Available online at www.sciencedirect.com
SCIENCE DIRECT
Solid-State Electronics 50 (2006) 733–740
SOLID-STATE ELECTRONICS
www.elsevier.com/locate/sse

Design guideline for high-speed InP/InGaAs SHBT using a practical scaling law

Daekyu Yu [a,*], Kyungho Lee [a], Kwangsik Choi [a], Bumman Kim [a], H. Zhu [b], K. Vargason [b], J.M. Kuo [b], P. Pinsukanjana [b], Y.C. Kao [b]

[a] *Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), San 31, Hyoja-dong, Nam-gu, Pohang, Gyeongbuk 790-784, North Korea*

[b] *Intelligent Epitaxy Technology (IntelliEPI), Inc. 1250E. Collins Blvd., Richardson, TX 75081, USA*

Received 4 April 2005; received in revised form 6 March 2006; accepted 13 March 2006
Available online 2 May 2006

The review of this paper was arranged by Prof. S. Cristoloveanu

Abstract

For many years, HBTs have been vertically and laterally scaled down to improve high-frequency performance. For the very small devices of recent process, some parameters cannot be scaled down properly and an alternative scaling-law is required. In this paper, we describe the optimization issues for high-speed InP/InGaAs SHBTs and offer a design guideline to accommodate the scaling limit. From a 0.25 μm SHBT designed by the scaling law, the maximum extrapolated f_{max} of about 687 GHz with f_T of 215 GHz can be achieved. We also investigate the effect of key geometrical parameters such as emitter geometry and base/collector layer thicknesses on the device RF performance.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: Heterojunction bipolar transistors; Collector–base capacitance; Base resistance; High-speed; Scaling

1. Introduction

Significant progress has been made for the development of high-speed InP-based HBTs in the 2000's. Rodwell's group at University of California, Santa Babara (UCSB) demonstrated a $0.4 \times 6\ \mu m^2$ InAlAs/InGaAs single heterojunction bipolar transistor (SHBT) with f_{max} of 1.08 THz and f_T of 204 GHz using the transferred-substrate technique [1], which is the highest f_{max} ever reported from any transistors. They could successfully minimize the collector capacitance using the structure, but the process is very complicated and may not be a manufacturable technique. In Postech, we have developed our own process technique for reduced collector capacitance using a new undercut technique and reduced emitter resistance by emitter metal widening and thick air-bridge processes in a conventional mesa-type structure [2,3]. Recently, a 0.25 μm InP/InGaAs SHBT with an f_{max} of about 687 GHz and f_T of 215 GHz was reported. It is the highest f_{max} from standard processed HBTs. Feng's group at University of Illinois at Urbana-Champaign (UIUC) used a vertically very thin base (25 nm) and collector (75 nm) layers to reduce the transit time. The $0.35 \times 12\ \mu m^2$ InP/InGaAs SHBT delivers an f_T of 506 GHz and an f_{max} of 219 GHz [4]. These HBTs have been vertically and laterally scaled down for enhanced high-frequency performance using the fundamental scaling law developed by Rodwell et al. [5]. However, due to the extremely fine geometry, several physical parameters cannot be reduced further to satisfy the scaling law [6] and an alternative scaling law is required. In this paper, we have introduced a new scaling law for the devices. Based on the rule, we describe the optimization

* Corresponding author. Tel.: +82 54 279 5584; fax: +82 54 279 8115.
E-mail address: sun13@postech.ac.kr (D. Yu).
URL: www.postech.ac.kr/ee/mmic (D. Yu).

0038-1101/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.sse.2006.03.005

issues and offer a design guideline for high-speed HBTs. Also, we analyze the effect of key geometrical parameters such as emitter geometry and base/collector layer thicknesses on RF performance. We believe that our new design guideline can offer a convenient way for implementation of scaled down HBTs for high-speed applications.

2. Scaling law of HBT

2.1. Review of fundamental scaling theory

Two measures of high-frequency performance of HBTs are given by

$$f_{max} = \sqrt{\frac{f_T}{8\pi R_B C_{CB}}}, \quad (1)$$

$$\frac{1}{2\pi f_T} = \frac{C_{JE} + C_{CB}}{g_m} + \tau_B + \tau_{SC} + (R_E + R_C)C_{CB}, \quad (2)$$

where τ_B and τ_{SC} are the base and collector transit times, respectively; C_{JE} and C_{CB} are the base–emitter and collector–base junction capacitances, respectively; g_m ($=qI_C/kT$) is the transconductance; and R_E, R_B and R_C are emitter, base and collector resistances, respectively.

A conventional mesa structure for HBTs is illustrated in Fig. 1. The large-sized HBT is scaled down by a scaling factor (γ) to produce a smaller HBT with lower resistance, inductance and capacitance. To ensure that the operation bandwidth increases by γ for all digital and analog circuits, all transit times and all capacitances must be reduced by γ, while maintaining the values for all resistances, transconductance, and the collector bias current (I_C). These relationships are summarized in Table 1 along with the scaling behavior of the important physical parameters.

2.2. Scaling limits on HBT in conventional HBT process

2.2.1. Limit to the enhancement of f_{max} (ρ_{BC}, $W_{J,BC}$ and $W_{S,EB}$)

As shown in Eq. (1), the base resistance (R_B) and collector–base capacitance (C_{CB}) are very important parameters for achieving high maximum oscillation frequency (f_{max}).

