Nano-scale Type-II InP/GaAsSb DHBTs to Reach THz Cutoff Frequencies

William Snodgrass and Milton Feng

University of Illinois at Urbana-Champaign, 208 North Wright Street, Urbana, IL 61801
mfeng@uiuc.edu (217) 333-8080

Keywords: Indium Phosphide, Heterojunction Bipolar Transistor, Terahertz, Scaling

Abstract: We demonstrate vertically and laterally scaled GaAsSb/InP type-II DHBTs with f_T = 670 GHz at 10.3 mA/μm^2 emitter current density and off-state collector-emitter breakdown voltage BV_{CEO} = 3.2 V. Small-signal modeling is used to extract delay terms and to identify material design and device fabrication requirements for next-generation devices with > 1 THz cutoff frequencies.

Introduction

InP based heterojunction bipolar transistors (HBTs) have the highest current gain cutoff frequencies of any transistor technology but the limits of the material for this application have not yet been reached. The reported breakdown voltages versus current gain cutoff frequencies of modern SiGe HBT, InP pHEMT, and InP HBT transistors are plotted in Fig. 1, revealing scaling trends for each technology. As illustrated in Fig. 1, InP HBTs maintain a breakdown advantage over SiGe HBTs and InP HEMT devices as the structures of each are scaled to achieve higher cutoff frequencies at the expense of breakdown voltage. Double heterojunction HBT (DHBT) designs using wide-gap InP collector layers improve the breakdown voltage and power handling of vertically scaled devices. Type-I DHBTs using InGaAs base layers have a current-blocking conduction band discontinuity at the base-collector junction that is alleviated through the use of setback and superlattice or step-graded layers in the collector. These devices benefit from ballistic electron injection and high electron mobility through the InGaAs base. Because type-I DHBTs must incorporate narrow band gap material at the base-collector interface they possess a limit to vertical scaling that results in breakdown voltages approaching those of the SHBT for vertically scaled designs with f_T greater than 500 GHz. The type-II InP DHBT with GaAsSb as the base layer is an alternative having a base-collector junction energy band alignment that naturally favors electron collection—allowing for an all-InP collector layer to be used to achieve higher breakdown voltage and thermal dissipation in scaled material structures. Fig. 2 plots reported transistor breakdown voltage versus collector layer thickness of several high performance InP SHBT, type-I DHBT, and type-II DHBT devices where this trend and the potential type-II advantage becomes apparent below 1000 Å.

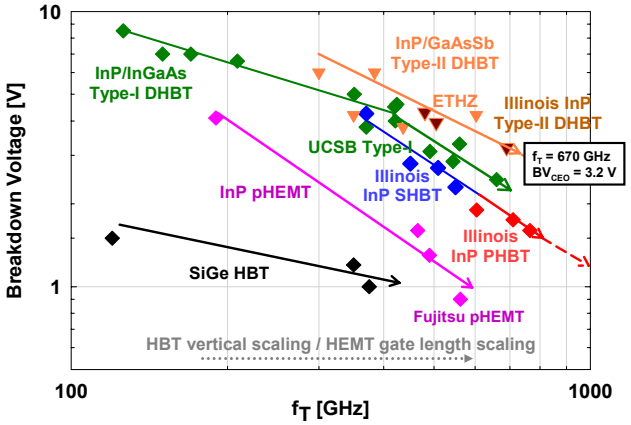

Figure 1. Breakdown voltage (BV_{CEO} at 1 kA/cm^2) versus current gain cutoff frequency f_T of InP HBT, InP pHEMT, and SiGe HBT devices.

