International Journal of High Speed Electronics and Systems
Vol. 17, No. 1 (2007) 61–80

World Scientific
www.worldscientific.com

PRESENT STATUS AND FUTURE DIRECTIONS OF SiGe HBT TECHNOLOGY

Marwan H. Khater[a], Thomas N. Adam[b], Rajendran Krishnasamy[c], Mattias E. Dahlstrom[c], Jae-Sung Rieh[d], Kathryn T. Schonenberg[a], Bradly A. Orner[c], Francois Pagette[a], Kenneth Stein[b], and David C. Ahlgren[b]

[a]*IBM T. J. Watson Research Center, 1101 Kitchawan Road, P. O. Box 218, Yorktown Height, New York 10598, USA*
mkhater@us.ibm.com

[b]*IBM Microelectronics, 2070 Route 52, Hopewell Junction, New York 12533, USA*

[c]*IBM Microelectronics, 1000 River Street, Essex Junction, Vermont 05452, USA*

[d]*School of Electrical Engineering, Korea University, Seoul, Korea 136-701*

The implementation of challenging novel materials and process techniques has led to remarkable device improvements in state-of-the-art high-performance SiGe HBTs, rivaling their III-V compound semiconductor counterparts. Vertical scaling, lateral scaling, and device structure innovations required to improve SiGe HBTs performance have benefited from advanced materials and process techniques developed for next generation CMOS technology. In this work, we present a review of recent process and materials development enabling operational speeds of SiGe HBTs approaching 400 GHz. In addition, we present device simulation results that show the extendibility of SiGe HBT technology performance towards half-terahertz and beyond with further scaling and device structure improvements.

Keywords: SiGe HBTs; vertical scaling; lateral scaling; self-aligned structure; raised extrinsic base.

1. Introduction and Overview

The improvement in SiGe HBT transistor performance, especially the operation speed, is an essential requirement for increased bandwidth and data transfer rates in modern network communication systems. SiGe HBTs have been favored for RF/analog/mixed-signal applications owing to their advantages in transconductance, $1/f$ noise, device matching, and power performance, as compared to CMOS transistors. In addition, a monolithic integration compatibility with standard CMOS technologies and high reliability has made them an attractive and low-cost alternative to III-V HBT technologies. Recent developments in SiGe HBT transistor technology allowed operation speeds approaching 400 GHz[1-4], minimum gate delays below 3.3 ps[3-5], and enabled circuits operation at 60 GHz[6]. The cut-off frequency (f_T) and maximum oscillation frequency (f_{MAX}) of SiGe HBTs have been improved significantly by vertical and lateral scaling enabled by modern CMOS process techniques readily available for SiGe HBT

and BiCMOS technologies. However, scaling of SiGe HBTs has limitations similar to those encountered in CMOS technology that can only be overcome by the advancement of new materials, process techniques, and structural innovations[3,4,7,8,9].

The performance of SiGe HBTs in recent years, compared to current InP-based HBT technologies[10-17], is shown in Fig. 1. Note that the common trend in HBT technology development is to achieve a device with balanced f_T and f_{MAX}. Aggressive vertical scaling and impurity-profile engineering of the SiGe base, enabled by modern epitaxial growth techniques, allowed SiGe HBTs to operate at f_T near 400 GHz[2,4], approaching frequencies once thought achievable only with III-V material based HBTs. In addition, as shown in Fig. 1, the capability of SiGe HBTs operation at f_T above 500 GHz was recently demonstrated by operating the device at a temperature of 4.5 K[18]. The improvement in f_T, however, trades off and limits f_{MAX} of the device due to increased parasitic capacitance and resistance caused by vertical scaling according to the approximate relation

$$f_{MAX} \approx \sqrt{\frac{f_T}{8\pi R_B C_{CB}}} \tag{1}$$

where, R_B is the total base resistance and C_{CB} is the collector-to-base capacitance. The parasitic (i.e. extrinsic) components of R_B and C_{CB} can be optimized to further improve f_{MAX} by lateral scaling and device structure modification enabled by advanced CMOS-compatible lithography and process techniques. A significant structural improvement of SiGe HBTs is the implementation of a raised extrinsic base self-aligned to the emitter, which allows reduction of R_B and C_{CB} independently[7,8]. Lateral scaling and device structure improvements enabled SiGe HBTs to operate at f_{MAX} up to 350 GHz[3].

In this paper, state-of-the-art SiGe HBTs performance is reviewed with emphasis on scaling, materials, process techniques, and structural modifications enabling the device performance improvement. In addition, future directions of SiGe HBT technology are discussed based on process and device simulations of further scaling, new process technologies, and materials implementation to further improve the device performance.

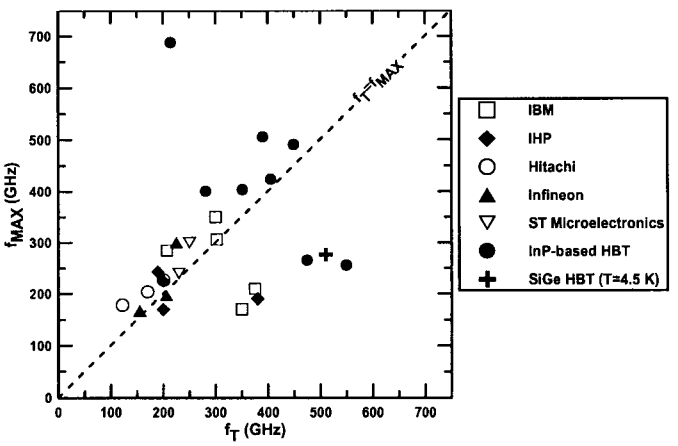

Fig. 1. Evolution of f_T and f_{MAX} of SiGe HBTs and InP-based HBTs.

2. SiGe HBT Device Structure

The schematic and SEM cross-section views of a modern SiGe HBT device are shown in Fig. 2. Combined deep and shallow trenches (DT and STI) provide device isolation. A buried subcollector and an n^- epitaxial layer form the collector region along with the selectively-implanted collector (SIC) pedestal. A boron-doped SiGe:C base layer is grown by non-selective UHV/CVD, and a boron-doped polysilicon raised extrinsic base is formed self-aligned to the in-situ phosphorus-doped emitter. A cobalt silicide ($CoSi_2$) formed on the raised extrinsic base polysilicon and collector reach-through serves as an ohmic low-resistance contact layer to the base and collector.

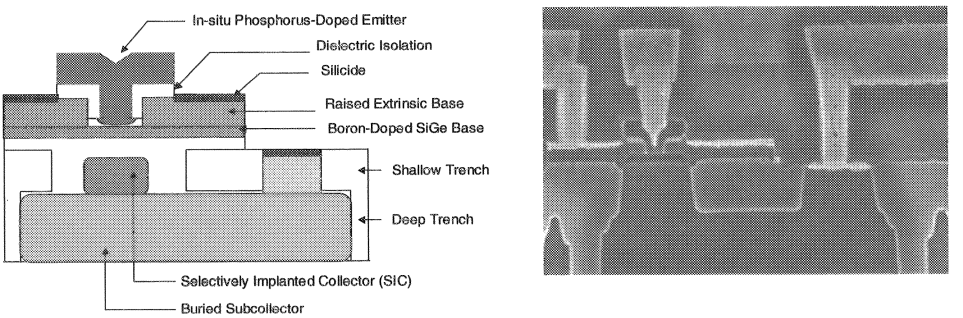

Fig. 2. Schematic and SEM cross section views of SiGe HBT with raised extrinsic base.

