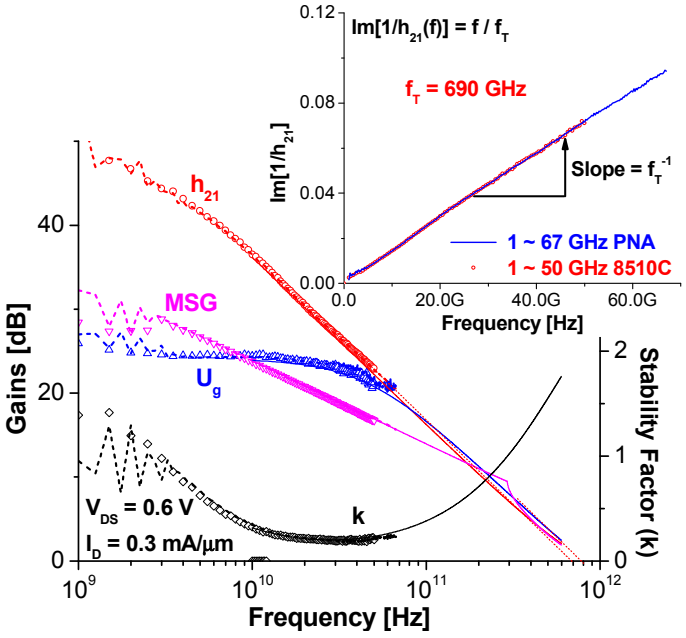

Fig. 4 RF gains and k vs. frequency of L_g = 40 nm MHEMTs: 1) 1-50 GHz using 8510C (symbols); 2) 1-67 GHz using PNA (dashed lines); 3) 10-600 GHz from the small-signal model (lines). Inset is f_T extraction by Gummel's method.

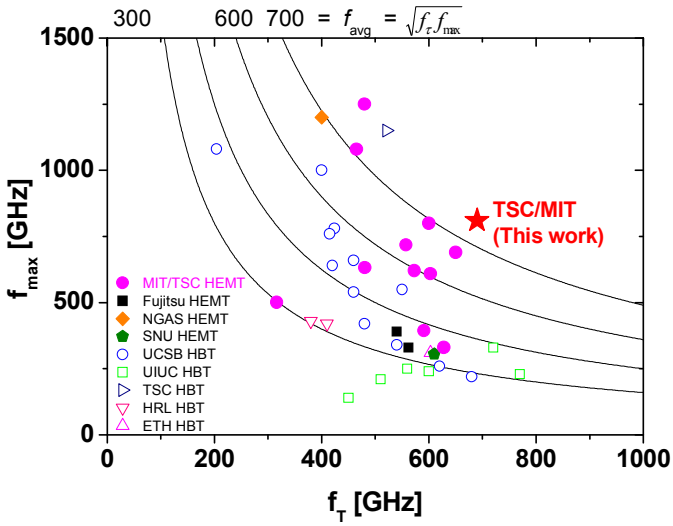

Fig. 5 f_{max} against f_T for our L_g = 40 nm MHEMTs, together with recent reports on HEMTs and HBTs in the literature.

Analytical f_T Model

To understand the origin of the record high-frequency response in our devices, we have developed an analytical model for f_T based on the small-signal equivalent circuit. **Fig. 6** shows a generic schematic of a III-V HEMT which can be partitioned in intrinsic parameters g_{mi}, g_{oi}, C_{gsi} and C_{gdi} and extrinsic parameters C_{gs_ext}, C_{gd_ext}, R_s and R_d.

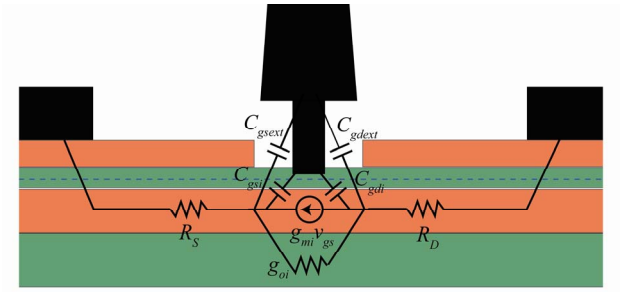

Fig. 6 Generic schematic of a III-V HEMT.

