InP Bipolar ICs: Scaling Roadmaps, Frequency Limits, Manufacturable Technologies

Doubling the bandwidth of phosphide heterojunction bipolar integrated circuits, used for imaging, radio astronomy and spectroscopy, appears to be feasible.

By Mark J. W. Rodwell, *Fellow IEEE*, Minh Le, *Member IEEE*, and Berinder Brar, *Member IEEE*

ABSTRACT | Indium phosphide heterojunction bipolar transistors (HBTs) find applications in very wide-band digital and mixed-signal integrated circuits (ICs). Devices fabricated in high-yield process flows at 500 nm feature size obtain ~450 GHz cutoff frequencies and ~5 V breakdown and enable high yield fabrication of integrated circuits having more than 3000 transistors. Laboratory devices at 250 nm feature size obtain 755 GHz $f_{\max}$. We describe device and circuit bandwidth limits associated with HBTs, develop scaling roadmaps for HBTs having lithographic minimum feature sizes between 512 and 64 nm, and identify key technological challenges in realizing 480-GHz digital ICs and 1000-GHz amplifiers. Key features of manufacturable self-aligned dielectric sidewall processes are described in detail.

KEYWORDS | Heterojunction bipolar transistors; millimeter-wave circuits; millimeter-wave devices; submillimeter-wave circuits; submillimeter-wave devices

I. INTRODUCTION

Compared to SiGe bipolar transistors and Si metal–oxide–semiconductor field-effect transistors, InP double hetero-junction bipolar transistors (DHBTs) [1]–[6] attain higher bandwidth at a given lithographic feature size and attain higher device breakdown voltage at a given device bandwidth. These advantages result from the high electron mobility of the InGaAs base [7], [8], the high base doping made feasible by strong emitter-base heterojunctions [9], and the high peak electron velocity and high breakdown field of the InP collector [10], [11]. InP HBTs find application in medium-scale integrated circuits (ICs) requiring wide bandwidth and high dynamic range, specifically gigahertz mixed-signal ICs.

We review recent rapid progress [12]–[18] in transistor bandwidth, circuit bandwidth, and large-scale IC fabrication. InP DHBTs at ~500 nm emitter feature size have attained ~450 GHz balanced ($f_\tau \cong f_{\max}$) cutoff frequencies and have enabled both 178-GHz amplifiers [14] and 150-GHz digital circuits [13], [14]. Three-thousand-transistor ICs have been fabricated [19], a scale sufficient for analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) having sampling rates in the microwave frequency range. These results have been obtained through scaling, through transistor design for digital circuit bandwidth, and through substantial innovations in the IC fabrication processes.

We will also here address the requirements—and our present plans—for a further doubling of device and circuit bandwidth, thus enabling terahertz transistor cutoff frequencies, 600-GHz wireless transceivers, and small digital ICs operating at 300 GHz. Finally, we will explore, to our present understanding, the frequency limits of bipolar integrated circuits.

Wide-band HBTs will enable increased resolution and bandwidth in gigahertz mixed-signal ICs, including ADCs, DACs, and microwave direct digital frequency synthesizers (DDSs). They will enable 100-GHz gain-bandwidth product operational amplifiers for low-distortion amplification at ~1–10 GHz, will enable 160-Gb/s [20] and 320-Gb/s optical-fiber links, and will enable sub-mm-wave (300 GHz–1 THz) ICs for imaging, radio astronomy, and spectroscopy.

M. J. W. Rodwell is with the Electrical and Computer Engineering Department, University of California, Santa Barbara, CA 93106 USA (e-mail: rodwell@ece.ucsb.edu).
M. Le was with Vitesse Semiconductor, Camarillo, CA 93102 USA. He is now with Evergreen Solar, Marlboro, MA 01752 USA (e-mail: minhsyle@gmail.com).
B. Brar is with Teledyne Scientific (formerly Rockwell Scientific), Thousand Oaks, CA 91360 USA (e-mail: bbrar@teledyne.com).

II. BIPOLAR TRANSISTOR DESIGN

Initial research [21], [22] regarding III–V HBTs identified semiconductors with the highest available carrier mobilities and high-field velocities and quantified the role of transient transport [1], [10]; further improvements are now obtained by *scaling* [6]. After first identifying transistor performance metrics, we develop HBT scaling laws, propose a roadmap extending through the 64 nm (1.5 THz) generation, and identify key difficulties in reaching this performance.

A. Performance Metrics and Goals

Although widely reported, the short-circuit current gain (f_τ) and power gain ($f_{\max}$) cutoff frequencies are of limited value in predicting the bandwidth of mixed-signal ICs. Transistors designed exclusively for high f_τ at the expense of other parameters will perform poorly in circuits.

For reactively tuned amplifiers used in radio transceivers, $f_{\max}$ defines the highest frequency at which power gain can be obtained, and therefore defines the transistor's useful frequency range. f_τ is of secondary importance, as $f_\tau : f_{\max}$ ratios below ~1 : 2 make impedance matching difficult. Practical radio-frequency (RF) amplifiers are tuned for minimum noise figure or for maximum efficiency and maximum saturated output power. Expressions for these parameters are complex, and we use instead $f_{\max}$ as a performance metric, assuming that useful gain is obtained at $f_{\text{signal}} \leq f_{\max}/2$.

In HBT mixed-signal ICs, many analog and digital blocks are constructed from differential pairs and Gilbert cells. An emitter-coupled logic (ECL) master–slave (M–S) latch [Fig. 1(a)] is a representative small circuit and a digital performance benchmark. When configured as a 2 : 1 frequency divider, its maximum clock frequency $f_{\text{clk,max}} = 1/2T_{\text{gate}}$ is approximated by [6], [23], [24]

$$T_{\text{gate}} \approx (\Delta V_L/I_C)(C_{je} + 6C_{cbx} + 6C_{cbi}) + \tau_f$$
$$+ R_{ex}(0.5C_{cbx} + 0.5C_{cbi} + 0.5\tau_f I_C/\Delta V_L)$$
$$+ R_{bb}(0.5C_{je} + C_{cbi} + 0.5\tau_f I_C/\Delta V_L). \quad (1)$$
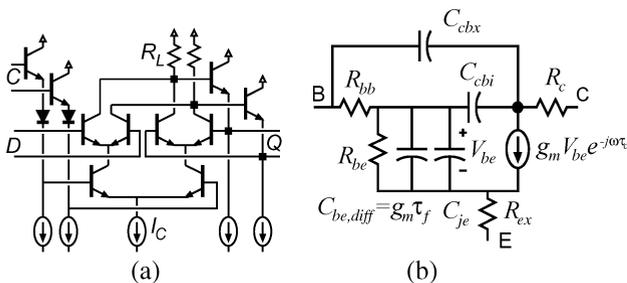
Fig. 1. *(a) ECL latch. A master–slave latch is formed from two cascaded latches with opposite clock phases. C = clock, D = data, Q = latch output. (b) Simplified HBT small-signal equivalent circuit.*

$\Delta V_L = I_C R_L$ is the logic voltage swing, R_L the load resistance, I_C the collector current, R_{ex} and R_{bb} the emitter and base parasitic resistances, C_{je} the emitter depletion capacitance, C_{cbx} and C_{cbi} the components of the collector-base capacitance C_{cb} external to and internal to R_{bb} [6], [25], and τ_f the sum of base τ_b and collector τ_c transit times [Fig. 1(b)]. The large 6 : 1 multipliers of $C_{cb}\Delta V_L/I_C$ in (1) result from Miller effect, fan-in, and fan-out.