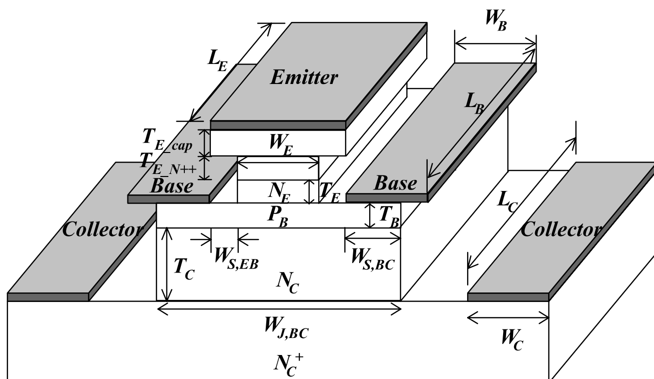

Fig. 1. The structure of conventional mesa-type HBT.

Table 1
Fundamental scaling rules on HBT

Physical parameters	Symbol	Scaling factor
Emitter thickness	T_E	γ^0
Emitter doping	N_E	γ^1
Emitter–base junction width	W_E	γ^{-2}
Emitter–base spacing	$W_{S,EB}$	γ^{-1}
Emitter–base junction length	L_E	γ^0
Emitter contact resistivity	ρ_{EC}	γ^{-2}
Base thickness	T_B	$\gamma^{-0.5}$
(Base doping) × (base hole mobility)	$(P_B) \times (\mu_{PB})$	$\approx\gamma^{0.5}$
Base contact width	W_B	γ^{-2}
Base contact length	L_B	γ^0
Base contact resistivity	ρ_{BC}	γ^{-2}
Collector thickness	T_C	γ^{-1}
Collector doping	N_C	γ^0
Collector–base junction width	$W_{J,BC}$	γ^{-2}
Collector–base junction length	L_C	γ^0
Collector current	I_C	γ^0
Collector current density	J_C	γ^2
Bias voltage	V_{CE}, v_{CE}	γ^0

C_{CB} can be minimized by narrowing the base–collector junction area or the base ohmic contact width. However, this reduction must be carried out with enhancement of the base specific contact resistivity (ρ_{BC}) because R_B can increase drastically as the base contact width becomes narrower than the contact transfer distance L_T ($= \sqrt{\rho_{BC}/R_{SB}}$). The contact resistivity of InGaAs ($p+$) has been reported to as low as $1 \sim 2 \times 10^{-7}\,\Omega\,cm^2$ and it is very difficult to reduce it any further. Therefore, both ρ_{BC} and collector base junction width ($W_{J,BC}$) cannot follow the γ^{-2} scaling rule. And, there is a trade-off for the scaling of R_B and C_{CB} in the conventional mesa-type HBT. Also, Self-Aligned Base Metal (SABM) technology has been widely used to reduce the base resistance. In the conventional HBT process with wet chemical etching using an emitter electrode mask, there is a practical limit to the minimum size of the emitter–base spacing ($W_{S,EB}$), preventing scaling.

2.2.2. Limit to the enhancement of f_T (ρ_{EC}, J_C and BV_{CEO})

Eq. (2) shows that, to improve the cut-off frequency, not only the base and collector transit times must be reduced through vertically scaling, but also the collector current density (J_C) should be increased and R_E be reduced. However, due to base push-out (or Kirk effect) at high current densities, the effective base thickness is increased. This results in an increase in the carrier transit time and a decrease in f_T. To mitigate this problem, the collector layer should be more highly doped and thinner, but the penalties are an increase of the collector–base capacitance and a reduction of the breakdown voltage (BV_{CEO}). For InP/InGaAs SHBTs, in contrast to InP double heterojunction bipolar transistors (DHBTs) and GaAs-based HBTs, the breakdown voltage drastically lowers as the collector doping level (N_C) increases, because of the complicated breakdown mechanisms such as avalanche multiplication, Zener tunneling and thermal generation of the low bandgap

InGaAs material (0.75 eV). For the low doped InGaAs collector, the doping level can be assumed constant for scaling law of InP SHBTs. Therefore, there are trade-offs for J_C, C_{CB} and BV_{CEO} in the conventional mesa-type HBT. Similarly, ρ_{BC} and ρ_{EC} cannot be reduced continuously, following a scaling factor of γ^{-2}.

2.2.3. Design issues for high-speed HBT

The schematic cross-section and high-speed optimization issues are described in Fig. 2. To obtain a high f_T and f_{max}, HBT design must achieve low base and collector transit times, properly trade-off for the required fast collector transit time with low BV_{CEO} and high C_{CB}, and minimize R_B, the extrinsic C_{CB}, and R_E. The general approach to improve the f_T is the use of thin base/collector layers. However, the approach can only be employed at the expense of reduced f_{max} due to the increased R_B and C_{CB}, and may not be suitable for applications demanding both high f_T and f_{max} [7].

Bandgap engineering (or grading scheme) is another important issue for higher f_T, especially for DHBTs. The compositional and/or doping graded base layer is very helpful for reducing the base transit time. Also, various grading schemes for the base–collector junction have been proposed to avoid the current blocking effect, such as composite collector [8], pn-pair doping [9], InGaAsP graded layers [10], staggered band lineup of InP/GaAsSb/InP [11], and optimization of ballistic carrier transport. The emitter resistance and inductance should be minimized to reduce parasitic delay. R_B is the most important factor for higher f_{max} and has three components: contact resistance, gap resistance, and intrinsic spread resistance. The minimum base resistance can be achieved using a highly doped base layer, scaling of the emitter width, self-alignment, and optimization of base ohmic contact. The C_{CB} reduction technique, independent of base contact width, is necessary for achieving both high f_T and f_{max}.

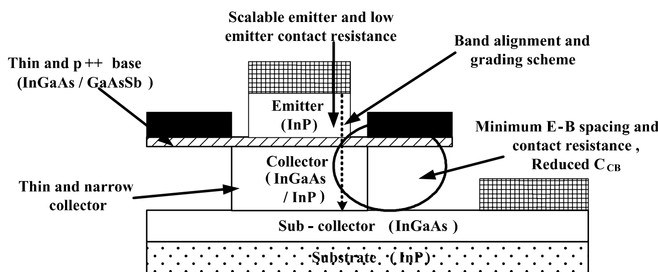

Fig. 2. Schematic representation of high-speed HBT.