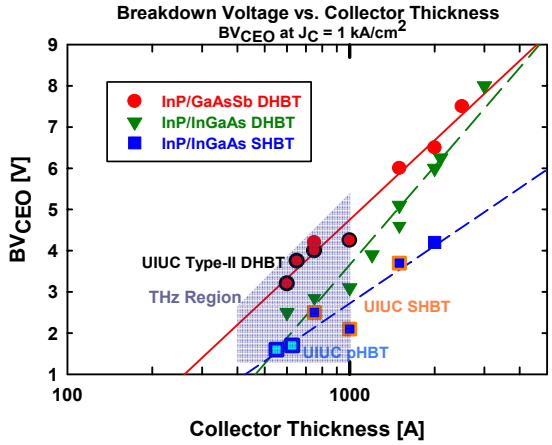

Figure 2. Reported breakdown voltage (BV_{CEO} at 1 kA/cm^2) versus collector layer thickness for InP SHBTs, type-I DHBTs, and type-II DHBTs. After [1].

Design and Fabrication

Low electron mobility in the type-II GaAsSb base layer typically limits the f_T of these devices, despite electron velocities exceeding 4.5 x 10^7 in the InP collector [2]. To enhance electron transport and lower total transit delay, we incorporate an InGaAsSb compositional grading—creating a built-in potential difference across the base layer of approximately 55 meV. Device structures were grown at the University of Illinois using a gas-source molecular beam epitaxy system. The epitaxial structures of the DHBT consisted of a 350 nm InP subcollector doped to $n = 3 \times 10^{19}$ cm^{-3}, a 25 nm InGaAs subcollector doped to $n = 4 \times 10^{19}$ cm^{-3}, an InP collector doped to $n = 3 \times 10^{16}$ cm^{-3}, a carbon-doped GaAsSb-InGaAsSb graded base with $p = 8 \times 10^{19}$ cm^{-3}, a 40 nm Si-doped InP emitter, and a 40 nm graded InGaAs emitter cap. Electron beam lithography and wet chemical etching were used to fabricate devices with emitter widths as small as 250 nm and lengths ranging from

2-8 µm. Extrinsic parasitics are minimized by depositing a silicon nitride emitter-base junction sidewall spacer after base contacts to self-align the base-collector etch. The extrinsic base contact post is electrically isolated from the intrinsic device. Devices are planarized by BCB spin coating prior to RF probe pad metallization. An energy band diagram with material composition and doping is presented in Fig. 3. A 20 nm base, 60 nm collector structure was designed to minimize transit delay and achieve high f_T. A 30 nm base, 100 nm collector structure was also grown to reduce parasitic resistance and capacitance for balanced f_T and f_{MAX}.

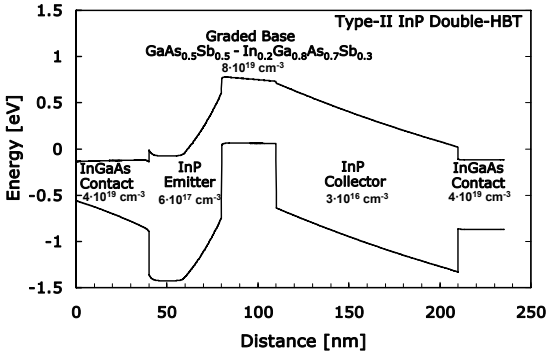

Figure 3. Illustration of UIUC graded-base type-II DHBT including material composition and doping of each region.

MEASUREMENT RESULTS

Transistor S-parameters were measured from 0.5 to 50 GHz using an HP8510C vector network analyzer. Off-wafer SOLT calibrations were used and measured on-wafer open and short probe pad parasitics were subsequently de-embedded from the measurements. A 60 nm collector device with emitter area 0.52 × 7.6 µm^2 has extrapolated f_T = 670 GHz and simultaneous f_{MAX} = 185 GHz when operated at current density J_E = 10.3 mA/µm^2. Measured current and unilateral power gain versus frequency of this device are plotted in Fig. 4. DC common-emitter and Gummel I-V characteristics of the device are presented in Fig. 5 and exhibit low offset voltage, small knee voltage