For adequate logic noise margin, $\Delta V_L = I_C R_L$ must be at least $\sim 4(kT/q + R_{ex}I_C)$, and is typically $\sim 10 \cdot kT/q$. Comparing expressions for gate delay (1) and f_τ

$$1/2\pi f_\tau = (kT/qI_c)(C_{je} + C_{cbx} + C_{cbi})$$
$$+ \tau_f + (R_{ex} + R_c)(C_{cbx} + C_{cbi}) \quad (2)$$

it is seen that as a consequence of large-signal operation ($\Delta V_L \gg kT/q$), fan-in, and fan-out, depletion capacitance charging times ($C\Delta V_L/I_C$, $C \cdot kT/qI_C$) have much stronger impact upon T_{gate} than upon f_τ, and dominate gate delay. Gate delay correlates well with neither f_τ nor $f_{\max}$; instead, R_{ex} must be minimized to permit small ΔV_L, and low $(C_{je} + C_{cb})/I_C$ is imperative.

B. Scaling Laws and Scaling Roadmaps

We develop HBT scaling laws [6], extending earlier analyses of Si devices [26]–[28]. For clarity, the treatment is approximate, modeling transport by the drift-diffusion equation and neglecting contact pad parasitics and metal resistances. Reference [6] provides more detail. We assume the dimensions, cross-section, and layer structure of Figs. 2 and 3.

In scaling, we seek to increase by $\gamma : 1$ the bandwidth of an arbitrary circuit using the transistor, requiring a $\gamma : 1$ reduction of all capacitances and transit delays while maintaining constant all parasitic resistances, all bias and signal voltages, the transconductance g_m, and the operating current I_C.

The collector transit time is $\tau_c \sim T_c/2v_{\text{eff}}$ and the base transit time $\tau_b \sim T_b^2/2D_n + T_b/v_{\text{eff}}$ [6], [10], where v_{eff} is the collector effective high-field velocity and D_n the electron diffusivity in the base. Reducing T_c by $\gamma : 1$ and T_b by slightly more than $\gamma^{1/2} : 1$ will reduce these transit delays in the desired proportion.

Reducing T_c by $\gamma : 1$ would *increase* C_{cb} if junction areas were held constant. Reducing the emitter and junction areas in proportion to $\gamma^2 : 1$ then results in the desired reduction in C_{cb}. With constant I_C but with the emitter junction area reduced in proportion to $\gamma^2 : 1$, the emitter current density necessarily increases: $J_E \propto \gamma^2$. This is feasible within the limits imposed by the Kirk effect, as $J_{\text{Kirk}} \propto 1/T_C^2 \propto \gamma^2$. Note that C/I ratios are reduced by *reducing* T_C.

The base-emitter depletion layer must also be thinned to avoid both series resistance effects and space-charge storage effects [6].
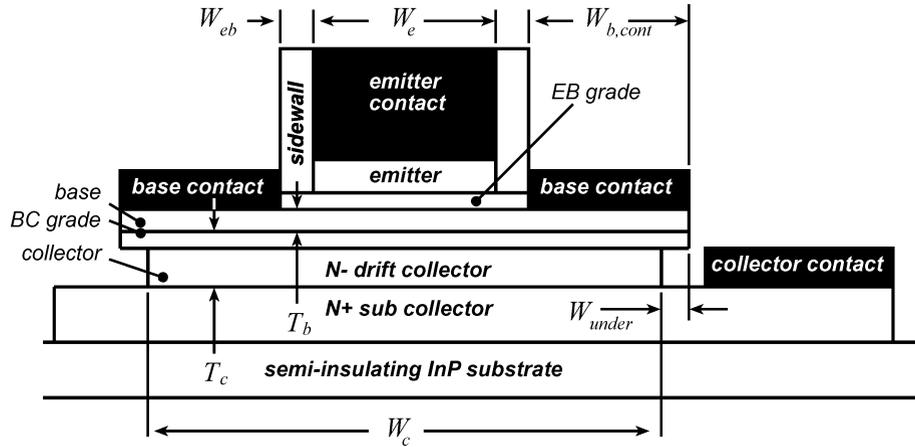
Fig. 2. *HBT cross-section and critical dimensions. The emitter stripe extends a distance L_e perpendicular to the figure. W_e is the emitter junction width, W_{eb} the base-emitter sidewall spacer thickness, $W_{b,\mathrm{cont}}$ the width of the base ohmic contact, and W_{under} the undercut of the collector junction under the base contacts.*

Because the emitter parasitic resistance R_{ex} must remain constant in the presence of an emitter junction area $A_E = L_E W_E \propto \gamma^{-2}$, the normalized emitter access resistivity $\rho_{ex} = R_{ex} A_E$ must be reduced rapidly, with $\rho_{ex} \propto \gamma^{-2}$. Note that ρ_{ex} includes both contact resistivity and bulk resistivity of the N+ emitter cap layer.

Emitter and collector junction areas can be reduced by reducing either junction widths or lengths. The variation of base resistance and thermal resistance with HBT geometry will determine this choice.

Consider the base resistance R_{bb} of a typical HBT with negligible base-collector undercut $W_{\mathrm{under}} \ll W_{b,\mathrm{cont}}$. Here R_{bb} is the sum of spreading resistance $R_{\mathrm{spread}} = \rho_s W_e / 12 L_e$, link resistance $R_{\mathrm{link}} = \rho_s W_{eb} / 2 L_e$, and contact access resistance $R_{\mathrm{contact}} = (\rho_s^{1/2} \rho_{b,v}^{1/2} / 2 L_E) \cdot \coth(W_{b,\mathrm{cont}}/L_c)$. Here ρ_s is the base sheet resistance, $\rho_{b,v}$ the specific (vertical)

contact resistivity, and $L_c = (\rho_{b,v}/\rho_s)^{1/2}$ the transfer length. The hyperbolic function is well approximated as $\coth(x) \cong 1/x + x/3$ for $x < 2$; hence

$$R_{bb} \cong \frac{\rho_{b,v}}{2} \frac{1}{L_e W_E} \frac{W_E}{W_{b,\mathrm{cont}}} + \frac{\rho_s}{2} \frac{W_E}{L_e} \left(\frac{1}{3} \frac{W_{b,\mathrm{cont}}}{W_E} + \frac{W_{eb}}{W_E} + \frac{1}{6} \right). \tag{3}$$

Assume all width ratios in (3) remain fixed. From the first (contact resistivity) term in (3), the scaling law for the base contact resistivity is obtained: because $W_E L_E \propto \gamma^{-2}$, the contact resistivity *must* vary in proportion to $\rho_{b,v} \propto \gamma^{-2}$ to obtain constant R_{bb}.

To maintain a constant R_{bb}, it is sufficient for the second (sheet resistivity) term in (3) to remain constant. This requires constant W_E/L_E, implying that $W_E \propto \gamma^{-1}$ and $L_E \propto \gamma^{-1}$. HBT bandwidth then varies as the inverse of the process minimum feature size, a desirably rapid variation. This scaling strategy is, unfortunately, not viable given thermal constraints.