A significant structural improvement of SiGe HBTs is the implementation of a polysilicon raised extrinsic base self-aligned to the emitter, where the emitter-to-base spacing is determined by a spacer width, as shown in Figs. 2, rather than implanted extrinsic base[19,20]. The extrinsic base is usually doped by boron implantation. The ion implant conditions for an implanted extrinsic base have to be optimized to balance the trade-off between R_B and C_{CB}. Low R_B requires high dose and energy implant, whereas low C_{CB} requires low dose and energy implant. In other words, R_B and C_{CB} are coupled in an implanted extrinsic base. In addition, implanting the extrinsic base limits lateral scaling of the device, where the implant creates silicon interstitials which cause the base dopants to diffuse at high temperature processing[20,21]. This limits the proximity of the extrinsic base to the intrinsic base portion of the device and thus limits base and emitter scaling. On the other hand, a raised extrinsic base, which allows independent optimization of R_B and C_{CB}, minimizes both of these effects by locating the low resistance region above the intrinsic base. In this case, both the base implant defects and the intersection of the collector and base dopants in the intrinsic device are eliminated. In addition, rather than implanted, the raised extrinsic base polysilicon could be in-situ doped with boron using CVD techniques to eliminate implant defects and further improve the device performance. High boron doping levels on the order of 10^{19}-10^{21} cm^{-3} can be achieved by implantation or in-situ doping techniques to obtain polysilicon extrinsic base with sufficiently low sheet resistance.

The critical lateral scaling dimensions and components of R_B and C_{CB} for the self-aligned device structure with raised extrinsic base are shown in Fig. 3. The emitter width (W_E) is defined by lithography and a spacer width (W_S) formed in a similar fashion

implemented in CMOS technology. In addition, the collector width (W_C), SIC implant width (W_{SIC}), and base silicide-to-emitter spacing (D) are also defined by lithography.

Total C_{CB} components include the intrinsic capacitance (C_{int}) and the extrinsic capacitance (C_{ext}). C_{int} is determined by the SIC doping levels and width as well as the emitter width, W_E. C_{ext} includes many components that are governed by the device structure and process techniques. The capacitance under and adjacent to the spacer (C_{SIC}) is influenced by the SIC implant width, W_{SIC}, and dopant lateral diffusion due to processing at high temperatures. The link capacitance component (C_{link}) is controlled by the link area overlap between the collector and base. The capacitance over STI (C_{STI}) is determined by the overlap area between the collector and base over STI and the STI material dielectric properties.

Total R_B components include the intrinsic base resistance (R_{int}) and the extrinsic base resistance (R_{ext}). R_{int} is determined by the intrinsic SiGe base thickness and doping levels as well as the emitter width, W_E. R_{ext} includes many components that are determined by the device structure and process techniques. The base resistance under the spacer (R_{spr}) is determined by the spacer width, W_S. The link resistance component (R_{link}) is controlled by the contact area and quality of the interface between the intrinsic and extrinsic base. The extrinsic base polysilicon resistance component (R_{poly}) is governed by the boron doping level, polysilicon thickness, and spacing D between the silicide and emitter. The silicide contact resistance component (R_{sc}) is controlled by the silicide to silicon interface properties, while the silicide resistance component (R_{sil}) is determined by the silicide type.

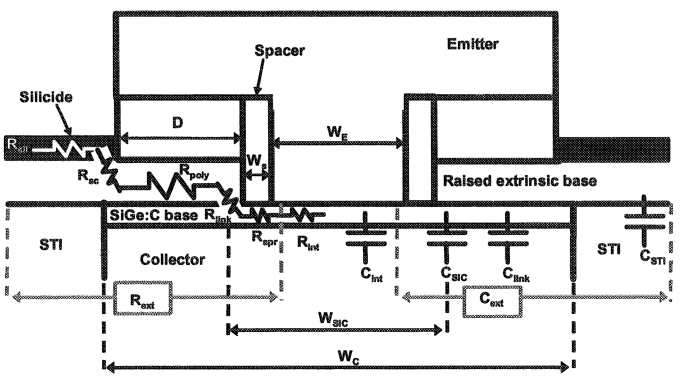

Fig. 3. Critical lateral scaling dimensions and components of R_B and C_{CB} for a self-aligned SiGe HBT device structure with raised extrinsic base.

3. SiGe HBT Technology: Current Status

3.1. *Vertical Scaling and Impurity Profile Engineering*

Vertical scaling and dopant profile optimization of the collector and SiGe base are key factors that help improve the device f_T. More specifically, the thicknesses and doping levels determine the delay times in the "*intrinsic*" device. f_T is improved as carrier transit-times and sheet resistances are reduced by thinning the collector and base layers and by boosting their doping levels.

Collector vertical scaling and dopant profile optimization are performed by controlled changes during the formation of the in-situ doped subcollector and selectively-implanted collector (SIC) region, as schematically depicted in Fig. 4. An in-situ doped subcollector layer is epitaxially grown by chemical vapor deposition (CVD). Subsequently, an n^- layer is epitaxially grown to separate the active device from the subcollector region. The doping level in the n^- layer is lower at the collector-base junction to reduce C_{CB} in the extrinsic device. However, the collector doping in the intrinsic device needs to be sufficiently high to reduce the collector resistance (R_C) in order to improve f_T. This is achieved by selective implantation of the intrinsic device using lithographic techniques, which can be properly scaled (i.e. W_{SIC}) to reduce R_C without significant impact on C_{CB} in the extrinsic device. A carefully designed SIC implant is adequate to provide a sufficient link conductance between the base and a subcollector with thickness below 0.5 μm. SIC implant energies and doses are usually in the range between 30 to 120 keV and 10^{14} to 10^{16} cm^{-2}, respectively.

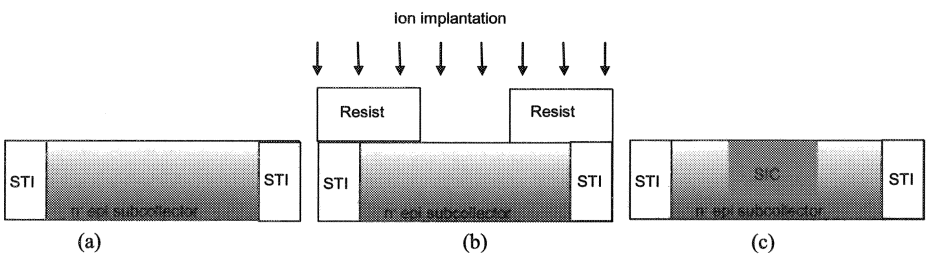

Fig 4. Schematic depiction of SiGe HBT collector formation: (a) epitaxial growth of n^- and subcollector layers and STI formation, (b) and (c) SIC formation using lithography.

The effect of normalized SIC total dose on C_{CB}, f_T, and f_{MAX} is shown in Fig. 5. The device SiGe base was optimized with 100% SIC total dose to obtain balanced f_T and f_{MAX} of 300 GHz[22]. As can be seen from Fig. 5, as the SIC total dose is increased to 100% to reduce R_C, C_{CB} increases by 30% (from 4.5 to 6.4 fF) leading to about 43% improvement in f_T (from 178 to 310 GHz) without a significant effect on f_{MAX}. In this case, the increase in C_{CB} is mainly due to an increase in both the intrinsic component C_{int} and the extrinsic component C_{SIC} since the implant is confined within W_{SIC}, which leads to a slight degradation in f_{MAX} by 15 GHz. In addition, no appreciable change in R_B was observed over the investigated SIC dose range. Recently, a similar technique has been implemented in the fabrication of InGaAs/InP double HBTs, where the subcollector was formed by a blanket Fe implant followed by a patterned Si implant to form the collector pedestal in order to reduce the extrinsic C_{CB}[17].

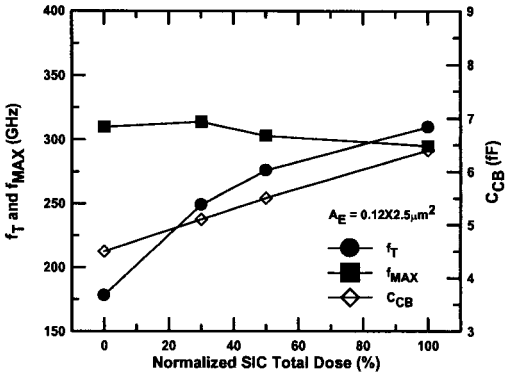

Fig. 5. C_{CB}, f_T, and f_{MAX} as a function of normalized SIC total dose.