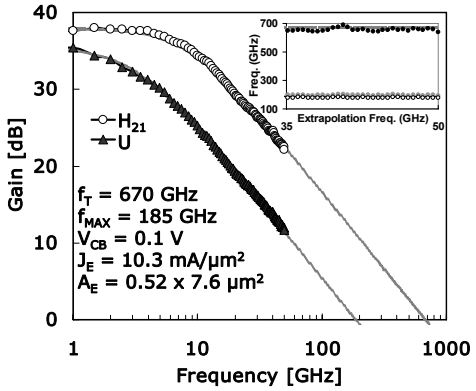

Figure 4. Measured 0-50 GHz RF current and power gains with extrapolations to f_T = 670 GHz and f_{MAX} = 185 GHz of device with A_E = 0.52 x 7.6 µm^2 on 20nm base, 60nm collector structure biased at 10.3 mA/µm^2. Extrapolated f_T and f_{MAX} assuming -20 dB/decade versus extrapolation frequency (inset).

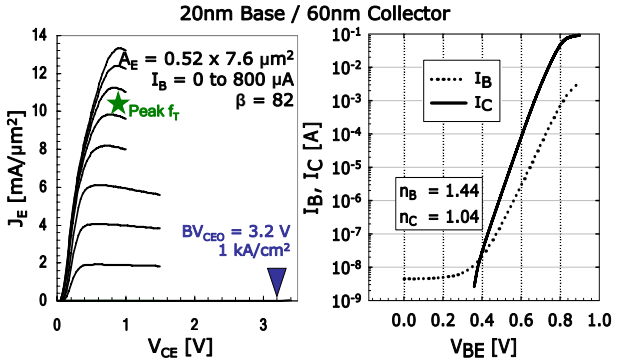

Figure 5. DC common-emitter family and forward Gummel plots of f_T = 670 GHz A_E = 0.52 x 7.6 µm^2 device on 20nm base, 60nm collector structure.

progression, and BV_{CEO} = 3.2 V defined at 1 kA/cm^2 collector current density. Peak differential β ≈ 82, and the collector and base ideality factors are n_C = 1.04 and n_B = 1.44. The $f_T \times BV_{CEO}$ product for this device is 2144 GHz-V. A smaller device on the same wafer with emitter dimensions 0.45 × 3.6 µm^2 has simultaneous f_T = 630 GHz and f_{MAX} = 350 GHz when operated at peak performance current density J_E = 12.1 mA/µm^2. At reduced current density of J_E = 5 mA/µm^2, f_T = 530 GHz and f_{MAX} = 340 GHz.

A device on the 100 nm collector structure with emitter area 0.46 × 3.1 µm^2 achieves f_T = 480 GHz and simultaneous f_{MAX} = 420 GHz when operated at J_E = 7.6 mA/µm^2. Measured RF gain curves and current density dependence of extrapolated f_T and f_{MAX} for the device are plotted in Fig. 6. DC measurements are shown in Fig. 7. The thicker collector layer increases BV_{CEO} to 4.3 V. Offset voltage is less than 0.1 V and the knee voltage is less than 0.75 V at peak performance current density. Peak differential β ≈ 38, and the collector and base ideality factors are n_C = 1.01 and n_B = 1.84. Extrapolations of current and power gain are determined from a least-squares fit to a single-pole transfer function which takes the cutoff frequency as a parameter [3]. These functions are plotted in Fig. 4 and Fig. 6 for comparison to the measured data.

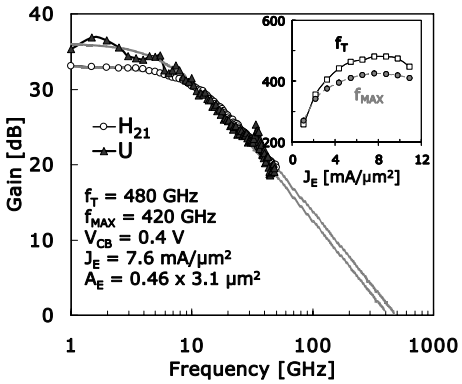

Figure 6. Measured 1-50 GHz RF current and power gains with extrapolations to f_T = 480 GHz and f_{MAX} = 420 GHz for device with A_E = 0.46 x 3.1 µm^2 on 30nm base, 100nm collector structure biased at 7.6 mA/µm^2. Extrapolated f_T and f_{MAX} as a function of emitter current density (inset).