2.3. Alternative scaling laws for process-limited HBTs

As we mentioned in the previous sections, a large HBT is scaled down by a scaling factor (γ) to produce a smaller HBT with smaller resistance, inductance and capacitance. To ensure that the operation bandwidth increases by γ, all transit times and all capacitances must be reduced by γ, while maintaining the values for all resistances, transconductance, and collector bias current (I_C). These scalings are modeled following Rodwell et al. [5], applying extensive scaling rule for all parameters without considering the process limit. But the collector–base junction width is not equal to emitter–base junction width for the mesa-type HBTs scaled by Rodwell et al. These parameters can be identical if the collector layer is properly undercut. In this case, the scaling becomes ideal for high-speed design and is listed as *Ideal case* in Table 2.

Actually, HBT parameters cannot be scaled down continuously due to the problems related with HBT fabrication processes for fine geometry. For process-limited devices, we need a more practical scaling law shown in Table 2. In this case, the listed parameters ($W_{S,EB}$, W_B, ρ_{EC}, ρ_{BC}, and $W_{J,BC}$) are not scaled (which is called *Normal process*), that is, have scaling factor of "zero". A large part of C_{CB} originates from the extrinsic base area in the mesa structured HBT, which can be reduced by the collector undercut process [3]. For InP HBTs which have a scalable collector using the undercut process, we propose a new rule (called *New process* in the table) scaling for collector–base junction width but without scaling the other limited parameters.

2.4. HBT performances scaled by the four rules

To compare the RF performance of HBTs as the devices are scaled down following the four rules, we have calculated equivalent circuit model parameters of the device. Based on the model, f_T and f_{max} versus emitter width are evaluated. The dimensions and epitaxial structure of the HBT used for the calculation are previously shown in

Table 2
Four different scaling rules

Physical parameters	Fundamental rules for process unlimited HBT		Practical rules for process limited HBT	
	Collector-undercut (*Ideal case*)	No-undercut (*Rodwell*)	Collector-undercut (*New process*)	No-undercut (*Normal process*)
$W_{S,EB}$	γ^{-1}		γ^0	
ρ_{EC}	γ^{-2}		γ^0	
W_B	γ^{-2}		γ^0	
ρ_{BC}	γ^{-2}		γ^0	
J_C	γ^2		$\gamma^{1\sim2}$	
$W_{J,BC}$	γ^{-2} ($W_{J,BC} = W_E$)	γ^{-2} ($W_{J,BC} > W_E$)	$\approx\gamma^{-2}$ ($W_{J,BC} \approx W_E$)	$\approx\gamma^0$ ($W_{J,BC} > W_E$)

Fig. 1 and described in Section 3. The operation condition is: $V_{BE} = 0.8$ V, $V_{CE} = 1.5$ V, and $J_C = 3\times10^5$ A/cm^2 (in the case of Fundamental Rule, it increases as the scaling factor of γ^2). The first step is to calculate the values of the resistive and capacitive elements. Resistances (R_E, R_B, and R_C) can be calculated through the majority carrier mobility, resistivity and sheet resistance, determined by doping level, material parameters and DC biases. The capacitances (C_{JE} and C_{CB}) are calculated by the parallel-plate approximation.

R_E consists of three components: emitter contact resistance ($R_{E,cont}$), resistances of the InGaAs emitter–cap layer ($R_{E,cap}$) and undepleted InP active emitter layer ($R_{E,InP}$). These parasitics can be calculated by emitter geometry, emitter doping level and emitter–base bias. R_B consists of three terms of spreading resistance under the emitter (R_{spread}), base–emitter gap resistance (R_{gap}) and contact resistance (R_{cont}), defined by [12]

$$R_B = \frac{R_{SB}W_E}{12L_E} + \frac{R_{SB}W_{S,EB}}{2L_E} + \frac{\sqrt{\rho_{BC}R_{SB}}}{2L_E}\coth\left(W_{S,BC}\sqrt{\frac{R_{SB}}{\rho_{BC}}}\right), \quad (3)$$

where L_E and W_E are the emitter–base junction length and width, respectively; $W_{S,EB}$ is the emitter–base spacing; ρ_{BC} and R_{SB} are the base specific contact resistivity and the base sheet resistance, respectively. We use $2\times10^{-7}\ \Omega\,\mathrm{cm}^2$ and $1.65\times10^{-7}\ \Omega\,\mathrm{cm}^2$ as the emitter and base contact resistivities, respectively, in 1.0 μm emitter geometry, and they are scaled down as the scaling factor of γ^{-2} in the case of *Fundamental Rule* but are kept constant in the case of *Practical Rule*. C_{CB} for f_{max} consists of three terms of the collector junction capacitance lying under the emitter ($C_{CB,E}$), under the gap between the emitter and the base contact ($C_{CB,gap}$), and under the base ohmic contact ($C_{CB,ext}$), and is given by

$$C_{CB} = \varepsilon_{InGaAs}\frac{W_E L_E}{T_C} + 2\varepsilon_{InGaAs}\frac{(W_{S,EB}+W_{S,BC})L_E}{T_C} + 2\varepsilon_0\frac{(W_B - W_{S,BC})L_E}{T_C}, \quad (4)$$

where T_C is the collector thickness; W_B and $W_{S,BC}$ are the base contact width and the collector–base junction width, respectively; ε_0 and ε_{InGaAs} are the permittivity of free space and InGaAs, respectively. Although the base–collector junction parasitics are distributed in nature, we simply compute the τ_{CB} as follows [13]:

$$\tau_{CB} = C_{CB,E}(R_{cont} + R_{gap} + R_{spread}) + C_{CB,gap}(R_{cont} + R_{gap}/2) + C_{CB,ext}R_{cont}. \quad (5)$$