Approximating heat flow as half-cylindrical at radii $r < L_E/2$ and as hemispherical at greater distances [29], [30], the junction temperature rise of an isolated HBT on a thick substrate is

$$\Delta T \approx \frac{P}{\pi K_{\mathrm{InP}} L_E} \ln \left(\frac{L_E}{W_E} \right) + \frac{P}{\pi K_{\mathrm{InP}} L_E}. \tag{4}$$

K_{InP} is the substrate thermal conductivity and P the dissipated power. A first strategy scales $W_E \propto \gamma^{-1}$ and $L_E \propto \gamma^{-1}$. Given that R_{bb} must not increase with scaling, this is the minimum rate at which W_E must be reduced to obtain increased bandwidth. Unfortunately, this strategy
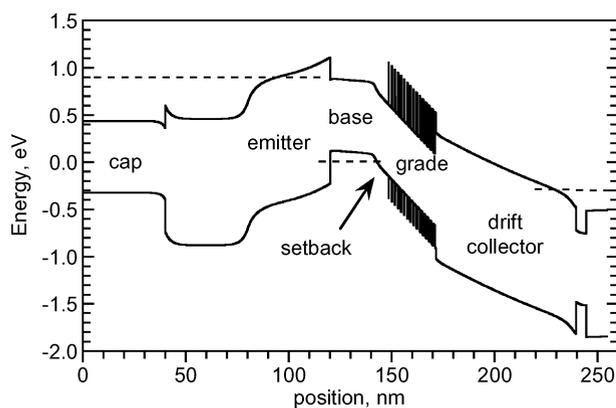
Fig. 3. *HBT semiconductor layer structure and band diagram. The emitter cap is InGaAs, the emitter InP, the base and setback layer are InGaAs, the grade an InAlAs/InGaAs superlattice, and the drift collector and subcollector InP.*

results in $\Delta T \propto \gamma$, a rapid increase in temperature with scaling.

A second scaling strategy scales $W_E \propto \gamma^{-2}$ and $L_E \propto \gamma^0$. Minimum feature size then decreases in proportion to the inverse square of IC bandwidth. This results in a moderate logarithmic increase with scaling in the temperature *of an isolated HBT*.

A third strategy scales W_E and L_E such that ΔT remains constant while junction area scales as $L_E W_E \propto \gamma^{-2}$. An exact solution is complex and uninformative; nearly constant ΔT is obtained over the next several scaling generations by taking $W_E \propto \gamma^{-3}$ and $L_E \propto \gamma^1$. This strategy maintains constant junction temperatures and constant thermal stability factors [31] with scaling [32] but requires that lithographic dimensions vary as the inverse *cube* of IC bandwidth, placing severe demands on IC fabrication.

We select the intermediate strategy ($W_E \propto \gamma^{-2}$, $L_E \propto \gamma^0$), providing a balance between thermal and lithographic constraints, and propose below (Table 1) a roadmap extending through 480-GHz digital clock rate.

The table assumes parameters feasible for NPN InGaAs/ InP HBTs. The temperature rise ΔT is calculated at $V_{CE} = 1.5$ V and $L_E = 2$ μm and assumes the calculation of (4). W_{eb} and W_{under} are both set at 8% of W_E. The calculations are from [6], with $v_{eff} \sim 3.5 \cdot 10^7$ cm/s and $D_n \sim 40$ cm^2/s selected to obtain the best fit of projected f_τ with the experiment. This fit results in good correlation with measured digital clock rate $f_{clock,max}$ but with calculated $f_{max} \sim 20\%$ higher than the experiment. Because base sheet resistance terms are *reduced* by scaling, base contact width and contact resistivity scaling have been slightly relaxed while still attaining the desired bandwidth.

C. Scaling Limits: Contact and Thermal Resistivities

Contact and thermal resistivities are presently the most serious barriers to scaling. Contact resistivities must decrease in proportion to the *square* of circuit bandwidth; 5 $\Omega - \mu m^2$ base $\rho_{v,b}$ contact resistivity and 2 $\Omega - \mu m^2$ emitter ρ_{ex} access resistivity are required for 480-GHz digital clock rate. On our best HBT samples to date, we

Table 1 InP HBT Scaling Laws and Proposed Scaling Roadmap

Parameter	scaling law	Gen. 2 (512 nm)	Gen. 3 (256 nm)	Gen. 4 (128 nm)	Gen 5 (64 nm)
MS-DFF speed	γ^1	150 GHz	240 GHz	330 GHz	480 GHz
Amplifier center frequency	γ^1	245 GHz	430 GHz	660 GHz	1.0 THz
Emitter Width	$1/\gamma^2$	512 nm	256 nm	128 nm	64 nm
Resistivity	$1/\gamma^2$	16 Ω-μm^2	8 Ω-μm^2	4 Ω-μm^2	2 Ω-μm^2
Base Thickness	$1/\gamma^{1/2}$	300Å	250 Å	212 Å	180 Å
Contact width	$1/\gamma^2$	300 nm	175 nm	120 nm	60 nm
Doping	γ^0	7 10^{19} /cm^2	7 10^{19} /cm^2	7 10^{19} /cm^2	7 10^{19} /cm^2
Sheet resistance	$\gamma^{1/2}$	500 Ω	600 Ω	708 Ω	830 Ω
Contact ρ	$1/\gamma^2$	20 Ω-μm^2	10 Ω-μm^2	5 Ω-μm^2	2.5 Ω-μm^2
Collector Width	$1/\gamma^2$	1.2 μm	0.60 μm	0.36 μm	0.18 μm
Thickness	$1/\gamma$	1500 Å	1060 Å	750 Å	530 Å
Current Density	γ^2	4.5 mA/μm^2	9 mA/μm^2	18 mA/μm^2	36 mA/μm^2
$A_{collector}/A_{emitter}$	γ^0	2.4	2.4	2.9	2.8
f_τ	γ^1	370 GHz	520 GHz	730 GHz	1.0 THz
f_{max}	γ^1	490 GHz	850 GHz	1.30 THz	2.0 THz
$V_{BR,CEO}$		4.9 V	4.0 V	3.3 V	2.75 V
ΔT		39 K	50 K	61 K	72 K
I_E / L_E	γ^0	2.3 mA/μm	2.3 mA/μm	2.3 mA/μm	2.3 mA/μm
τ_f	$1/\gamma$	340 fs	240 fs	180 fs	130 fs
C_{cb}/I_c	$1/\gamma$	400 fs/V	280 fs/V	240 fs/V	170 fs/V
$C_{cb}\Delta V_{logic}/I_c$	$1/\gamma$	120 fs	85 fs	74 fs	52 fs
$R_{bb}/(\Delta V_{logic}/I_c)$	γ^0	0.76	0.47	0.34	0.26
$C_{je}(\Delta V_{logic}/I_C)$	$1/\gamma^{3/2}$	380 fs	180 fs	94 fs	50 fs
$R_{ex}/(\Delta V_{logic}/I_c)$	γ^0	0.24	0.24	0.24	0.24

measure $5 \ \Omega - \mu\text{m}^2$ for the emitter resistivity and less than $5 \ \Omega - \mu\text{m}^2$ for the base resistivity. The emitter thus presents the more serious challenge.

In addition to the effects of doping and barrier potential, resistivity of ex-situ deposited contacts is strongly influenced by surface oxides and cleaning procedures. For the base contacts, we observe the lowest resistivity with Pd solid-phase-reaction contacts [33], which penetrate oxides. We have recently investigated methods to reduce the emitter contact resistivity. Details will be reported elsewhere [34]; several processes have been developed that provide between 0.5 and $1.0 \ \Omega - \mu\text{m}^2$ contact resistivity to N+ InGaAs or InAs layers.