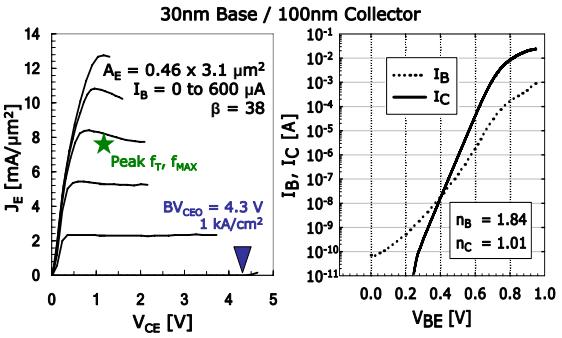

Figure 7. DC common-emitter family and forward Gummel plots for f_T = 480 GHz A_E = 0.46 x 3.1 μm² device on 30nm base, 100nm collector structure.

DEVICE MODELING

An equivalent small-signal circuit model for each device is extracted from high-frequency parameter extraction techniques and device geometry. Fig. 8 illustrates the T-model small-signal equivalent circuit and specific component values extracted for the 0.52 x 7.6 μm² device on the 20nm base, 60nm collector structure having f_T = 670 GHz. The simulated scattering parameters agree well with the measured data as shown in Fig. 8 Modeling reveals that despite lower mobility through the GaAsSb base layer, our vertically scaled type-II DHBTs have similar total electron transit times to fully pseudomorphic InP/InGaAs single heterojunction devices [4]. Total transit time for the 60 nm collector structure is 136fs ± 2% and increases to

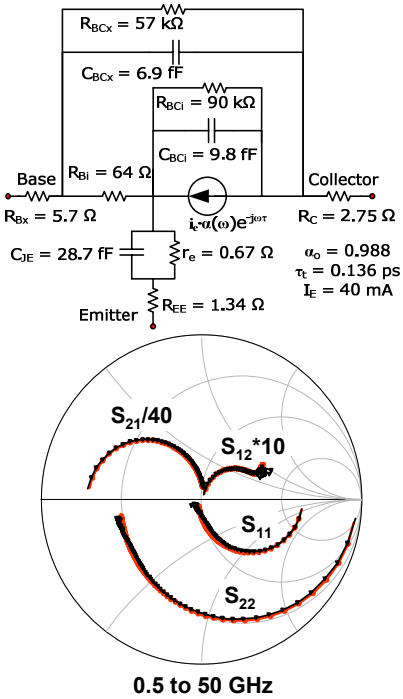

Figure 8. Equivalent small-signal circuit model of f_T = 670 GHz A_E = 0.52 x 7.6 μm² device on 20nm base, 60nm collector structure with component values extracted from measured high-frequency device parameters (top). Measured device S-parameters (black triangles) compared to simulated parameters (red circles) from 0.5 to 50 GHz (bottom).

230fs ± 3% for the 100 nm collector structure. The total f_T delay for the devices is partitioned into its base-collector junction charging, emitter-base junction charging, and electron transit delay components in Fig. 9. In addition, a uniform base device with 25 nm base and 65 nm collector is included in Fig. 9 for comparison. The graded base significantly reduces transit delay by enhancing base transport and also reduces charging delays as a result of higher operating current densities.

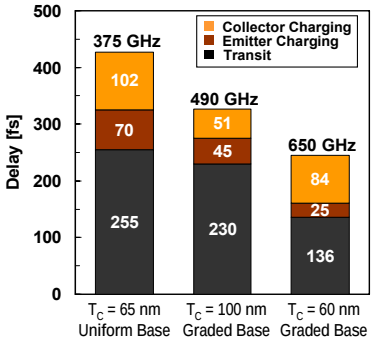

Figure 9. Total f_T delay of devices from three separate type-II DHBT material structures partitioned into base-collector RC charging, emitter-base RC charging, and electron transit delay components.