To obtain f_T, we calculate the charging times and transit times as follows:

$$\frac{1}{2\pi f_T} = \frac{C_{JE}+C_{CB}}{g_m} + \tau_B + \tau_{SC} + (R_E + R_C)C_{CB}, \quad (6)$$

where

$$\tau_B = \frac{T_B^2}{2D_{nb}}, \quad \tau_{SC} = \frac{X_{dep,CB}}{2V_{sat}}, \quad (7)$$

where T_B is the base thickness; $X_{dep,CB}$ is the collector–base depletion thickness. We calculate base and collector transit times using the parameters of minority carrier diffusivity in the InGaAs base layer C-doped to 8.0×10^{19} cm^{-3} as $D_{nb}\approx 70$ cm^2/s and saturation velocity in InGaAs collector layer as $V_{sat}\approx 3\times10^7$ cm/s. Then, we can calculate f_{max} using the above results

$$f_{max} = \sqrt{\frac{f_T}{8\pi\tau_{CB}}}. \quad (8)$$

Fig. 3 shows the calculated RF performance (f_T and f_{max}) of an HBT designed for high f_{max} operation. The emitter length (L_E) is fixed to 10 μm. As expected, the *Ideal Case* delivers the best RF performance with f_{max} above 1 THz for 0.1 μm emitter geometry. The HBT scaled by the *New* rule shows good RF performance due to the low R_B and C_{CB} through the undercut process, although ρ_{BC} and it $W_{J,BC}$, which are the scaling limited parameters, are not scaled. However, the scaled HBT using the *Normal process* has rather low f_T and f_{max} due to the high extrinsic C_{CB}.

To achieve a high f_T, the device must be vertically scaled down further using thin base and collector layers. In this case, the f_{max} is inevitably reduced due to the increased R_B and C_{CB}. Fig. 4 shows the comparison of the RF performances (f_T and f_{max}) of the device optimized for high f_T using the thin layers. The HBT scaled using the *Ideal case* rule delivers the best RF performance of about 600 GHz f_T and 900 GHz f_{max} from a 0.1 μm emitter HBT. These high f_T and f_{max} for the *Ideal case* are achieved by scaling of J_C, R_B and $C_{CB,ext}$. As shown, the maximum f_T is reached for a 0.1 μm HBT because the emitter

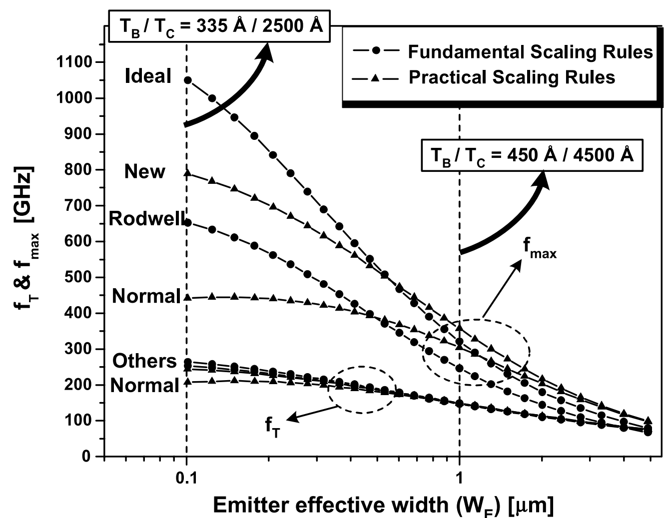

Fig. 3. Comparison of the RF performances (f_T and f_{max}) as the device is scaled down following the four scaling rules. The HBT is designed for high f_{max} operation.

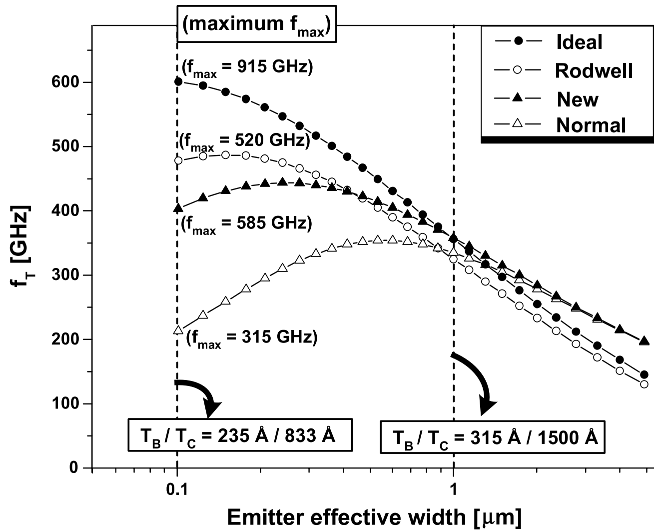

Fig. 4. Comparison of f_T as the device is scaled down following the four scaling rules. The HBT is designed for high f_T operation.

resistance (R_E), in spite of emitter contact resistivity (ρ_{EC}) scaling, is not perfectly scaled down. For the small emitter size HBT, the emitter epitaxial layer resistance ($R_{E,epi}$) is no less important than the emitter contact resistance ($R_{E,cont}$). And to prevent f_T degradation due to emitter charging time, $R_{E,epi}$ must be kept small by scaling the emitter layer thickness. In the *New* scaling rule, f_T falls off for a small geometry due to the large emitter resistance (R_E) and constant collector current density (J_C). As the device is vertically scaled down, the emitter charging time ($\tau_E = (C_{JE} + C_{CB})/g_m$) and collector charging time ($\tau_C = (R_E + R_C)C_{CB}$) become very important, comparable to the base–collector transit time ($\tau_B + \tau_{SC}$). For the *Normal* scaling rule, f_T fall-off is very significant due to the increase of $C_{CB,ext}$.