The crystalline SiGe base layer is one of the most demanding parts of the HBT and technological challenges arise during "classical" device scaling. Vertical scaling of the SiGe base thickness and doping profile is achieved with modern low temperature UHV-CVD techniques, which enabled epitaxial growth of aggressively in-situ doped thin SiGe layers (below 50 nm) with excellent doping profile control[23-26]. Ultra-low deposition temperatures are required for a) thermal budget considerations during BiCMOS integration, b) dopant and alloy abruptness, and c) film and interface cleanliness. Typically, temperatures well below 600°C are necessary for a technology requiring a near-atomic dopant and alloy control. However, as thicknesses are reduced and dopant profiles become narrower, emitter and collector diffusion as well as base widening during subsequent activation and re-crystallization anneals severely limited further device scaling. The incorporation of carbon in the in-situ boron-doped SiGe base (i.e. SiGe:C) was found to suppress thermal boron diffusion caused by high thermal-budget processing and thus enabled continued and extendible base scaling[27]. A graded carbon doping in the base was implemented in fabricating InGaAs/InP double HBTs to reduce the base sheet and contact resistivities in order to improve f_{MAX}[10].

A typical SiGe base profile and scaling to improve f_T is shown in Fig. 6, where carbon profile is not shown. The total SiGe base thickness includes the neutral base layer and both emitter and collector intrinsic (here undoped) silicon layers (i.e. i-layer). The neutral SiGe base layer, containing the boron and carbon dopants, has a graded-Ge profile which creates a quasi electric field accelerating electrons across the base and hence improving f_T. The collector and emitter intrinsic layers allow dopant diffusion during the emitter thermal anneal to form the collector-base and emitter-base junctions. In addition, a sufficient collector i-layer thickness is critical for high quality epitaxial SiGe base layer growth otherwise degraded by residual contamination at the metallurgical collector-base interface. The neutral SiGe base width W_B determines the base transit time and requires optimization to improve f_T. For example, if W_B is reduced without adjusting the boron doping concentration, it may significantly increase R_B and degrade f_{MAX}. Therefore, efforts for minimizing W_B need to be carried out simultaneously with doping concentration increases, which is a challenging task. Two growth parameters are widely used to control W_B: diborane flow rate (i.e. boron concentration) and as-grown base width. The base profile in Fig. 6(b) was scaled down as compared to Fig. 6(a) to improve f_T by a) reducing the thicknesses of emitter i-layer, neutral base width, and

collector i-layer, b) adjusting boron dose and width, and c) increasing the Ge gradient for a higher built-in quasi-static drift field in the neutral base.

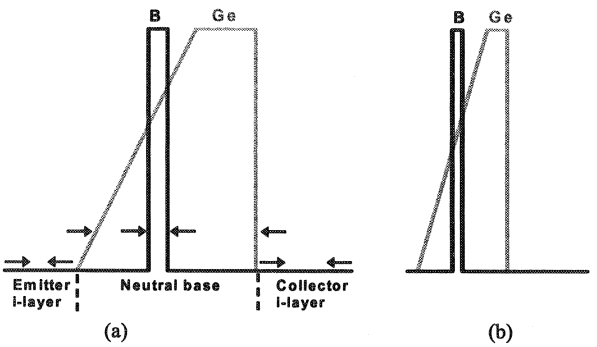

Fig. 6. Typical SiGe:C base doping profiles and scaling to improve f_T. Carbon profile is not shown.

The effect of normalized collector intrinsic layer thickness on C_{CB}, f_T, and f_{MAX} is shown in Fig. 7. The collector i-layer thickness modulates the distance between the SIC and the base layer. Thickening the collector i-layer would effectively decrease both the intrinsic and the extrinsic components of C_{CB}, unlike the case with reduced SIC total dose (Fig. 5), which affects only the intrinsic component C_{int} and the extrinsic component C_{SIC}. As a result, the increase in collector i-layer thickness would have inverse effect on f_T and f_{MAX}, which can be seen from Fig. 7. As the collector i-layer thickness is increased by 60%, C_{CB} decreases by about 20% (from 7.2 to 6 fF), with no appreciable change in R_B, leading to about 43 GHz (~27%) increase in f_{MAX} (from 159 to 202 GHz). On the other hand, such an increase in the collector i-layer thickness results in about 10 GHz decrease in f_T (from 341 to 331 GHz) caused by an increase in R_C (i.e. transit time).

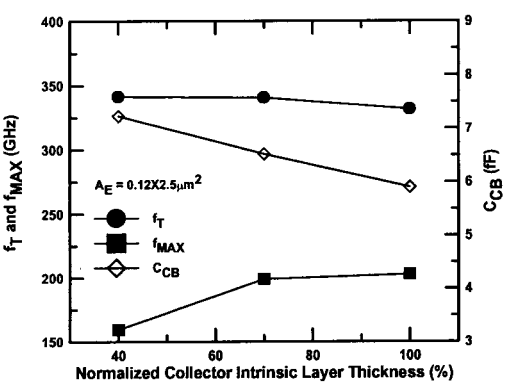

Fig. 7. C_{CB}, f_T, and f_{MAX} and as a function of normalized collector intrinsic layer thickness.

The effect of normalized as-grown base width W_B on R_B, f_T, and f_{MAX} is shown in Fig. 8. As W_B is increased by about 45%, R_B decreases by about 16% (from 42.2 to 35.4 Ω) leading to 20 GHz increase in f_{MAX} (from 306 to 326 GHz). In this case, the decrease in R_B is mainly due to a decrease in R_{int}, where the base current flows

laterally through a wider W_B towards the base contact. On the other hand, f_T decreases by 24 GHz (from 303 to 279 GHz) due to an increase in the base transit time caused by reduced Ge ramp slope for wider W_B. The peak Ge percentage in the base determines the Ge ramp slope (Fig. 6) and thus affects the quasi electric field and carrier transport across the base region. As the peak Ge percentage is increased for a nominal W_B, the electric field increases leading to a reduction in the base transit time and improvement in f_T. For example, increasing the peak Ge fraction by 28% resulted in a moderate increase in f_T by ~11 GHz whereas f_{MAX} degraded by ~13 GHz due to 20% increase in R_B. f_T and f_{MAX} can be simultaneously improved with W_B reduction accompanied with an increase in the boron doping concentration, which can be achieved by increasing the diborane flow during the SiGe base growth. A two-step Ge profile in the base and boron in-situ doping of the emitter i-layer was shown to increase f_T at low currents allowing ultra-low power operation of SiGe HBT at high speeds[28].

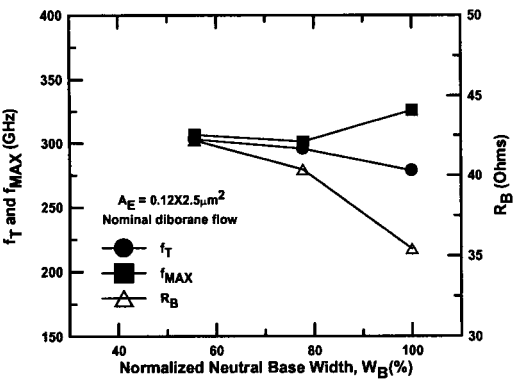

Fig. 8. R_B, f_T, and f_{MAX} as a function of normalized neutral base width (W_B).