SCALING FOR THZ CUTOFF FREQUENCIES

While vertical scaling and energy band engineering successfully reduce the traditionally dominant electron transport delays in InP HBTs, parasitic RC charging delays now represent 35-50% of total f_T delay in vertically scaled HBTs with collector thicknesses < 1000 Å. Also, rising base-collector junction capacitance and base resistance lowers the power-gain cutoff frequency f_{MAX}. To offset the increasing significance of extrinsic parasitic RC charging delay, lateral (fabrication) scaling beyond the current 0.25 μm node is needed. Equation (1) represents the HBT f_T delay as the sum of two transit and two RC charging terms. Here τ_B and τ_C are the base and collector transit times. R_E, R_B, and R_C are the total emitter, base and collector resistances, respectively. C_{JE} is the base-emitter junction capacitance and C_{BC} is the base-collector junction capacitance. The $\eta kT/qI_C$ term is the dynamic resistance of the forward-biased emitter-base junction.

$$\frac{1}{2\pi f_T} = \tau_B + \tau_C + \frac{\eta kT}{qI_C}C_{JE} + (R_C + R_E + \frac{\eta kT}{qI_C})C_{BC} \qquad (1)$$

Fig. 10 illustrates a cross-section through a typical mesa HBT that is representative of the current UIUC fabrication process. An idealized HBT cross-section is also presented in Fig. 10 to illustrate potential solutions to the scaling limitations of the traditional approach. The parameters labeled in Fig. 10 are described in Table 1 along with the scaling factor that is required of each parameter to improve f_T by a factor of κ. Assuming the f_T delay is approximately equally partitioned between transit delay and charging delay, as we have shown for aggressively scaled devices, both types

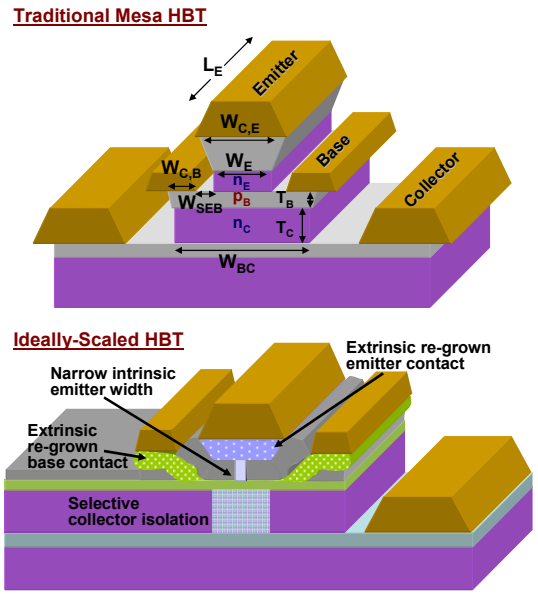

Figure 10. Illustration of traditional mesa-HBT with key design parameters and fabricated dimensions annotated (top). Ideally-scaled HBT with advanced fabrication steps to eliminate scaling trade-offs [5](bottom).

of delay must be scaled as κ^{-1}. In the case of purely diffusive transport, base transit delay is proportional to the square of the base thickness, which must therefore be scaled by $\kappa^{0.5}$. Collector transit delay is directly proportional to collector thickness, however scaling T_C as κ^{-1} will increase C_{BC} as κ^{1}. To counter this capacitance increase and achieve the desired reduction in total capacitance, the device area must then be scaled as κ^{-2} by reducing each contact and junction width.