When biased at $18 \ \text{mA}/\mu\text{m}^2$, the 125-nm scaling generation HBT has a nearly degenerate electron concentration at the peak of the InP/InGaAs energy barrier in the base-emitter junction. Approximating to first order in I_c the resulting deviation in V_{be} from the $V_{\text{be}} \propto (kT/q)\ln[I/I_o]$ Boltzmann characteristics, the effect can be modeled as an increase[1] in R_{ex}, proportional to $1/m^*_{\text{InP}}$, of $\sim 1 \ \Omega - \mu\text{m}^2$.
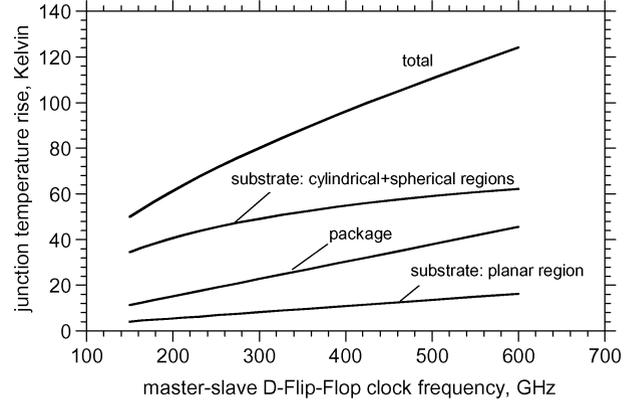
IC and package thermal resistance appear the most serious scaling challenges. On the IC, transistors are arranged in dense arrays. The transistor spacings D are small and must scale as $1/\gamma$ to scale wiring delays. Heat flow is then approximately half-cylindrical at radii $r < L_E/2$, hemispherical at radii $L_E/2 < r < D/2$, and planar at radii $D/2 < r < T_{\text{sub}}$, where T_{sub} is the substrate thickness. The substrate temperature rise is then

$$\Delta T_{\text{InP}} \sim \Delta T_{\text{cylindrical}} + \Delta T_{\text{hemisperical}} + \Delta T_{\text{planar}}$$
$$= \left(\frac{P}{\pi K_{\text{InP}} L_E}\right) \ln\left(\frac{L_e}{W_e}\right)$$
$$+ \left(\frac{P}{\pi K_{\text{InP}}}\right)\left(\frac{1}{L_E} - \frac{1}{D}\right)$$
$$+ \left(\frac{P}{K_{\text{InP}}}\right)\frac{T_{\text{sub}} - D/2}{D^2}. \quad (5)$$

In addition to the logarithmic temperature increase arising from narrow emitters, because $D \propto 1/\gamma$ at fixed T_{sub}, scaling causes ΔT_{planar} to scale as $\Delta T_{\text{planar}} \propto \gamma^2$. The planar term can be reduced by extreme thinning of the substrate, with $T_{\text{sub}} \propto 1/\gamma$, using wafer lapping or thermal vias [35].

Given a square IC of linear dimensions W_{chip} on a large copper heat sink, the package thermal resistance is approximated by planar and spherical regions, giving $\Delta T_{\text{package}} \cong (1/4 + 1/\pi)(P_{\text{chip}}/K_{Cu} W_{\text{chip}})$. Because $D \propto 1/\gamma$, $W_{\text{chip}} \propto 1/\gamma$; hence $\Delta T_{\text{package}} \propto \gamma$.

Fig. 4 shows temperature rise versus clock rate computed from (5). The calculation scales, according to

---

[1]The authors are indebted to G. Liang and M. Lundstrom (Purdue University, West Lafayette, IN) for their unpublished analyses.

Fig. 4. *Calculated package and substrate temperatures rise, as a function of digital clock rate, for a 2048-HBT CML integrated circuit.*

Table 1, the transistor, circuit, and layout parameters of a $\sim$150 GHz current-mode-logic (CML) M–S latch [51]. This latch operates HBTs at $I_E = 10$ mA (when on), $V_{CE} \sim 1$ V, and $R_L = 25 \ \Omega$. Dissipation of the 16-HBT circuit is 210 mW and the die area is $300 \cdot 300 \ \mu\text{m}^2$, of which the HBTs lie in a central $120 \cdot 120 \ \mu\text{m}^2$ region at an average $D = 30 \ \mu\text{m}$ separation. Given the $\sim$180 μm/ps propagation velocity, internal gate wiring delays are $\sim$0.7 ps (20% of T_{gate}) while wiring delays between adjacent gates are $\sim$1.7 ps (50% of T_{gate}); scaling $D \propto 1/\gamma$ maintains wiring delays at a fixed fraction of T_{gate}. We assume a 2048-HBT circuit (256 M–S latches); IC dimensions and power are obtained by multiplying those of the latch. The substrate is thinned aggressively: $T_{\text{sub}} = 40 \ \mu\text{m} \cdot (150 \ \text{GHz}/f_{\text{clock}})$.

Constraints are application specific; hence Fig. 4 only illustrates ΔT calculation. It appears, however, that 480-GHz f_{clock} is thermally feasible in 2000-transistor ICs, an integration scale typical of many ADCs and DACs. Employing HBTs with $2:1$ smaller L_E and I_E, and $2:1$ increased R_L, would reduce power.

As the collector is thinned, breakdown voltage $V_{\text{br,ceo}}$ is reduced. Table 1 gives measured $V_{\text{br,ceo}}$ versus for InP/InGaAs/InGaAlAs/InP DHBTs having 150-, 106-, and 75-nm T_C; $V_{\text{br,ceo}}$ at 53 nm T_C is extrapolated from 63-nm measured data. If their thicknesses are an excessive fraction of the total collector depletion layer thickness, the narrow-bandgap InGaAs setback and InGaAs/InAlAs grade layers (Fig. 3) may reduce V_{br} through increased Zener tunneling. When necessary, these layers can be thinned. $V_{\text{br,ceo}}$ is not the sole limit to applied voltage; device heating often imposes more stringent limits to the applied V_{CE}.

Because the bipolar transistor scaling laws presented here assume the drift-diffusion equation, the high-frequency limits of this approximation should be briefly examined. Transient transport in the HBT base and velocity overshoot in the collector depletion region are

addressed in [10]. There is no expectation that v_{eff} will decrease in HBTs with very thin base or collector depletion layers. In *bulk* (nonplanar) point-contact Schottky diodes, the finite skin depth $\delta_s = (2/\omega\mu_o\sigma)^{1/2}$ in doped semiconductor layers, the dielectric relaxation frequency $\omega_d = \sigma/\varepsilon$, the scattering frequency (the inverse of the momentum relaxation time) $\omega_s = q/m^*\mu$, and the plasma resonance frequency $\omega_p = (\omega_s\omega_d)^{1/2}$ all impair the diode responsivity at frequencies near 20 THz [36]. Here $\sigma = 1/\rho = q\mu n$ is the conductivity, μ_o the permeability, ε the permittivity of the semiconductor, and m^* the carrier effective mass. For the HBT emitter and collector layers, n-InGaAs at $\sim 3.5 \cdot 10^{19}/\text{cm}^3$ doping, $\omega_s/2\pi = 7$ THz, $\omega_s/2\pi = 7$ THz, and $\omega_p/2\pi = 74$ THz. For HBT base layers, p-InGaAs at $\sim 7 \cdot 10^{19}/\text{cm}^3$ doping, $\omega_s/2\pi =$, $\omega_s/2\pi = 12$ THz, and $\omega_p/2\pi = 31$ THz. Even at 5 THz, the skin depth in such layers well exceeds the base and subcollector layer thicknesses; hence semiconductor skin effect has negligible effect on HBT bandwidth. Because ω_d and ω_p far exceed anticipated HBT bandwidths, the emitter, base, and subcollector bulk resistivities all can be approximated as $\rho_{j\omega} = (1 + j\omega/\omega_s)$ [36]. Even this effect is negligible. Because the HBT lateral dimensions must be reduced in proportion to the inverse square of increases in device bandwidth, even given bulk resistivities increasing with frequency, base and collector contact resistivities will dominate over bulk semiconductor resistivities. To reiterate, achievable metal–semiconductor contact resistivities and device thermal resistivity are the primary limits faced in scaling HBTs for terahertz bandwidths.

D. Representative Mesa HBT Results

We show two representative results for HBTs fabricated in a simple mesa fabrication sequence. This process flow (Fig. 5) employs mesa etches to define junctions and metal depositions to define self-aligned contacts. Mesa processes incur difficulties in yield of large ICs, but their simplicity facilitates exploratory device fabrication.