3. Experimental results and discussions

3.1. Device technology

We have fabricated small-size HBTs using our practical scaling rule (*New process*). The epitaxial layer of the HBTs is grown by Solid Source Molecular Beam Epitaxy on a Fe-doped semi-insulating (1 0 0) InP substrate. The layer structure includes, from the top, InGaAs emitter contact layer, InGaAlAs graded layer, InP emitter layer (700 Å, Si-doped to 7.0×10^{17} cm^{-3}), InGaAs base layer with Indium mole fraction graded from of 0.46 to 0.53 (400 Å, C-doped to 8.0×10^{19} cm^{-3} with 20 Å spacer), InGaAs collector layer (2500 Å, Si-doped to 2.0×10^{16} cm^{-3}), and InGaAs subcollector layer (6000 Å, Si-doped to 1.0×10^{19} cm^{-3}). Lastly, an InP etch-stop layer is inserted in the middle of the collector layer for the collector undercut process.

The devices are fabricated using a standard mesa process. Emitters are defined using contact lithography, resulting in a minimum emitter metal width of about 0.5 μm. The emitter is slightly undercut, resulting in an effective 0.25 μm size. The self-aligned Pt/Ti/Pt/Au base metal having 1 μm width is evaporated. The base and collector layers are etched with a citric-based etchant. In this etching process, the collector undercut process is also carried out. Next, a Ti/Au emitter widening metal is evaporated. The subcollector is etched for device isolation. In this etching process, the active base area and the base area for the interconnecting base pad are isolated. Next, the Ti/Pt/Au collector metal and metal pad are evaporated. Lastly, an Au air-bridge is formed [14]. The whole process flow and detailed process are described in [15]. SEM pictures of a fabricated HBT are shown in Fig. 5.

There are three geometrical variables associated with scaling: emitter width (W_E), emitter length (L_E) and base/collector thickness (T_B/T_C). We compare results of InP/InGaAs SHBTs with: $A_E = 0.25 \times 10$ μm^2, 0.5×10 μm^2, 1.0×10 μm^2 (emitter width (W_E) variation); 0.5×8 μm^2, 0.5×10 μm^2, 0.5×15 μm^2 (emitter length (L_E) variation); and 1.0×20 μm^2 with $T_B/T_C = 400$ Å/2500 Å, 400 Å/4000 Å, and 600 Å/6000 Å (base/collector thickness (T_B/T_C) variation). The microwave performance of these devices is characterized by on-wafer *S*-parameter measurements for 0.5–40 GHz (and 50–75 GHz for devices with 0.25 μm emitter width) band using an Agilent 8510C network analyzer calibrated by through-reflect-line method.

To analyze the effect of the geometrical parameters (emitter geometry and base/collector layer thicknesses) on RF performance, we have extracted the small signal equivalent circuit parameters. The junction depletion capacitances and bias independent parasitic capacitances are extracted from "cold" *S*-parameters under reverse and low forward biases. The series resistances are directly determined from "over-driven" *S*-parameters [16]. Also, the values for the transit times are calculated from the measured *S*-parameters. By plotting f_T vs. $1/I_C$ and fitting a straight line, the intersection at the y axis provides the value for $\tau_F = \tau_B + \tau_{SC} + (R_E + R_C)C_{CB}$ and the slope leads to the emitter charging time τ_E.

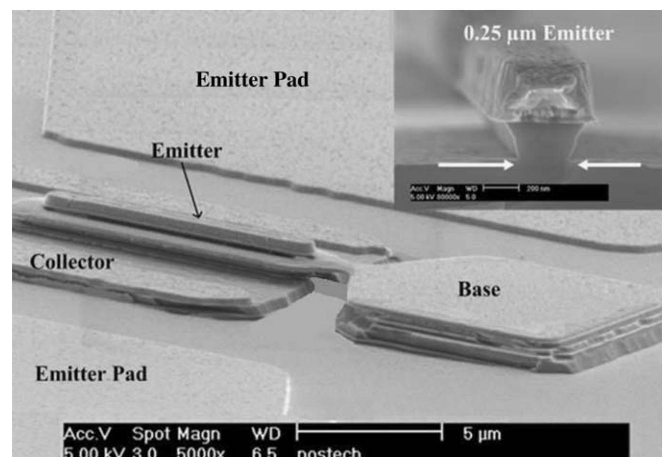

Fig. 5. SEM pictures of the fabricated HBT with a 0.25×10 μm^2 emitter area.

3.2. Lateral scaling for emitter size variation (W_E and L_E)

The effects of the emitter width (W_E) variation on RF performance are shown in Fig. 6. Emitter length (L_E) and base/collector thickness (T_B/T_C) are held constant at 10 μm and 400 Å/2500 Å, respectively. As W_E decreases from 1.0 to 0.25 μm, f_{max} is improved considerably due to reduced R_B and C_{CB} by lateral scaling. However, f_T decreases slightly because the emitter charging time increases (from 0.13 to 0.15 ps) due to a decrease in transconductance (g_m, from 0.92 to 0.32 Ω^{-1}) and because the collector charging time increases (from 0.021 to 0.025 ps) due to a higher R_E (from 1.6 to 6.7 Ω) even though C_{CB} decreases from 6.7 to 1.9 fF. Still, the effect of emitter and collector charging times on f_T is small, about 24% (emitter and collector charging times are 0.15 and 0.03 ps, respectively, and $\tau_B + \tau_{SC} = 0.57$ fF for 0.25 μm emitter device). These devices have a large base–collector transit time due to the thick collector for high f_{max} suitable for analog/RF applications. For comparison, the calculated values with the scaling rules outlined above are also plotted in Fig. 6, and they are in good agreement with the experimental results.