Table 1. HBT design parameters and scaling rules for traditional mesa-HBT fabrication

Symbol	Parameter	Scaling factor
T_B	Base layer thickness	$1/\sqrt{\kappa}$
T_C	Collector layer thickness	$1/\kappa$
n_E	Emitter doping	κ
p_B	Base doping	constant (~max.)
n_C	Collector doping	κ
$W_{C,E}$	Emitter contact width	$1/\kappa^2$
W_E	Emitter-base junction width	$1/\kappa^2$
$W_{C,B}$	Base contact width	$1/\kappa^2$
W_{SEB}	Base-emitter separation	constant (~min.)
W_{BC}	Base-collector junction width	$1/\kappa^2$

In Table 2, key design and performance parameters for the current UIUC type-II DHBT and two future scaling generations are presented assuming traditional mesa HBT fabrication with realistic contact resistances and operating current densities. The model assumes electron transport remains the same as in existing devices while device dimensions and material properties scale as described in the table. As emitter widths are reduced to < 200 nm, f_T and f_{MAX} are brought into balance through reduction of base resistance and C_{BC} and ~ 1 THz simultaneous cutoff frequencies should

Table 2. UIUC type-II DHBT design and performance evolution using traditional mesa-HBT fabrication.

Parameter	300-500 nm (This work)	150-300 nm (2^{nd} Gen.)	65-150 nm (3^{rd} Gen.)
T_C	60 nm	80 nm	55 nm
T_B	20 nm	18 nm	16 nm
$W_E/W_{C,E}$	450 nm	130 nm	90 nm
ρ_{CE}	6 Ω-μm^2	4 Ω-μm^2	3 Ω-μm^2
$W_{C,B}$	150 nm	105 nm	100 nm
ρ_{CB}	9 Ω-μm^2	6 Ω-μm^2	5 Ω-μm^2
R_{SB}	2750 Ω/□	1300 Ω/□	1440 Ω/□
W_{BC}	840 nm	400 nm	350 nm
J_C	12.1 mA/μm^2	14 mA/μm^2	22 mA/μm^2
f_T	630 GHz	740 GHz	900 GHz
f_{MAX}	350 GHz	1.05 THz	1.2 THz
BV_{CEO}	3.2 V	3.95 V	3.05 V

become possible. To move beyond this milestone, the limitations of the traditional fabrication must be overcome using the ideally-scaled HBT concepts described in Fig. 10. By incorporating re-grown emitter and base contact regions. $W_{C,E}$ would no longer be determined by W_E, and similarly $W_{C,B}$ becomes independent of W_{BC}. Because of this decoupling, parasitic capacitance can be eliminated without a contact resistance penalty.

CONCLUSIONS

The type-II InP/GaAsSb device technology with low transport delays and high breakdown voltage shows real potential for achieving > 1 THz cutoff frequencies using a practical, manufacturable material design. If we are to continue to benefit from vertical scaling, InP HBTs will require more advanced process technologies resembling those currently used in SiGe HBT fabrication in order to control parasitic charging delays.

ACKNOWLEDGEMENTS

This work was supported by DARPA-MTO under the SWIFT and TFAST programs directed by Dr. Mark Rosker and Dr. Steve Pappert.

REFERENCES

[1] H. G. Liu, et al., Electron Devices, IEEE Transactions on, **53**, 559 (2006).
[2] H. G. Liu, et al., Electron Device Letters, IEEE, **25**, 769 (2004).
[3] H. K. Gummel, Proceedings of the IEEE, **57**, 2159 (1969).
[4] W. Hafez, et al., Applied Physics Letters, **86**, 152101 (2005).
[5] J. C. Zolper, 2004 IEEE International Electron Devices Meeting, IEDM Technical Digest., 30.1.1 (2003).

ACRONYMS

HBT: Heterojunction Bipolar Transistor
SHBT: Single Heterojunction Bipolar Transistor
DHBT: Double Heterojunction Bipolar Transistor
SOLT: Short-Open-Load-Through
BCB: bisbenzocyclobutene
f_T: Currrent gain cutoff frequency
f_{MAX}: Power gain cutoff frequency