Fig. 6 shows a device scanning electron microscope (SEM) image and Fig. 7 results. Griffith *et al.* [37] reported initial results with HBTs at 250-nm W_e—having
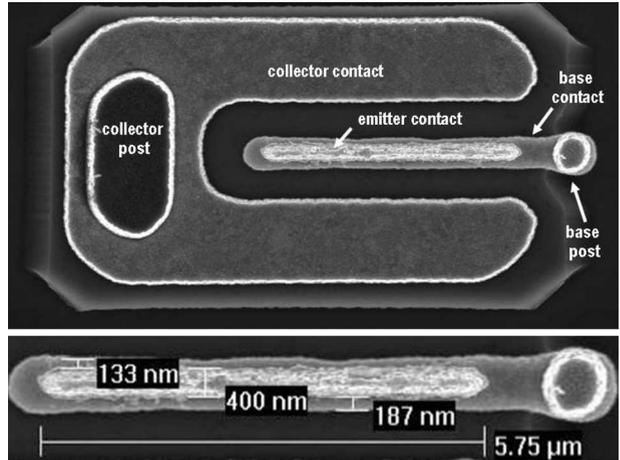
Fig. 6. *SEM images (top view and closeup of emitter-base junction) of a simple mesa DHBT. The emitter and base contact widths are 400 and $\sim$150 nm, respectively. The emitter-base junction width is $\sim$300 nm.*
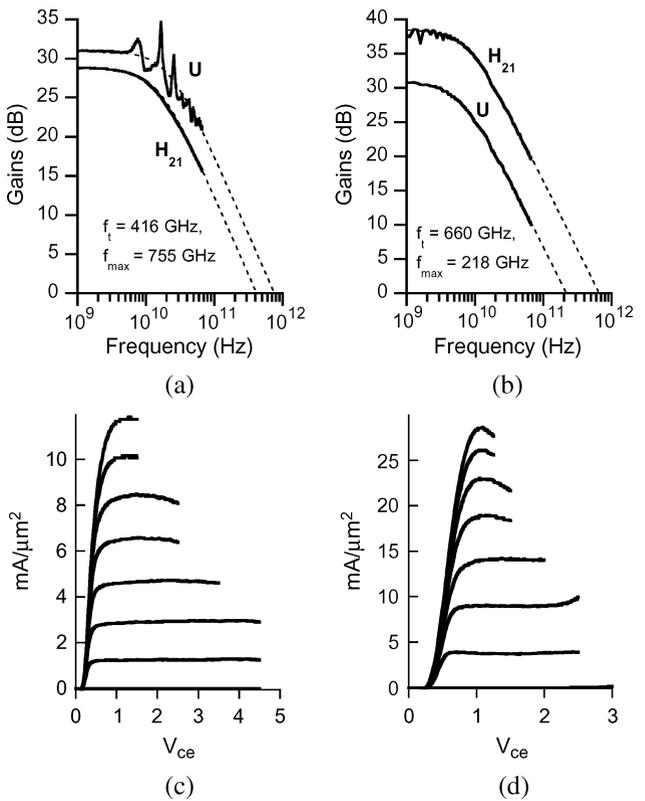
Fig. 7. *Measured microwave gains and common-emitter characteristics of representative mesa DHBTs. (a) Gains of a DHBT having $T_c = 150$ nm, $T_b = 30$ nm, and $W_e = 250$ nm, biased at $J_e = 12$ mA/μm^2. (b) Gains of a DHBT having $T_c = 60$ nm, $T_b = 14$ nm, and $W_e = 400$ nm, biased at $J_e = 13$ mA/μm^2. DC common-emitter characteristics of the HBT having (c) $T_c = 150$ nm and (d) $T_c = 60$ nm are also shown.*
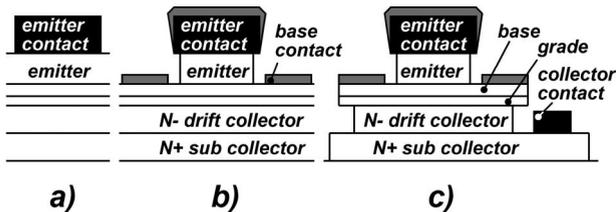
Fig. 5. *Mesa HBT process flow. (a) Emitter metal deposition. (b) Emitter etch and self-aligned base contact metal deposition by liftoff. (c) Base mesa etch, collector mesa etch, and collector contact deposition.*

larger T_C than devices of the same scaling generation in Table 1—exhibiting 755-GHz f_{max} and 416-GHz f_τ. These devices exhibit $V_{br,ceo} = 4$ V at $I_c/W_E L_E = 0.1$ mA/μm^2 leakage. To explore variation of breakdown with collector thickness, devices with 60-nm T_c were fabricated by Griffith *et al.* [38]. f_τ, at 660 GHz, is lower than that expected of properly scaled devices (Table 1) because of both excessive $(5 \ \Omega - \mu m^2) \ \rho_{ex}$ and excessive emitter junction width that indirectly limits (4) the bias current density. The measured breakdown $V_{br,ceo} = 3$ V at 10 kA/cm^2 indicates that 2.5–3 V breakdown can be obtained even at the 62-nm scaling generation ($\sim$1–1.5 THz f_τ and f_{max}).

Examples of mesa HBT IC results include 172-GHz medium-power amplifiers using 500-nm W_e devices with 300-GHz f_{max} and 150-GHz static frequency dividers [14].

III. MANUFACTURABLE HBT PROCESSES

To fabricate large ICs, high transistor yield is necessary. For acceptable IC power dissipation, junction areas must be small. Without these two key parameters, a wide-band transistor technology has no relevance to ICs of significant integration scale. To permit self-aligned base metal liftoff, the mesa process (Fig. 5) requires an emitter semiconductor significantly thicker than the base contact metal. Excessive undercut during the emitter semiconductor etch then becomes a serious yield difficulty for HBTs with narrow emitters, as do base-emitter short-circuits associated with the base metal liftoff.

More reliable and higher yield process flows are needed for fabrication of narrow base contacts self-aligned to the emitter-base junction. One approach is to adapt to III–V semiconductors [39]–[41] the emitter regrowth processes used for SiGe HBTs [42]. Emitter dielectric sidewall processes [13], [15], [43] have shown greater success and will be described in detail.

A second challenge is junction passivation; reliably controlling the effect of the junction surface both on leakage currents and on breakdown. A third challenge arises from the base contact pad capacitance, which becomes dominant in the short-emitter-length HBTs employed in lower power circuits.
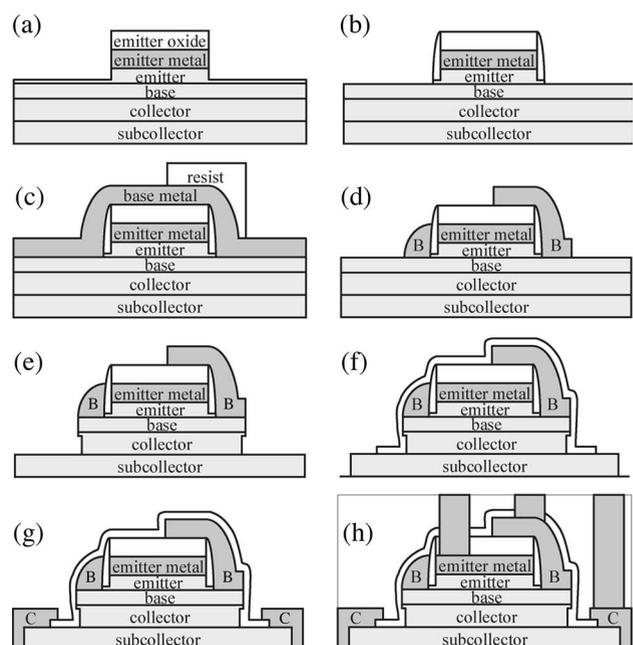
Below are described two IC fabrication processes, developed independently at Vitesse Semiconductor and at Teledyne Scientific. Both processes separate the emitter and base contacts using dielectric sidewall spacers, thereby avoiding base contact liftoff. The two processes diverge in the definition of the base contact; the Vitesse process defines a base contact in the form of a metal sidewall through isotropic metal deposition and anisotropic etching, while the Teledyne process defines the base contact by selective-area plating, in a process similar to the self-aligned silicide processes of VLSI. In the Vitesse process, the base contact pad area is eliminated by defining above the emitter, while in the Teledyne

process the pad capacitance is reduced by a collector pedestal implant.