By putting these into the equation of f_T, we can partition the total delay ($t = 1/2\pi f_T$) into three components, as following:

$$f_T = \frac{1}{2\pi}\frac{g_{mi}}{C_{gs}+C_{gd}+g_{mi}C_{gd}(R_S+R_D)[1+(1+\frac{C_{gs}}{C_{gd}})\frac{g_{oi}}{g_{mi}}]}$$

$$C_{gs} = C_{gsi} + C_{gsext} \qquad C_{gd} = C_{gdi} + C_{gdext}$$

$$\tau = \frac{1}{2\pi f_T} = \tau_t + \tau_{ext} + \tau_{par}$$

$$\tau_t = \frac{C_{gsi}+C_{gdi}}{g_{mi}} = \frac{L_g}{v_e}$$

$$\tau_{ext} = \frac{C_{gsext}+C_{gdext}}{g_{mi}}$$

$$\tau_{par} = C_{gd}(R_S+R_D)[1+(1+\frac{C_{gs}}{C_{gd}})\frac{g_{oi}}{g_{mi}}]$$

τ_t is the transit time under the gate which in a simple model is given as L_g/v_e, where v_e is the average electron velocity under the gate. τ_{ext} is the parasitic charging delay through C_{gs_ext} and C_{gd_ext}. Of course, the lower the extrinsic capacitances, the smaller τ_{ext}. Alternatively, improving carrier transport by increasing g_{mi} is also an effective way to mitigate τ_{ext}. τ_{par} is the delay mainly associated with R_s and R_d. **Fig. 7** illustrates the extracted C_{gs} and C_{gd} against L_g at V_{DS} = 0.6 V. Both C_{gs} and C_{gd} are linearly dependent upon L_g, as expected. Note that C_{gs_ext} is larger than C_{gd_ext} due to a higher value of reverse bias at the drain side of the gate. **Fig. 8** plots the extracted g_{mi} and g_{oi} against L_g at V_{DS} = 0.6 V. As L_g decreases, g_{mi} initially increases and then saturates at L_g of around 60 nm, whereas g_{oi} continues to increase.

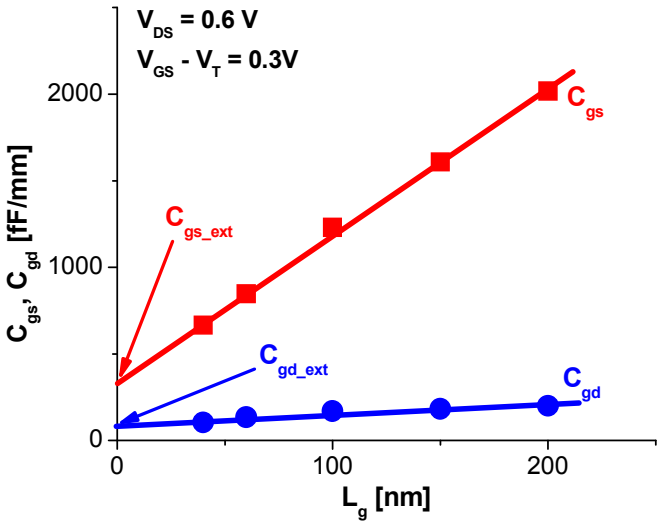

Fig. 7 C_{gs} and C_{gd} as a function of L_g at V_{DS} = 0.6 V. Both C_{gs} and C_{gd} are linearly dependent upon L_g.