The emitter resistance (R_E) is an additional important parameter that impacts f_T of SiGe HBTs. Recent studies have shown that a significant reduction in R_E as well as better controllability is achieved with in-situ doped emitter polysilicon compared to conventional implanted emitter polysilicon[29]. Phosphorous or arsenic in-situ doped polysilicon with high doping levels are implemented as emitter layers in modern SiGe HBT devices to improve f_T and control the device DC gain (β). Typical emitter polysilicon layer doping levels are on the order of 10^{20} cm^{-3}. The emitter polysilicon is usually deposited by rapid-thermal reduced-pressure CVD (RTCVD) for lowest thermal budget. Crystalline re-alignment, and dopant activation and out-diffusion from the emitter polysilicon are usually achieved with rapid thermal annealing (emitter RTA), in the range of 900-1000°C, which reduces R_E and forms the emitter-base junction. Cross-sectional TEM of a re-aligned emitter after activation anneal is shown in Fig. 9. Typical emitter layer sheet resistances are between 1-4 mΩ-cm. A metal emitter, made by complete silicidation of a mono-crystalline emitter (i.e. fully silicided emitter), was shown to reduce R_E and improve the device f_T and breakdown voltage[9].

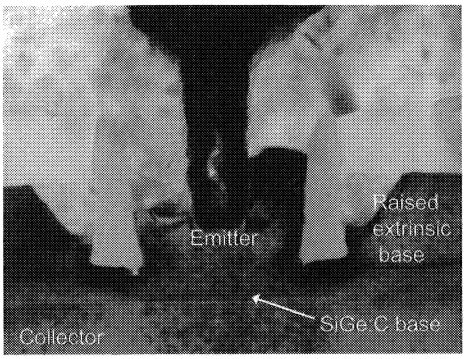

Fig. 9. Cross-sectional TEM of re-aligned emitter after activation anneal.

3.2. *Lateral Scaling and Device Structure Modifications*

SiGe HBT device performance can be further improved by lateral scaling and device structure modifications. Lithography techniques can be used, in a similar fashion implemented in CMOS technology, to reduce W_{SIC} and W_C thus minimizing C_{CB}. Furthermore, lithography and process techniques can be used to optimize the silicide-to-emitter spacing, D, and spacer width, W_S, to reduce R_B. In addition, the emitter width, W_E, in modern SiGe HBT technologies can be scaled down to sub-100 nm dimensions using advanced lithography techniques to reduce R_B and C_{CB} in order to improve the device speed.

As the SIC dimension W_{SIC} is reduced, C_{CB} decreases and R_C increases, while R_B remains approximately constant. However, the time delay product $R_C.C_{CB}$, which determines the device speed, was shown to decrease with SIC lateral dimension[30]. In this case, the reduction in W_{SIC} reduces the capacitance components C_{SIC}. As a result, f_T increases due to the decrease in delay time and f_{MAX} increases due to the reduction in C_{CB}. Furthermore, the time delay $R_C.C_{CB}$ could be optimized by lateral scaling of the collector width, W_C. The effect of W_C lateral scaling on f_T and f_{MAX} is shown in Fig. 10. As W_C is reduced from 0.72 μm to 0.40 μm, C_{CB} reduces by about 25% (from 12.2 to 9.15 fF) leading to an increase in f_T by 16 GHz (~5%) due to the decrease in delay time and an increase in f_{MAX} by 10 GHz (~6%). In this case, reduction in W_C reduces the capacitance component C_{link} and increases the component C_{STI}. However, C_{link} decrease is more effective in reducing total C_{CB}, since the contribution of C_{STI}, where the base and collector are separated by a thick oxide in the STI region, is negligible. A novel collector structure was shown to reduce C_{CB} by selectively undercutting the collector region to reduce the capacitance components C_{SIC} and C_{link}[4].

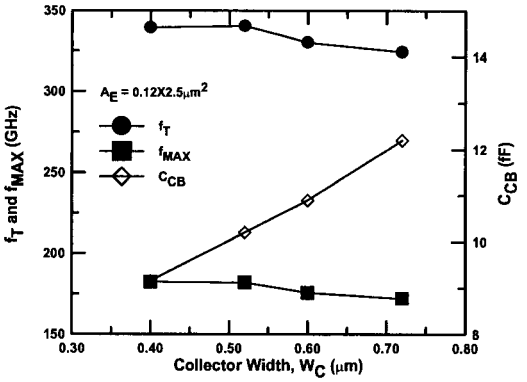

Fig. 10. f_T and f_{MAX} as a function of collector width (W_C).

State-of-the-art lithography tools employed in CMOS technology scaling to sub-100 nm dimensions are used to reduce the emitter dimension in SiGe HBTs to lower the device R_B and C_{CB}. The decrease in emitter (i.e. intrinsic device) area in effect reduces the intrinsic components R_{int} and C_{int}. In addition, reducing W_E leads to an increase in the link contact area between the base and collector leading to a decrease in R_{link} and a moderate increase in C_{link}. However, the decrease in C_{int} is more effective in reducing total C_{CB}, since the increase in C_{link} for the investigated W_E range is negligible. The effect of W_E scaling on R_B, C_{CB}, f_T, and f_{MAX} is shown in Fig. 11. Reducing W_E from 0.2 μm to 0.12 μm results in a reduction of about 9% and 12% in R_B and C_{CB}, respectively, which leads to a significant increase of 30 GHz in f_{MAX} and a moderate increase of 8 GHz in f_T. The emitter width could readily be reduced below 100 nm to further improve SiGe HBT performance. However, aggressive lateral reduction of W_E may increase R_E and R_C in the intrinsic device, which leads to f_T degradation[31]. This can be compensated for by increasing the doping levels in the emitter and collector.

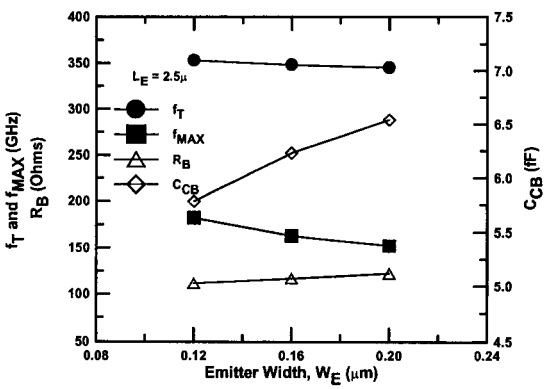

Fig. 11. R_B, C_{CB}, f_T, and f_{MAX}, as a function of emitter width (W_E).

Another technique to reduce the device parasitics is the optimization of the device layout. Two device layout configurations are shown in Fig. 12, where CBE and CBEBC represent the relative order of electrode contacts. SiGe HBT devices have traditionally

adopted the compact CBE configuration as the contact resistance does not limit performance because of the availability of silicide. However, as the device speed enters the hundreds-of-GHz operation regime, the effect of parasitic resistance and capacitance becomes more significant and the device layout needs to be considered for any performance enhancement. The CBEBC configuration improves f_T and f_{MAX}, compared to the CBE configuration, due to R_C reduction and the symmetric spread of injected electrons in the collector region as well as reduction of R_B component along the silicided region. The CBEBC configuration alone improved f_T and f_{MAX} by 25 GHz and 40 GHz, respectively[2].

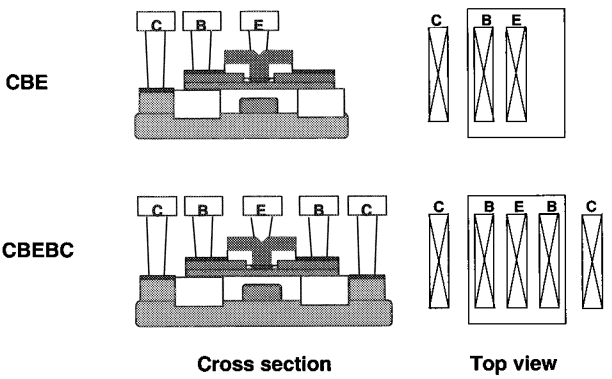

Fig. 12. Schematics of SiGe HBT device layout configurations: CBE and CBEBC.