A. Self-Aligned Metal Sidewall Technology

He *et al.* describe in [13] a DHBT fabrication process that avoids the difficulties, discussed above, associated with fabricating DHBTs using a series of semiconductor mesa etches and self-aligned contact metal liftoff steps. This process is referred to by its manufacturer, Vitesse Semiconductor, as the VIP-2 process. In this process, at no point in the process is metal liftoff employed for contact metal definition. A dielectric spacer defines the self-aligned emitter-base spacing, while a metal spacer defines the self-aligned base width, allowing narrow base contacts to be defined independent of lithographic constraints. The collector junction is self-aligned directly to the base. During the process flow, junction surfaces are passivated immediately after their formation, and ledge passivation is incorporated into both the EB and the BC junctions. The fabrication process is illustrated in Fig. 8 and is discussed below.

In Fig. 8(a), a tungsten-based emitter contact metal is deposited on the epitaxial wafer, followed by a dielectric hard mask that both insulates the base metal from the emitter metal and suppresses galvanic reactions during wet processes. The metal and hard mask layers are defined by lithography and a subsequent etch. An anisotropic plasma etch then removes most of the emitter epitaxial layer. A short wet etch completes the emitter etch and removes surfaces damaged by the plasma etch. Since there is no

Fig. 8. *Fabrication procedure of metal-sidewall DHBT. Process steps (a)–(h) are described in the text.*
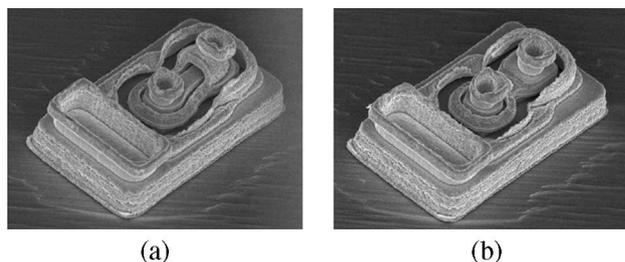
Fig. 9. *SEM images of metal-sidewall HBTs where base metal (a) does and (b) does not overlap emitter metal.*

need to accommodate a subsequent liftoff, the emitter epitaxial layer can be thin, and no emitter etch undercut is necessary. The emitter etch is therefore made significantly more controllable.

To permit placement of a contact via between the emitter metal and the IC interconnects, the emitter at one end expands to a width nearly twice the minimum width. It should be noted that the periphery/area ratio of a circular emitter of diameter D is the same as that of a long emitter stripe of width $W_E = D/2$. While the flared emitter contact pad area must entail some increase in the device $R_{bb}C_{cb}$ time constant and in the device thermal resistance, the devices exhibit RF performance consistent with that expected of the 512-nm scaling generation. The potential application of this process flow to fabrication of HBTs with 250 and 125 nm W_E will depend upon the ability to proportionally reduce the diameter of the emitter contact via and to proportionally improve its precision in placement.

Immediately after the emitter etch, a conformal dielectric layer is deposited to passivate the emitter junction. It is then plasma etched anisotropically to create a spacer separating the base and emitter contacts [Fig. 8(b)]. The spacer width is easily scaled by adjusting the dielectric deposition thickness.

After a short wet etch to remove the plasma damage from the spacer dry etch, base contact metal using platinum and tungsten is deposited. To define a pad upon which the base interconnect via will land [Fig. 8(c)], the metal is patterned by a base-metal mask. A dielectric hard mask is used in this step. To reduce collector junction capacitance, the base-metal pad may overlap the emitter metal since they are isolated by the emitter hard mask. An anisotropic plasma etch forms the base-metal pad and the self-aligned base metal around all emitter structures [Fig. 8(d)]. The width of the self-aligned base metal is determined not by lithography but by the thickness of the base metal and its hard mask.

The collector junction is defined by a plasma etch using the base metal as a self-aligned mask. A wet etch removes the plasma etch damage and completes the formation of the collector [Fig. 8(e)]. A dielectric layer is

then deposited and patterned immediately to passivate the collector junction. Subcollectors are then isolated by a plasma etch [Fig. 8(f)]. Tungsten-based collector metal is subsequently deposited, a mask defined by lithography, and the metal is etched [Fig. 8(g)]. Resistor metal is then deposited and patterned. After a final wet clean, a dielectric passivation layer is deposited to fully encapsulate the substrate. An aluminum-based four-layer interconnect process completes IC fabrication [Fig. 8(h)]. Fig. 9 shows SEM images of metal-sidewall devices with interconnect metal after all dielectric layers are removed to reveal device structures.

The passivation of junction surfaces greatly affects leakage currents, and hence device performance and reliability. To minimize surface leakage currents, the device incorporates passivation ledges—thin depleted layers of exposed wideband-gap material—in both the emitter-base and base-collector junctions. In the Vitesse sidewall process, the emitter etch leaves a thin ledge on which the dielectric spacer is deposited. The ledge layer is later removed after the spacer etch to confine the ledge to the space between the emitter semiconductor and the base contacts. The ledge thickness must be sufficiently thin to ensure full depletion.

To control the ledge thickness, and to facilitate inline manufacturing process monitoring and control, the ledge fabrication process is monitored by spectroscopic ellipsometry using ultraviolet and visible wavelengths. Quantitative relationships between ledge thickness and current gain were obtained using such ellipsometry measurements (Fig. 10). At ultraviolet and visible wavelengths there is sufficient contrast between the refractive indexes of InP and InGaAs to distinguish the ledge from the base. The strong absorption in this spectral region also improves the near-surface sensitivity of the measurement by suppressing interfering signals from deeper layers.

Similar to the emitter-base junction, ledge passivation of the base-collector junction reduces the base-collector leakage current. An exposed junction ledge is formed from
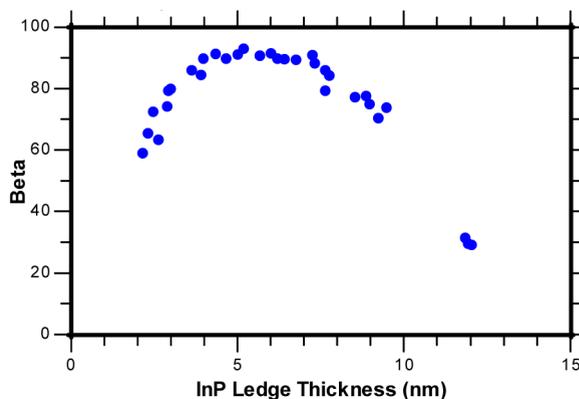
Fig. 10. *Common-emitter current gain versus ledge thickness.*
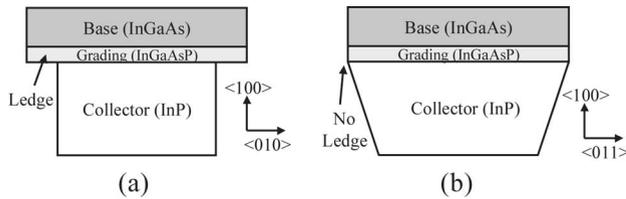
Fig. 11. *Ledge passivation of the base-collector junction (emitter-base junction not shown).*

the base-collector quaternary grading layer by slightly undercutting the InP collector from the base using a selective InP wet etch during the collector junction etch. In this technique, the orientation of the transistor can directly affect its BC ledge passivation due to the wet-etch facet characteristics. As illustrated in Fig. 11, a consistent undercut can be achieved when device is oriented in the ⟨010⟩ direction, but undercut may not be achieved for devices oriented in the ⟨011⟩ direction because of pinning of the etch facet by the InGaAs base layer. This orientation dependence is observed electrically when pairs of nearby base-collector junctions ($2.5 \times 2.5\ \mu m$ base-collector diode test structures) of different orientation are measured (Fig. 12).