The effects of L_E variation on RF performance are shown in Fig. 7. W_E and T_B/T_C are held constant at 0.5 μm and 400 Å/2500 Å, respectively. As L_E decreases from 15 to 8 μm, the reduction of C_{CB} (from 9.85 to 3.44 fF) leads to considerable improvement in f_{max}, in spite of a slight increase of R_B (from 7.8 to 10.4 Ω). Also, the emitter charging time increases (from 0.10 to 0.16 ps) due to the reduced transconductance (g_m, from 0.91 to 0.52 Ω^{-1}) and the collector charging time decreases slightly (from 0.053 to 0.037 ps) by the reduced C_{CB}, resulting in a slight decrease of f_T.

3.3. Vertical scaling for base/collector thicknesses

The effects of the base/collector thickness (T_B/T_C) variation on RF performance are shown in Figs. 8 and 9. The

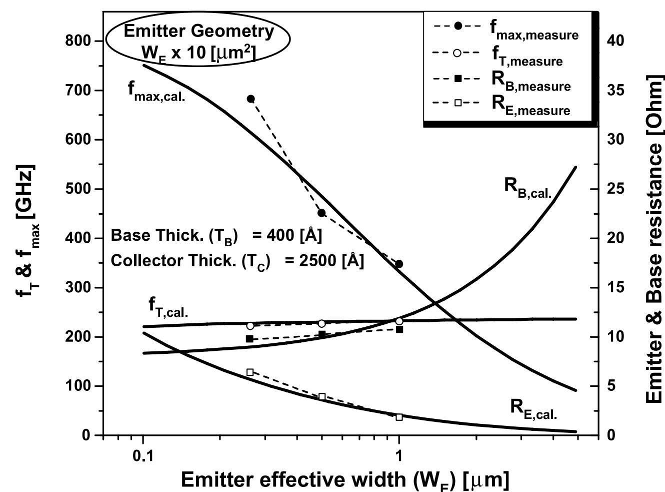

Fig. 6. RF performance (f_T and f_{max}) and resistances (R_E and R_B) versus emitter effective width (W_E) for the HBT with $T_B/T_C = 400$ Å/2500 Å and $L_E = 10$ μm.

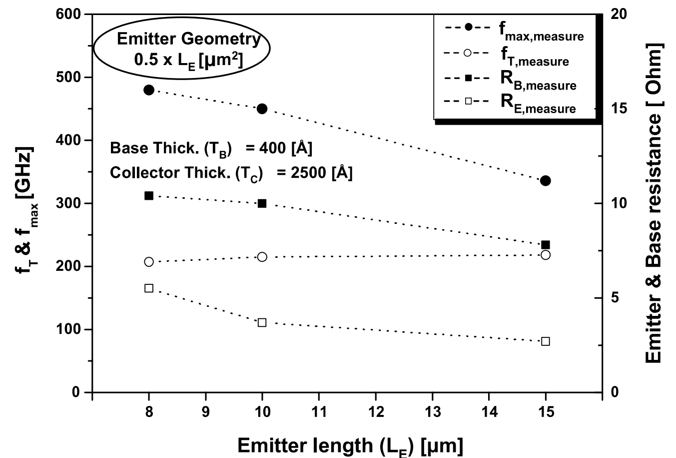

Fig. 7. RF performances (f_T and f_{max}) and resistances (R_E and R_B) versus emitter length (L_E) for the HBT with $T_B/T_C = 400$ Å/2500 Å and $W_E = 0.5$ μm.

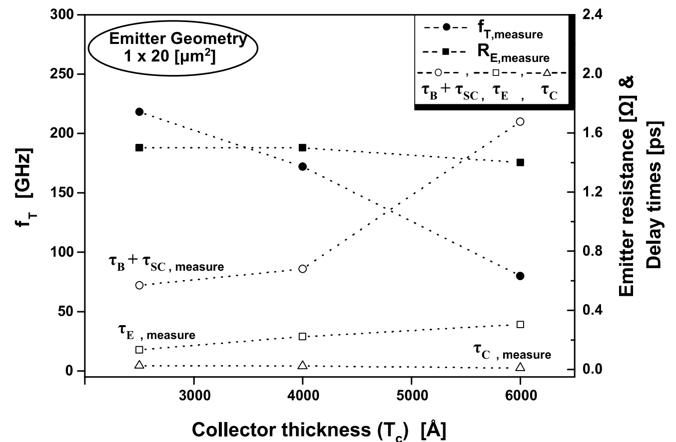

Fig. 8. Transistor current gain cut-off frequency (f_T), emitter resistance (R_E) and delay times ($\tau_B + \tau_{SC}$, τ_E, τ_C) versus base/collector thickness (T_B/T_C) for 1.0×20 μm² HBT.