3.3. *State-of-the-Art SiGe HBT Peroformance Path (IBM)*

Figure 13 shows the performance improvement through vertical scaling and device structure improvements for a SiGe HBT device with an emitter area A_E=0.12×2.5 μm^2. The first implementation of a self-aligned device structure with a raised extrinsic base and optimized vertical scaling[32] achieved f_T of 200 GHz and f_{MAX} of 285 GHz[7]. To improve f_T, the device vertical profile was scaled down by 1) collector vertical scaling, 2) SiGe base vertical scaling, which included thickness reduction, boron and Ge width reduction, and Ge gradient increase, 3) emitter activation anneal reduction. This resulted in an improvement in f_T to 280 GHz whereas f_{MAX} degraded to 170 GHz due to an increase in C_{CB} and R_B caused by collector vertical scaling and base layer thinning. To further improve f_T, additional collector vertical scaling to reduce R_C and an increase in peak Ge percentage (i.e. Ge gradient slope) were performed. Since these changes mainly affect the intrinsic device, f_T improved to 350 GHz while f_{MAX} remained approximately the same at 170 GHz[1]. Using the same vertical profile, the device layout was subsequently modified from CBE to CBEBC configuration, which improved f_T and f_{MAX} to 375 GHz and 210 GHz, respectively[2]. In an effort to achieve balanced f_T and f_{MAX}, the peak Ge percentage was decreased to reduce the quasi electric field in the base region in order to reduce f_T, while the base vertical profile was modified to reduce C_{CB} and improve f_{MAX}. In this case, a device with balanced f_T and f_{MAX} both of which exhibiting 300 GHz, was achieved[22]. Finally, the silicide-to-emitter spacing D was reduced to lower R_B (i.e. R_{poly}) in order to improve f_{MAX}. As a result, a device with f_{MAX} of 350 GHz, without affecting f_T of 300 GHz, was achieved[3]. f_T and f_{MAX} for this device are plotted in Fig. 14 as a function of collector current (I_C) and selected device parameters are summarized in

Table 1. As depicted in Fig. 13, vertical scaling results in a trade-off between the device performance and breakdown voltage as expected. Aggressive vertical scaling aimed to improve f_T and peak DC current gain (β) results in reduced collector-emitter breakdown voltage (BV_{CEO}) and collector-base breakdown voltage (BV_{CBO}).

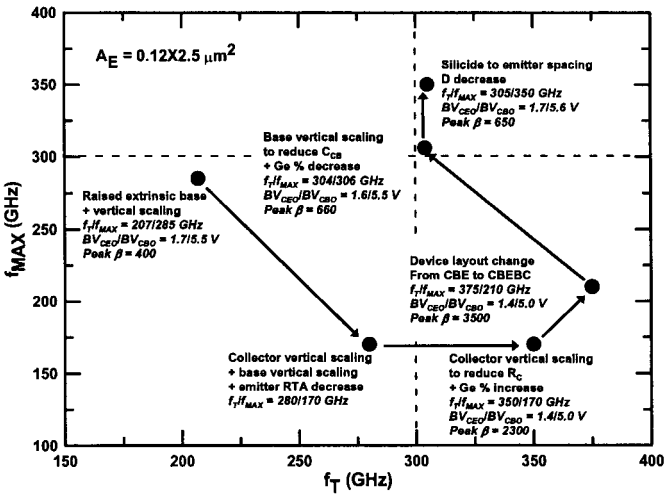

Fig. 13. State-of-the-art SiGe HBT technology performance path (IBM).

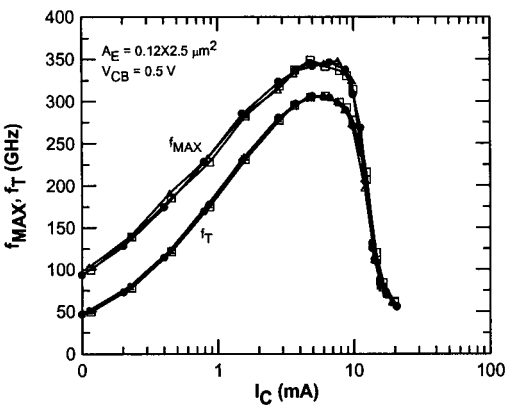

Fig. 14. f_{MAX} and f_T extrapolated from U and h_{21} at 40 GHz with -20 dB/dec slope.

Table 1. Selected device parameters of SiGe HBT with emitter area A_E=0.12×2.5 μm².

Parameter	Value
Peak f_{max}	350 GHz
Peak f_T	300 GHz
J_C @ peak f_{max} and f_T	19 mA/μm²
Peak β	650
BV_{CBO}	5.6 V
BV_{CEO}	1.7 V
BV_{EBO}	2.5 V

4. SiGe HBT Technology: Future Directions

SiGe HBT device performance can be further improved by vertical scaling, lateral scaling, and structure modifications enabled by modern process techniques and new materials developed for state-of-the-art CMOS technologies. In addition, technology simulation tools can be used to predict the impact of extended scaling and the implementation of novel materials on the device performance. In this section, we present simulation results to determine possible future directions to improve SiGe HBTs performance.

4.1. *SIC Implant Species*

Phosphorus (P) has been conventionally used as SIC implant species in modern SiGe HBTs due to its low activation thermal budget, which helps maintain high f_T. However, the high diffusion coefficient of phosphorus results in a significant lateral diffusion to the extrinsic device which leads to an increase in the extrinsic capacitance component C_{SIC}, hence degrading f_{MAX}. Other species with lower diffusion coefficient, such as arsenic (As) and antimony (Sb), can be implemented to reduce the lateral diffusion of SIC implant to maintain a lower C_{SIC}. Simulation results in Fig. 15 show the effect of phosphorus, arsenic, and antimony SIC implants on C_{CB}, f_T, and f_{MAX}. The total dose and anneal temperature were the same for all implant species. As can be seen from Fig. 15, C_{CB} is lower for As and Sb (~ 17%) compared to P leading to an increase of about 54 GHz in f_{MAX} (from 316 GHz to 370 GHz). However, a slight decrease in f_T is observed due to a decrease in the vertical diffusion of As and Sb for the same anneal temperature compared to P. We note that As and Sb SIC implants had insignificant effect on R_B.

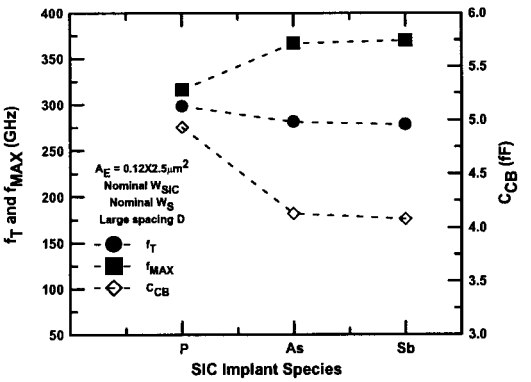

Fig. 15. Simulated C_{CB}, f_T, and f_{MAX}, for different SIC implant species: Phosphorus (P), Arsenic (As), and Antimony (Sb), for a device with large D.

4.2. *Lateral Scaling and Device Structure Modifications*

A further reduction in the extrinsic capacitance component C_{SIC} can be achieved by reducing the SIC implant width W_{SIC} using modern lithography techniques readily available for SiGe HBT technology. The simulation results shown in Fig. 16 demonstrate the effect of normalized W_{SIC} (nominal W_{SIC}=100%) scaling on C_{CB}, f_T, and f_{MAX} for a device with an emitter area A_E=0.12×2.5 μm^2, phosphorus SIC implant, nominal W_S, and

large spacing D. As can be seen from Fig. 16, a reduction of about 56% in W_{SIC} significantly reduces C_{SIC} and leads to about 28% decrease in C_{CB} (from 4.93 fF to 3.56 fF). As a result, f_{MAX} improves by 46 GHz (from 316 GHz to 362 GHz) while f_T decreases by about 9 GHz, which is believed to be due to an increase in R_C[30]. We note that R_B remained approximately constant with W_{SIC} reduction.