B. Metal Sidewall Technology: Device Results

The baseline metal-sidewall transistors have a 500-nm minimum emitter width as defined by lithography and deliver more than 300 GHz f_τ and f_{max} with $V_{br,ceo}$ over 5 V. Its epitaxial structure incorporates a compositionally graded strained InGaAs base for drift-assisted carrier transport. The base is 30 nm thick, p-type doped to $4.5 \cdot 10^{19}$ cm^{-3} with a 7% grade in Ga composition. The

InP emitter is 30 nm thick, n-type doped to $7 \cdot 10^{17}$ cm^{-3}, and capped with a 50-nm n+ contact layer that step-grades from InP to InGaAs and finally toward InAs to minimize emitter resistance. The InP collector is 150 nm thick, n-type doped to $\sim 8 \cdot 10^{16}$ cm^{-3}, and incorporates an InGaAsP grading with an associated delta doping layer to neutralize the quasi-field associated with the grade.

To demonstrate the potential for increased transistor bandwidth in the metal sidewall process, the epitaxial layer structure was thinned in some process runs. The base layer was thinned 20 nm and its doping increased to $8 \cdot 10^{19}$ cm^{-3}. The collector layer was thinned to 100 nm. Further, a 10-nm emitter-base junction grading layer was introduced, reducing the required base-emitter voltage at a given current density (Fig. 13), providing a smaller and more constant collector-current ideality factor, and increasing transconductance at high current density. The emitter-base grade also provides an etch stop on the ledge, greatly improving the control of the emitter etch and hence the ledge passivation. With this scaled epitaxial structure, we have measured approximately 400 GHz f_τ and 450 GHz f_{max}.

C. Metal Sidewall Technology: Reliability Results

High current density accelerated life tests were performed on metal-sidewall DHBTs of $1 \times 1\ \mu m^2$ drawn emitter size. The tests were carried out at a bias of 4 mA/μm^2 Zemitter current density and at junction temperatures of 222 and 272 °C for durations as long as 2770 h. Forty devices were stressed at each temperature. Although the junction temperature employed for these tests is similar to that reported for other InP device reliability investigations, we have employed a higher current density [44]–[48], reflecting the required increase in current density in HBT processes of progressively
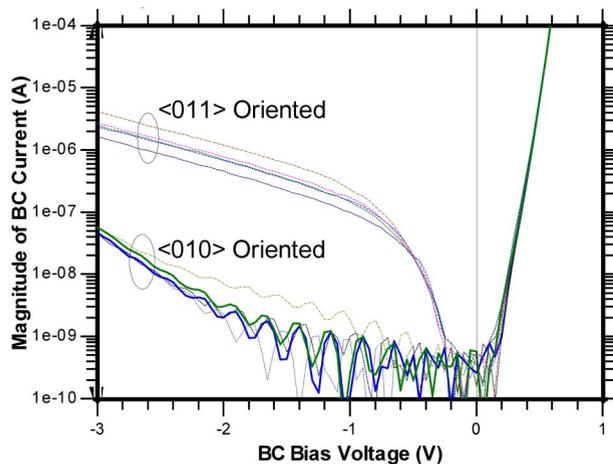
Fig. 12. *Base-collector leakage current of different device orientations across the same wafer ($2.5 \times 2.5\ \mu m$ base-collector diode test structures).*
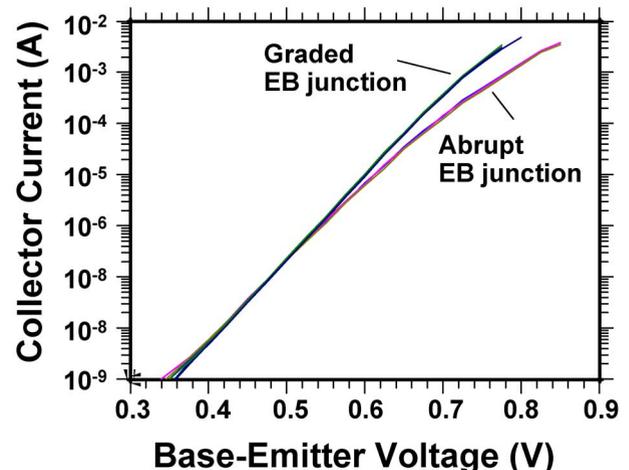

Fig. 13. *Forward-active collector current at $V_{cb} = 0$ V for wafers having abrupt and graded emitter-base junctions.*

increasing bandwidth. The failure criterion employed was a 20-mV change in the base-emitter voltage V_{be} required to maintain 4 mA/μm^2 emitter current density. The median time to failure was 700 h for devices stressed at 272 °C junction temperature and 15 000 h for devices stressed at 222 °C. Extrapolation of these data using the Arrhenius method indicates a failure activation energy of 1.4 eV. The mean time to failure is then extrapolated as 8.8 × 10^8 h (10 000 years) at 100 °C junction temperature and 4.4 × 10^6 h (502 y) at 150 °C junction temperature.

In other reports of InP HBT reliability [44]–[46], [48], degradation in dc current gain (β) is the most commonly cited, with reported activation energies between 1.0–1.5 eV. With the metal-sidewall transistors here reported, β generally increased by about 30% during accelerated testing and β degradation was not the primary failure mechanism. Kiziloglu *et al.* [47] reported HBT reliability data with V_{be} variation as the primarily failure mechanism, finding a failure activation energy of 1.5 eV for HBTs employing an InAlAs emitter.

D. Metal Sidewall Technology: Circuit Results

Commercial and defense applications include high-frequency mixed-signal integrated circuits. Highlighting the integration level possible with this technology is a 30-GHz clocked DDS with more than 3000 transistors and an average of 40-dB spurious-free dynamic range fabricated with over 50% integrated circuit yield [19]. Higher levels of integration are possible, and circuits with more than 10 000 transistors have been fabricated. Circuits operating at 160 Gb/s have been designed in the VIP-2 metal-sidewall process including a 4 : 1 multiplexer [20]. Track-and-hold circuits have been demonstrated at 10-G/s sampling rate with 20-GHz bandwidth and 1-V peak-to-peak differential signal. Fig. 14 shows a
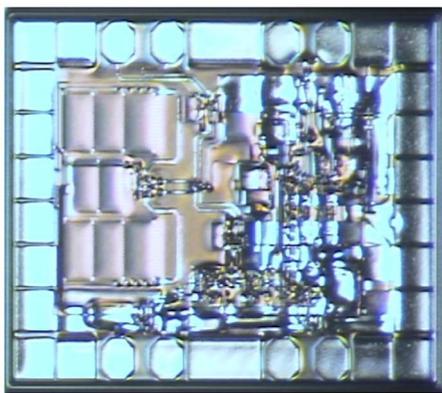
Fig. 14. *PLL die photograph shows VCO tank circuit on the left and loop filters and 32 : 1 divider on the right. This monolithic circuit operates at 80 GHz and exhibits −80 dBc (1 Hz) phase noise at 1-MHz offset from carrier.*

monolithic phase-locked loop (PLL) that includes a differential 80-GHz Colpitts voltage-controlled oscillator (VCO) with +/−4% tuning range, a 2-MHz bandwidth loop filter, and a 32 : 1 frequency divider. The measured phase noise is −80 dBc (1 Hz) at 1-MHz offset from the carrier. Devices with high breakdown voltages for microwave and millimeter-wave power amplification have also been investigated.