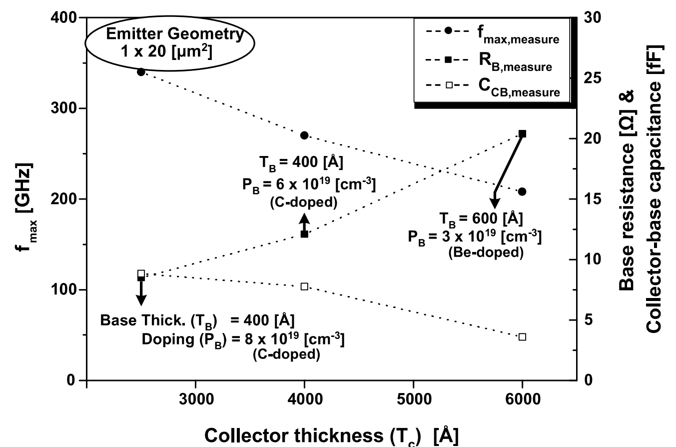

Fig. 9. Maximum oscillation frequency (f_{max}), base resistance (R_B) and collector–base capacitance (C_{CB}) versus base/collector thickness (T_B/T_C) for 1.0×20 μm² HBT.

emitter area (A_E) is held constant at $1.0 \times 20\ \mu m^2$ for constant R_E. As the device is vertically scaled down from 600 Å (T_B)/6000 Å (T_C) to 400 Å/2500 Å, the transit-time ($\tau_B + \tau_{SC}$) decreases from 1.68 to 0.57 ps and the emitter charging time also decreases from 0.30 to 0.13 ps, due to an increase in transconductance (g_m, from 0.55 to 1.85 Ω^{-1}). However, the collector charging time (τ_C) increases slightly from 0.01 to 0.03 ps for the larger C_{CB}. The reduction of the total delay time (τ_{total} from 1.99 to 0.73 ps) leads to a considerable improvement in f_T (from 80 to 218 GHz).

As indicated in Fig. 9, the base sheet resistance is optimized from 640 to 472 Ω/□ by increasing the base layer doping level (from 600 Å of $3 \times 10^{19}\ cm^{-3}$ Be doping to 400 Å of $8 \times 10^{19}\ cm^{-3}$ C doping). We changed the base layer dopant from Be (*Beryllium*) to C (*Carbon*), because C-doped base layer has higher mobility and better ohmic characteristic than Be-doped one. Therefore, the base resistance (R_B) is reduced from 20.4 to 8.5 Ω in spite of the thinner layer. With improvement of f_T (from 80 to 218 GHz), f_{max} is also increased, from 208 to 340 GHz.

3.4. Scaling trend and challenges for higher f_T and f_{max}

Fig. 10 shows the scaling trend of InP/InGaAs SHBT fabricated in Postech. We designate the 600 Å base and 6000 Å collector layers as the 1st generation InP/InGaAs SHBT because this structure is designed for high f_{max}. The 1st generation SHBT with 1.0 μm emitter geometry has 80 GHz f_T and 208 GHz f_{max}. Then the device is laterally and vertically scaled down for improved f_T and f_{max}. As a result, the high-frequency performance of $f_T =$ 154 GHz and $f_{max} = 478$ GHz is obtained from the 2nd generation SHBT with 0.5 μm emitter width, 400 Å base and 4000 Å collector thicknesses. Subsequently, vertical and lateral scalings of the device is continued by reducing T_C from 4000 Å to 2500 Å and shrinking W_E from 0.5 μm to 0.25 μm, respectively. For the 3rd generation HBTs, we find a maximum extrapolated f_{max} of about 687 GHz and f_T of 215 GHz, as shown in Fig. 11. This is the highest f_{max} ever reported for a mesa-type HBTs [17]. As the device is vertically and laterally scaled down further using the *New* rule, the SHBT is expected to deliver RF performance with $f_{max} \geqslant 800$ GHz and f_T around 300 GHz or $f_T \geqslant 500$ GHz and f_{max} around 600 GHz. For the *New* scaling law, f_T is limited by the emitter and collector charging times due to the large R_E. For further improvement, we must find a new process/structure scheme to reduce R_E and epi-design to increase the Kirk current level (J_K). To reduce $R_{E,cont}$, ρ_{EC} must be improved following the γ^{-2} scaling factor or the emitter contact area must be maintained large through the emitter regrowth process [18]. As the device is scaled down laterally, the emitter epitaxial resistance ($R_{E,epi}$) is no less important than $R_{E,cont}$, and $R_{E,epi}$ must be kept low with emitter layer thickness scaling.

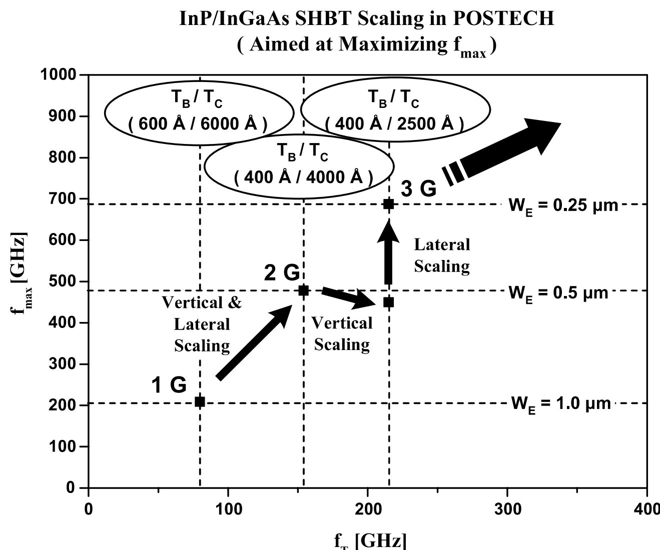

Fig. 10. Scaling trend of InP/InGaAs SHBT fabricated in Postech.

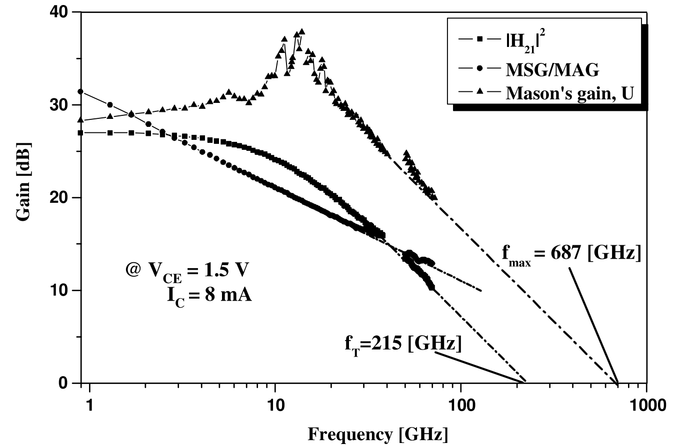

Fig. 11. Frequency dependencies of $|h21|^2$, MSG/MAG, and Mason's gain for the fabricated HBT with a $0.25 \times 10\ \mu m^2$ emitter area.