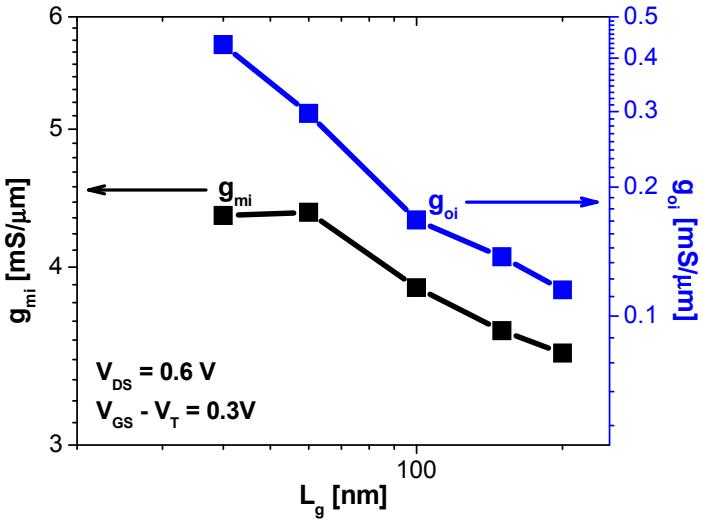

Fig. 8 g_{mi} and g_{oi} as a function of L_g at ($V_{GS} - V_T$) = 0.3 V and V_{DS} = 0.6 V.

Given the values of g_{mi}, g_{oi}, C_{gsi}, C_{gdi}, C_{gs_ext} and C_{gd_ext}, together with parasitic resistances (R_s and R_d), the delay components were computed, as above. **Fig. 9** plots extracted three delay components as a function of L_g. As L_g decreases, τ_t decreases accordingly. However, both τ_{ext} and τ_{par} do not scale with L_g and remain constant. In fact, at L_g = 40 nm, τ_{ext} becomes a dominant factor in the total delay. As a matter of fact, a sum of τ_{ext} and τ_{par} constitutes 65% in the L_g = 40 nm device, which significantly hampers further L_g scaling benefits in f_T. **Fig. 10** plots the electron average velocity (v_e) under the gate, as a function of L_g. As L_g decreases, the extracted v_e increases thanks to an increase in the electron injection velocity at the top of the potential barrier near the source and enhanced velocity overshoot effect at the drain side of the gate. At L_g = 40 nm, v_e = 5 × 10^7 cm/s was obtained from our analytical f_T model.

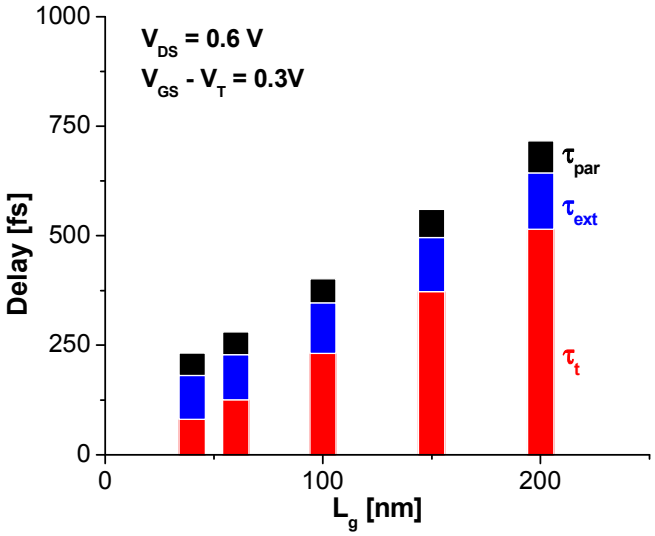

Fig. 9 Modeled delay components as a function of L_g.