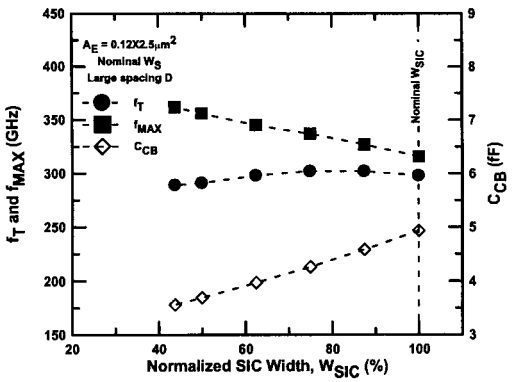

Fig. 16. Simulated C_{CB}, f_T, and f_{MAX} as a function of normalized SIC width (W_{SIC}) for a device with large D.

A significant lateral scaling parameter in SiGe HBTs is the reduction in the emitter width W_E, which can be scaled down to sub-100 nm dimensions using state-of-the-art lithography tools used in CMOS technologies. Simulation results plotted in Fig. 17 show the effect of W_E scaling on R_B, f_T, and f_{MAX} for a device with an emitter length L_E=2.5 μm, phosphorus SIC implant of nominal W_{SIC}, nominal W_S, and large spacing D. Reducing W_E from 0.12 μm to 0.07 μm results in about 33% reduction in R_B (from 43.3 Ω to 29 Ω) caused mainly due to a significant reduction in R_{int} and R_{link}. However, simulation results did not predict a reduction in C_{CB}, as expected from experimental results shown in Fig. 11, which remained about 4.9 fF for all simulated values of W_E. Nevertheless, the decrease in R_B with W_E results in a significant increase of 114 GHz in f_{MAX} (from 314 GHz to 428 GHz) and a moderate decrease in f_T by 11 GHz.

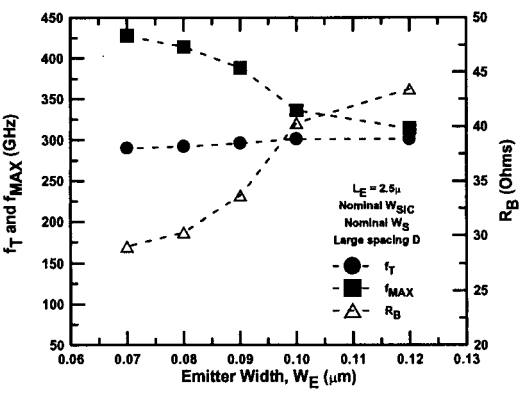

Fig. 17. Simulated R_B, f_T, and f_{MAX} as a function of emitter width (W_E) for a device with large D.

The extrinsic base components R_{spr} and R_{link} can be minimized to further reduce R_B by reducing the spacer width W_S. A narrower spacer in effect reduces the width of the lightly doped base region below the spacer, which reduces R_{spr}, and increases the overlap link area between the heavily doped polysilicon extrinsic base and the intrinsic base, which reduces R_{link}. Simulation results in Fig. 18 demonstrate the effect of normalized W_S (nominal W_S=100%) on R_B, f_T, and f_{MAX} for a device with an emitter area A_E=0.10×2.5 μm^2, phosphorus SIC implant of nominal W_{SIC}, and large spacing D. As can be seen in Fig. 18, R_B decreases significantly with W_S, which is a major advantage of a self-aligned device structure. As W_S is reduced by about 45% below its nominal value, R_B decreases by about 43% (from 40.3 Ω to 23.1 Ω), which results in an increase of 68 GHz in f_{MAX} from 336 GHz to 404 GHz (~20%). On the other hand, when W_S is increased by about 45% above its nominal value, R_B increases slightly by about 18% (from 40.3 Ω to 48.1 Ω), resulting in a small increase of 23 GHz in f_{MAX} (~7%). This indicates that for a large W_S, the extrinsic base resistance component R_{spr} becomes less effective in R_B optimization, in a device with a large spacing D, compared to the other extrinsic base components (e.g. R_{poly}).

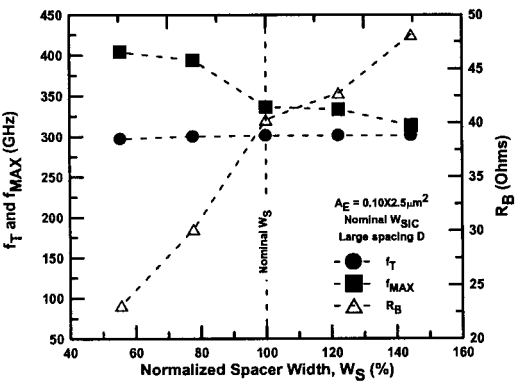

Fig. 18. Simulated R_B, f_T, and f_{MAX} as a function of spacer width (W_S) for a device with large D.

To reduce the extrinsic resistance component R_{poly}, the width of the un-silicided portion of the extrinsic base polysilicon can be minimized by reducing the silicide-to-emitter spacing D. In this case, the current spreads through a smaller portion of un-silicided extrinsic base polysilicon to reach the silicide edge, which effectively reduces R_B and improves f_{MAX}. Such a structure modification of a device with an emitter area A_E=0.12×2.5 μm^2 resulted in about 14% reduction in R_B and a significant increase of 50 GHz in f_{MAX} (from 300 to 350 GHz)[3], which is also predicted by simulation as shown in Fig. 19. Also shown is the effect of W_E and W_S scaling on R_B, f_T, and f_{MAX} for a device with small spacing D. As can be seen from Fig.19, a reduction in W_E from 0.12 μm to 0.10 μm results in a moderate decrease in R_B of about 6% (from 34.5 Ω to 32.5 Ω) and a slight increase of about 27 GHz in f_{MAX} (from 348 GHz to 375 GHz). However, as W_S is reduced by 45% below its nominal value for a device with an emitter area A_E=0.10×2.5 μm^2, R_B decreases significantly by about 55% (from 32.5 Ω to 14.5 Ω). This indicates the effectiveness of W_S reduction in optimizing R_B for a device with small spacing D, where

the contribution of the extrinsic resistance component R_{poly} becomes negligible. As a result, f_{MAX} increases by 65 GHz (from 375 GHz to 440 GHz).

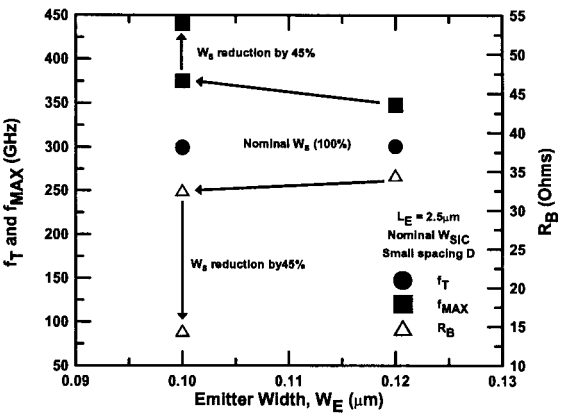

Fig. 19. Simulated R_B, f_T, and f_{MAX} as a function of emitter width (W_E) and spacer width (W_S) for a device with small D.

The extrinsic resistance component R_{sil} is determined by the silicide sheet resistance. Cobalt silicide has conventionally been used for SiGe HBT as a low-resistance ohmic contact to the base and collector. However, two high-temperature annealing steps in the range of 500°C-800°C are required to form the lowest resistance phase of cobalt disilicide $CoSi_2$. Such a high thermal budget causes a significant increase in R_B and C_{CB} due to dopants diffusion in the base and collector, which leads to adegradation in both f_T and f_{MAX}. Recent developments in nickel-based silicide implementation in CMOS technology can also be utilized to improve SiGe HBT device performance. The low-resistance mono-silicide phase (e.g. NiSi and NiPtSi) can be formed in one annealing step at low temperatures in the range 400°C-700°C, thus reducing the process thermal budget and dopant diffusion[33,34]. In addition, the lower resistivity and silicon consumption compared to cobalt silicide, is useful in SiGe HBTs to reduce R_B (i.e. R_{sil} component) and improve f_{MAX}.