4. Conclusion

For highly scaled down HBTs, several physical parameters cannot be reduced following the conventional scaling law [5]. Therefore, we have suggested an alternative scaling law which is practically achievable. Following the *New* rule, we have outlined a 0.25 μm SHBT with maximum extrapolated f_{max} of about 687 GHz and f_T of 215 GHz. As the device is vertically and laterally scaled down further, it will be possible to demonstrate RF performance of $f_{max} \geqslant 800$ GHz with f_T around 300 GHz or $f_T \geqslant$ 500 GHz and f_{max} around 600 GHz. The device scaling using the *New* law has a limit on f_T due to the large R_E and constant Kirk current level (J_K) producing a large emitter and collector charging times. Therefore, to achieve RF performance above the suggested values, we must improve the process scheme to reduce R_E and epi-design to increase J_K. We believe that our *New* scaling rule offers a practical way for the design of high-speed HBTs with ultrasmall geometry.

Acknowledgement

This work was supported in part by the Korean Ministry of Education under BK21 project and the center for Broadband OFDM Mobile Access (BrOMA) at POSTECH through the ITRC program of the Korean MIC, supervised by IITA (IITA-2005-C1090-0502-0008).

References

[1] Rodwell MJW, Betser Y, Jaganathan S, Mathew T, Sundararajan PK, et al. Submicron lateral scaling of HBTs and other vertical-transport devices: towards THz bandwidths. In: GaAs Conf 2000; Proc; 2000. p. 1–4.

[2] Lee K, Yu D, Chung M, Kang J, Kim B. New collector undercut technique using a SiN sidewall for low base contact resistance in InP/InGaAs SHBTs. IEEE Trans Electron Dev 2002;49(6):1079–82.

[3] Yu D, Choi K, Lee K, Kim B, Ontiveros D, et al. Realization of high-speed InP SHBTs using novel but simple techniques for parasitic reduction. In: IEEE Internal Conference Indium Phosphide and Related Materials (IPRM) 2004; Symp-2004. p. 753–6.

[4] Hafez M, Lai JW, Feng M. InP/InGaAs SHBTs with 75 nm collector and $f_T > 500$ GHz. Electron Lett 2003;39(20):1475–6.

[5] Rodwell MJW, Urteaga M, Mathew T, Scott D, Mensa D, et al. Submicron scaling of HBTs. IEEE Trans Electron Dev 2001;48(11):2606–24.

[6] Zolper JC. Super-scaled InP HBTs for 150 GHz circuits. In: IEEE Int Electron Dev Meeting (IEDM) 2003; Dig-2003. p. 30.1.1–4.

[7] Kim B, Lee K, Yu D. Current status of millimeter-wave transistor technology. Plenary talk, In: Topical Symposium on Millimeter Waves (TSMMW) 2004. p. 1–5.

[8] Feygenson A, Hamm RA, et al. High speed InGaAs/InP composite collector bipolar transistors. IEEE Trans Electron Dev 1992;39(11):2658.

[9] Kurishima K, Nakajima H, et al. High-speed InP/InGaAs double-heterostructure bipolar transistors with suppressed collector current blocking. Appl Phys Lett 1993;62(19):2372–4.

[10] Ida M, Kurishima K, et al. InP/InGaAs DHBTs with 341-GHz f_T at high current density of over 800 kA/cm^2. In: IEEE Int Electron Dev Meeting (IEDM) 2001; Dig-2001. p. 776–9.

[11] Bolognesi CR, et al. Investigation of high-current effects in staggered lineup InP/GaAsSb/InP hetero-junction bipolar transistors: temperature characterization and comparison to conventional type-I HBTs and DHBTs. In: IEEE Int Electron Dev Meeting (IEDM) 2001; Dig-2001. p. 768–71.

[12] Liu W. Handbook of III–V heterojunction bipolar transistor. New York: Wiley; 1998.

[13] Vaidyanathan M, Pulfrey DL. Extrapolated f_{max} of heterojunction bipolar transistors. IEEE Trans Electron Dev 1999;46(2):301–9.

[14] Kim W, Lee K, Chung M, Kang J, Kim B. High-speed AlGaAs/GaAs HBTs with a reduced base–collector capacitance. Electron Lett 2001;37(20):1259–61.

[15] Yu D, Lee K, Kim B, Ontiveros D, Vargason K, et al. Ultra high-speed InP/InGaAs SHBTs with f_{max} of 478 GHz. IEEE Electron Dev Lett 2003;24(6):384–6.

[16] Gobert Y, Tasker PJ, Bachem KH. A physical, yet simple, small-signal equivalent circuit for the heterojunction bipolar transistor. IEEE Trans Microwave Theory Tech 1997;45(1):149–53.

[17] Yu D, Choi K, Lee K, Kim B, et al. Ultra high-speed 0.25 μm emitter InP-InGaAs SHBTs with $f_{(max)}$ of 687 GHz. In: IEEE Int Electron Dev Meeting (IEDM) 2004; Dig-2004. p. 557–60.

[18] Scott DW, Wei Y, Dong Y, Gossard AC, Rodwell MJ. A 183 GHz f_T and 165 GHz f_{max} regrown-emitter DHBT with abrupt InP emitter. IEEE Electron Dev Lett 2004;25(6):360–2.