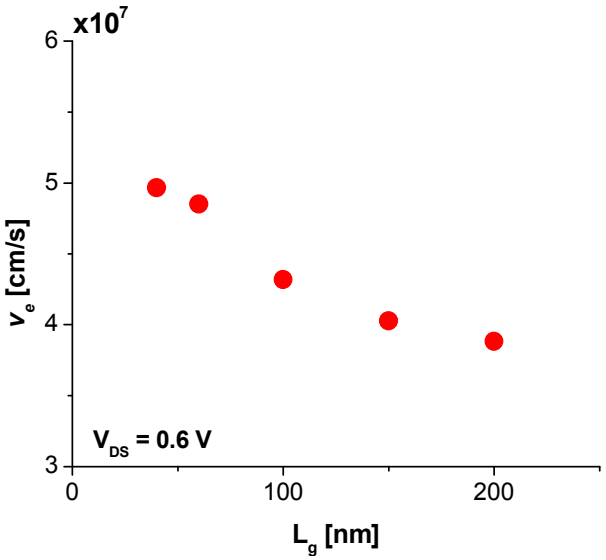

Fig. 10 Extracted average electron velocity (v_e) against L_g.

Finally, we have verified our analytical f_T model, by comparing it to the measured one. **Fig. 11** plots both measured and modeled f_T as a function of L_g at V_{DS} = 0.6 V. Excellent agreement is achieved which increases the credibility of our extraction process. The model helps us to discuss technological options to further improve f_T beyond 1 THz. For that, we added a model projection of f_T in **Fig. 11** with 30% reduction in both parasitic resistances R_s and R_d, and extrinsic capacitances C_{gs_ext} and C_{gd_ext} at V_{DS} = 0.6 V, together with L_g scaling down to 20 nm. This reveals that 1 THz f_T is an eminently attainable goal. This study highlights the importance of parasitic reduction to further boost f_T, rather than straight L_g scaling.

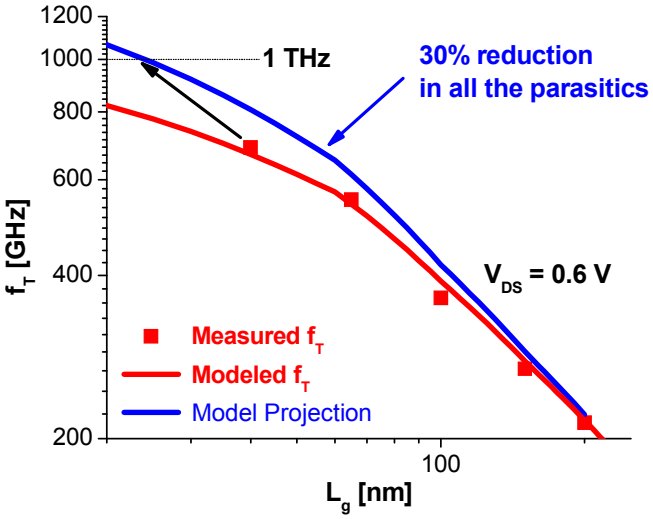

Fig. 11 Modeled (Red-line) and measured f_T (Rectangle) as a function of L_g at V_{DS} = 0.6 V, together with model projection (Blue-line) having 30 % reduction in all the parasitic components, such as R_s, R_d, C_{gs_ext} and C_{gd_ext}.

Conclusions

In summary, we have demonstrated a record f_T = 688 GHz in L_g = 40 nm $In_{0.7}Ga_{0.3}As$ MHEMTs on GaAs substrate with g_{m_max} in excess of 2.7 mS/μm. To the best of the author's knowledge, this is the highest f_T ever report in any FET on any material system. Besides, we have developed an analytical f_T model, which provides an excellent agreement with measured f_T. This in turns guides a realistic way to further improve f_T beyond THz.

References

[1] D.-H. Kim *et al.*, *IEEE EDL*, p. 830, 2008.
[2] R. Lai *et al.*, *IEDM*, p. 609, 2007.
[3] M. Urteaga *et al.*, *DRC*, pp. 281, 2011.
[4] A. Leuther *et al.*, *IPRM*, 2011.
[5] H. K. Gummel, *Pro. IEEE*, pp. 2159, 1969.

Acknowledgements

This work has been supported by the internal R&D program at TSC.