The advantages of R_{poly} and R_{sil} optimization, however, can be limited by the silicide-to-polysilicon contact resistance. The extrinsic resistance component $R_{sc}=\rho_c/A$ is determined by the silicide contact resitivity (ρ_c) and contact area between the silicide and polysilicon (A). The contact resistivity is determined by the doping level at the silicide/polysilicon interface, where high dopant concentration is critical to ensure a low contact resistance. Nickel-based silicides offer lower ρ_c due to lower thermal budget for NiSi and NiPtSi formation[33,34]. Recent results suggested that the contact resistance of NiSi on p-doped substrates is lower than that of $CoSi_2$[35,36], which can be beneficial to reduce R_B (i.e. R_{sc}) and improve f_{MAX} in SiGe HBTs.

4.3. *Extended SiGe HBT Technology Performance Path: Example Case*

Based on the simulation results, few options were selected as an example case, to extend the performance path of SiGe HBT technology. More specifically, as shown in Fig. 20, three options were added by simulation to the performance path in Fig. 13, which include

a) W_E reduction from 0.12 μm to 0.10 μm, b) W_{SIC} reduction by 37% of nominal width, and c) W_S reduction by 45% of nominal width. As a result, as illustrated in Fig. 20, f_{MAX} could be improved to 490 GHz with no significant impact on f_T, which remains about 300 GHz. In addition, further vertical scaling, lateral scaling, and device structure improvements, as described in the previous sections, are expected to enable the evolution of SiGe HBTs towards operational speeds in the THz regime.

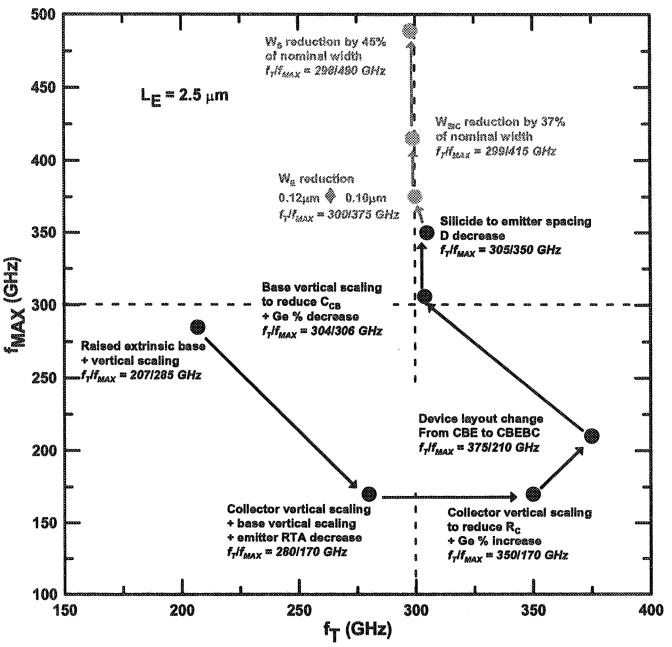

Fig. 19. Extended SiGe HBT technology performance path.

5. Summary

We presented a review of recent developments in SiGe HBT technology that led to a significant improvement in the device performance. High-performance SiGe HBTs operating at speeds approaching 400 GHz have been achieved by vertical scaling, lateral scaling, and device structure innovations enabled by modern CMOS-compatible materials and process techniques. We also presented device simulation results that showed the extendibility of the performance of SiGe HBTs towards half-terahertz and beyond.

6. Acknowledgments

The authors would like to acknowledge partial support of this work by DARPA under SPAWAR contract number N66001-02-C-8014. The authors also would like to thank Joseph Kocis, David Rockwell, Michael Longstreet, Karyn Hurley, and Robert Groves for their support in device process and test.

References

1. J.-S. Rieh, B. Jagannathan, H. Chen, K. Schonenberg, D. Angell, A. Chinthakindi, J. Florkey, F. Golan, D. Greenberg, S. –J. Jeng, M. Khater, F. Pagette, C. Schnabel, P. Smith, A. Stricker, K. Vaed, R. Volant, D. Ahlgren, G. Freeman, K. Stein, and S. Subbanna, SiGe HBTs with cut-off frequency of 350 GHz, in *IEDM Tech. Dig.*, 771-774 (2002).
2. J.-S. Rieh B. Jagannathan, H. Chen K. Schonenberg, S. –J. Jeng, M. Khater, D. Ahlgren, G. Freeman, and S. Subbanna, Performance and design considerations for high speed SiGe HBTs of f_T/f_{MAX} = 375 GHz/210 GHz, in *Proc. of International Conf. on Indium Phosphide and Related Materials*, 374-377 (2003).
3. M. Khater, J. –S. Rieh, T. Adam, A. Chinthakindi, J. Johnson, R. Krishnasamy, M. Meghelli, F. Pagette, D. Sanderson, C. Schnabel, K. Schonenberg, P. Smith, K. Stein, A. Stricker, S. –J. Jeng, D. Ahlgren, and G. Freeman, SiGe HBT technology with f_{MAX}/f_T= 350/300 GHz and gate delay below 3.3 ps, in *IEDM Tech. Dig.*, 247-250 (2004).
4. B. Heinemann, R. Barth, D. Bolze, J. Drews, P. Formanek, T. Grabolla, U. Haak, W. Hoppner, D. Knoll, K. Kopke, B. Kuck, R. Kurps, S. Marschmeyer, H. Richter, H. Rucker, P. Schley, D. Schmidt, W. Winkler, D. Wolansky, H. –E. Wulf, and Y. Yamamoto, A low-parasitic collector construction for high-speed SiGe:C HBTs, in *IEDM Tech. Dig.*, 251-254 (2004).
5. J. Bock, H. Schafer, H. Knapp, K. Aufinger, M. Wurzer, S. Boguth, T. Bottner, R. Stengl, W. Perndl, and T. Meister, 3.3 ps SiGe bipolar technology, in *IEDM Tech. Dig.*, 255-258 (2004).
6. B. Floyd, S. Reynolds, U. Pfeiffer, T. Zwick, T. Beukema, and B. Gaucher, SiGe bipolar transceiver circuits operating at 60 GHz, *IEEE J. Solid-State Circuits* **40**(1), 156-167 (2005)
7. B. Jagannathan, M. Khater, F. Pagette, J. –S. Rieh, D. Angell, H. Chen, J. Florkey, F. Golan, D. Greenberg, R. Groves, S. –J. Jeng, J. Johnson, E. Mengistu, K. Schonenberg, C. Schnabel, P. Smith, A. Stricker, D. Ahlgren, G. Freeman, K. Stein, and S. Subbanna, Self-aligned SiGe NPN transistors with 285 GHz f_{MAX} and 207 GHz f_T in a manufacturable technology, *IEEE Elec. Dev. Lett.* **23**(5), 258-260 (2002).
8. K. Washio, SiGe HBT and BiCMOS technologies, in *IEDM Tech. Dig.*, 113-116 (2003)
9. J. Donkers, T. Vanhoucke, P. Agarwal, R. Hueting, P. Meunier-Beillard, M. Vijayaraghavan, P. Magnee, M. Verheijen, R. De Kort, and J. Slotboom, Metal emitter SiGe:C HBTs, in *IEDM Tech. Dig.*, 243-246 (2004).
10. M. Dahlstrom, X. –M. Fang, D. Lubyshev, M. Urteaga, S. Krishnan, N. Parthasarathy, Y. M. Kim, Y. Wu, J. M. Fastenau, W. K. Liu, M. Rodwell, Wideband DHBTs using a graded carbon-doped InGaAs base, *IEEE Elec. Dev. Lett.* **24**(7), 433-435 (2003).
11. W. Hafez and M. Feng, 0.25 μm Emitter InP SHBTs with f_T= 550 GHz and BV_{CEO} > 2 V, in *IEDM Tech. Dig.*, 549-552 (2004).
12. T. Hussain, Y. Royter, D. Hitko, M. Montes, M. Madhav, I. Milosavljevic, R. Rajavel, S. Thomas, M. Antcliffe, A. Arthur, Y. Boegeman, M. Sokolich, J. Lee, and P. Asbeck, First demonstration of sub-0.25 μm width emitter InP DHBTs with > 400 GHz f_T and > 400 GHz f_{MAX}, in *IEDM Tech. Dig.*, 553-556 (2004).
13. D. Yu, K. Choi, K. Lee, B. Kim, H. Zhu, K. Vargason, J. M. Kuo, P. Pinsukanjana, and Y. Kao, Ultra high-speed 0.25 μm emitter InP-InGaAs SHBTs with f_{MAX} of 687 GHz, in *IEDM Tech. Dig.*, 557-560 (2004).
14. Z. Griffith, M. Dahlstrom, M. Rodwell, X. –M. Fang, D. Lubyshev, Y. Wu, J. Fastenau, and A. Liu, InGaAs-InP DHBTs for increased digital IC bandwidth having 391 GHz f_T and 505 GHz f_{MAX}, *IEEE Elec. Dev. Lett.* **26**(1), 11-13 (2005).
15. Z. Griffith, M. Rodwell, X. –M. Fang, D. Lubyshev, Y. Wu, J. Fastenau, and A. Liu, InGaAs/InP DHBTs with 120 nm collector having simultaneously high f_T and $f_{MAX} \geq$ 450 GHz, *IEEE Elec. Dev. Lett.* **26**(8), 530-532 (2005).

16. W. Snodgrass, B. –R. Wu, W. Hafez, K. –Y. Cheng, and M. Feng, Graded base type-II InP/GaAsSb DHBT with f_T = 475 GHz, *IEEE Elec. Dev. Lett.* **27**(2), 84-86 (2006).
17. N. Parthasarathy, Z. Griffith, C. Kadow, U. Singisetti, M. Rodwell, X. –M. Fang, D. Loubychev, Y. Wu, J. Fastenau, and A. Liu, Collector-pedestal InGaAs/InP DHBTs fabricated in a single-growth, triple-implant process, *IEEE Elec. Dev. Lett.* **27**(5), 313-316 (2006).
18. R. Krithivasan, Y. Lu, J. Cressler, J. –S. Rieh, M. Khater, D. Ahlgren, and G. Freeman, Half-Terahertz operation of SiGe HBTs, *IEEE Elec. Dev. Lett.* **27**(7), 567-569 (2006).
19. S. -J. Jeng, D. Greenberg, M. Longstreet, G. Hueckel, D. Harame, and D. Jadus, Lateral scaling of the self-aligned extrinsic base in SiGe HBTs, in *Proc. BCTM*, 15-18 (1996).
20. S. -J. Jeng, D. Ahlgren, G. Berg, B. Ebersman, G. Freeman, D. Greenberg, J. Malinowski, D. Nguyen-Ngoc, K. Schonenberg, K. Stein, D. Colavito, M. Longstreet, P. Ronsheim, S. Subbanna, and D. Harame, Impact of extrinsic base process on NPN HBT performance and polysilicon resistor in integrated SiGe HBTs, in *Proc. BCTM*, 187-190 (1997).
21. M. Hashim, R. Lever, and, P. Ashburn, Two dimensional simulation of transient enhanced boron out-diffusion from the base of a SiGe HBT due to an extrinsic base implant, in *Proc. BCTM*, 96-99 (1997).
22. J. -S. Rieh, D. Greenberg, M. Khater, K. Schonenberg, S. -J. Jeng, F. Pagette, T. Adam, A. Chinthakindi, J. Florkey, B. Jagannathan, J. Johnson, R. Krishnasamy, D. Sanderson, C. Schnabel, P. Smith, A. Stricker, S. Sweeney, K. Vaed, T. Yanagisawa, D. Ahlgren, K. Stein, and G. Freeman, SiGe HBTs for millimeter-wave applications with simultaneously optimized f_T and f_{MAX} of 300 GHz, in *IEEE RFIC Symp. Dig.*, 395-398 (2004).
23. B. Meyerson, Low-temperature silicon epitaxy by ultrahigh vacuum/chemical vapor deposition, *Appl. Phys. Lett.* **48**(12), 797-799 (1986).
24. B. Meyerson, UHV/CVD growth of Si and Si:Ge alloys: chemistry, physics, and device applications, *Proc. IEEE* **80**(10), 1592-1608 (1992).
25. D. Harame, J. Comfort, J. Cressler, E. Crabbe, J. Sun, B. Meyerson, and T. Tice, Si/SiGe epitaxial-base transistors-part I: materials, physics, and circuits, *IEEE Trans. Elec. Dev.* **42**(3), 455-468 (1995).
26. D. Harame, J. Comfort, J. Cressler, E. Crabbe, J. Sun, B. Meyerson, and T. Tice, Si/SiGe epitaxial-base transistors-part II: process integration and analog applications, *IEEE Trans. Elec. Dev.* **42**(3), 469-482 (1995).
27. L. Lanzerotti, J. Sturm, E. Stach, R. Hull, T. Buyuklimanli, and C. Magee, Suppression of boron outdiffusion in SiGe HBTs by carbon incorporation, in *IEDM Tech. Dig.*, 249-252 (1996).
28. M. Xu, S. Decoutere, A. Sibaja-Hernandez, K. Van Wichelen, L Witters, R. Loo, E. Kunnen, C. Knorr, A. Sadovnikov, and C. Bulucea, Ultra low power SiGe:C HBT for 0.18 μm RF-BiCMOS, in *IEDM Tech. Dig.*, 125-128 (2003).
29. A. Joseph, P. Geiss, X. Liu, J. Johnson, K. Schonenberg, A. Chakravarti, D. ahlgren, and J. Dunn, Emitter resistance improvement in high-performance SiGe HBTs, in *Proc. ISTDM*, 53-54 (2003).
30. A. Stricker, G. Freeman, M. Khater, and J. –S. Rieh, Evaluating and designing the optimal 2D collector profile for a 300 GHz SiGe HBT, *Mat. Sci. Semi. Proc.* **8**, 295-299 (2005).
31. J. -S. Rieh, D. Greenberg, A. Stricker, and G. Freeman, Scaling of SiGe heterjunction bipolar transistors, *Proc. IEEE* **93**(9), 1522-1538 (2005).
32. S. -J. Jeng, B. Jagannathan, J. -S. Rieh, J. Johnson, K Schonenberg, D. Greenberg, A. Stricker, H. Chen, M. Khater, D. Ahlgren, G. Freeman, K. Stein, and S. Subbanna, A 210 GHz f_T SiGe HBT with a non-self-aligned structure, *IEEE Elec. Dev. Lett.* **22**(11), 542-544 (2001).
33. T. Morimoto, T. Ohguro, H. Momose, T. Iinuma, I. Kunishima, K. Suguro, I. Katakabe, H. Nakajima, M. Tsuchiaki, M. Ono, Y. Katsumata, and H. Iwai, Self-aligned nickel-mono-silicide

technology for high-speed deep submicron logic CMOS ULSI, *IEEE Trans. Elec. Dev.* **42**(5), 915-922 (1995).

34. P. Lee, K. Pey, D. Mangelinck, J. Ding, D. Chi, and L. Chan, New salicidation technology with Ni(Pt) alloy for MOSFETs, *IEEE Elec. Dev. Lett.* **22**(12), 568-570 (2001).
35. J. Kittl, A. Lauwers, O. Chamirian, M. Van Dal, A. Akheyar, M. De Potter, R. Lindsay, and K. Maex, Ni- and Co-based silicides for advanced CMOS applications, *Microelec. Eng.* **70**(2-4), 158-165 (2003).
36. J. A. Kittl, A. Lauwers, O. Chamirian, M. Van Dal, A. Akheyar, O. Richard, J. Lisoni, M. De Potter, R. Lindsay, and K. Maex, Silicides for 65 nm CMOS and beyond, in *Mat. Res. Soc. Symp. Proc.* **765**, 267-278 (2003).