E. Self-Aligned Plated-Base-Contact Technology

In a mesa HBT process (Figs. 5 and 6), base metal contacts are self-aligned to the emitter junction by using metal liftoff. As noted above, such processes face yield limitations associated with difficulties in reliably forming breaks in the deposited metal film between the emitter and base contacts. The metal-sidewall process avoids this difficulty by covering the entire emitter contact metal stripe by dielectric sidewalls and a dielectric top cap; the base contact metal then covers the entire emitter contact. Consequently, in the metal-sidewall process, the emitter terminal must be accessed using a narrow via aligned to the emitter junction. An alternative to this process is to selectively deposit, via electroplating, the base contact metal only on the exposed surfaces of the base semiconductor. Narrow emitters can then be contacted using a planarization and etch-back process. We refer to this process flow as a self-aligned plated-base-contact technology. Such a process has been developed at Teledyne Scientific and is described by Urteaga *et al.* in [15]. As with the metal-sidewall process, the plated-base-contact process separates the emitter and base contacts using dielectric spacers.

The process flow is illustrated schematically in Fig. 15. The process starts with emitter contact formation. Because the emitter contact metal functions as a via to connect to first-level metal, this contact must be relatively tall, and a high height/width aspect ratio is therefore required. This is achieved by electroplating the emitter contact metal. Emitter junction areas are defined by patterning openings into photoresist and then electroplating to fill these with metal. Plating fully fills the resist opening, and the metal sidewall slope conforms to that of the photoresist. Emitter contacts can thus be formed with high height/width aspect ratios and nearly vertical sidewalls. This is in marked contrast to metal liftoff, for which sidewall slopes are typically ∼4 : 1 and for which attempts to form tall metal contacts to narrow emitter junctions result in a triangular metal cross-section with a metal height at most 4 : 1 larger than the emitter width. Because plating can produce tall metal contacts to even narrow emitters, the emitter contact metal can serve as the contact via/post even at scaling generations below 250-nm emitter width.

Subsequent to emitter metal definition, the emitter-base junction is defined by an HCl-based selective wet etch
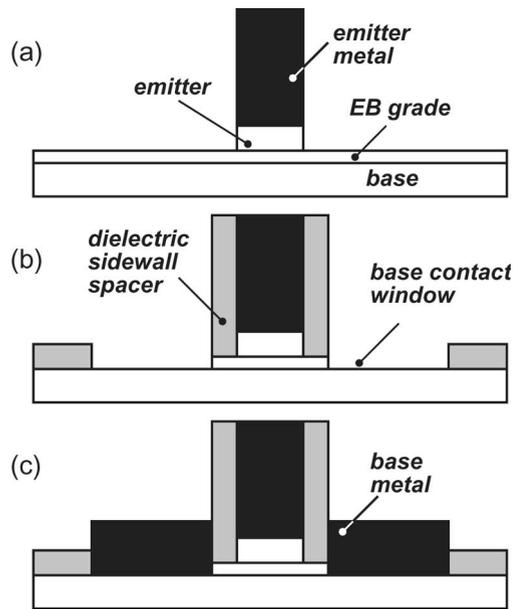
Fig. 15. *Schematic cross-section of the base-emitter process flow incorporating base-emitter dielectric spacers and selective-area electroplating of the emitter and base contacts. (a) Emitter contact metal plating and emitter semiconductor etch. (b) Dielectric sidewall definition and base contact lithography. (c) Base contact formation by selective-area electroplating.*

[Fig. 15(a)]. Because contact liftoff is not employed, the emitter semiconductor is thin, minimizing junction undercut during etching. The etch stops on the InAlAs/InGaAs superlattice grade in the base-emitter junction.

The base and emitter contacts are separated laterally by a dielectric sidewall formed by conformal (chemical vapor deposition) dielectric deposition and anisotropic (reactive ion etching) etch removal [Fig. 15(b)]. The dielectric sidewall contacts the InAlAs/InGaAs superlattice grade; because InAlAs exhibits a larger energy separation between the conduction band edge and the surface Fermi level, a potential barrier is formed that drives electrons away from the surface, e.g., a passivation ledge is formed.

Subsequent to sidewall formation, the base-emitter grade layers are removed and the base contacts formed [Fig. 15(c)] by electroplating of the base metal directly onto the base semiconductor. The base metal plating proceeds in windows in a plating mask with a size significantly larger than the minimum lithographic limit, nominally centered around the emitter. At this step in the process, the top of the emitter contact metal and the base semiconductor layer are exposed, whereas the sides of the emitter metal are covered with the dielectric sidewall. The base-emitter barrier potential prevents plating onto the top of the emitter contact; the base contact metal selectively electroplates only onto the base semiconductor. A subsequent process step places a base contact post onto the plated base

metal such that the post height is coincident with the top of the emitter contact.

The base-collector junction, the collector contacts, and the N+ subcollector mesa are then defined wet-etch and metal evaporation processes. A dielectric layer is deposited after the subcollector mesa etch to reduce variations in the topography of the wafer. Thin-film resistors are deposited onto the dielectric layer, followed by contact posts to the collector and resistors. The tops of emitter, base, and collector contact posts lie in the same horizontal plane. Device SEM images are shown in Fig. 16.

The back-end interconnect process (Fig. 17) is initiated by spin-casting BenzoCycloButene (BCB) to planarize the wafer. The BCB is then etched back to expose the contact posts. The remaining interconnect steps follow a repetitive process of interconnect metal deposition, BCB dielectric deposition, and via etching. The interconnect metals and vias are all electroplated. Metal–insulator–metal capacitors with Si_xN_y dielectric are provided between first and second interconnect metal.

First-level interconnect metal uses 1-μm lines to contact posts. The second and third metal use 2-μm lines and 2-μm square vias. The interlayer dielectric is BCB.

F. Plated-Base-Contact Technology: Device Results

Transistors and integrated circuits were fabricated in the plated-base-contact process using 500-nm minimum
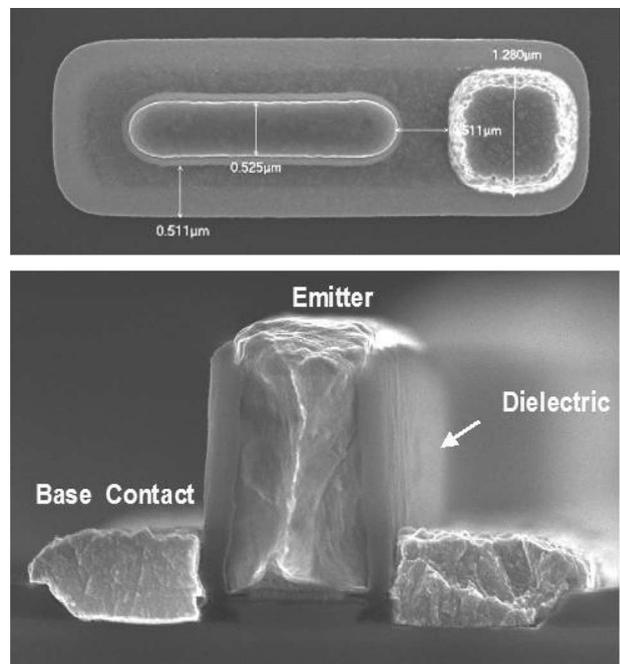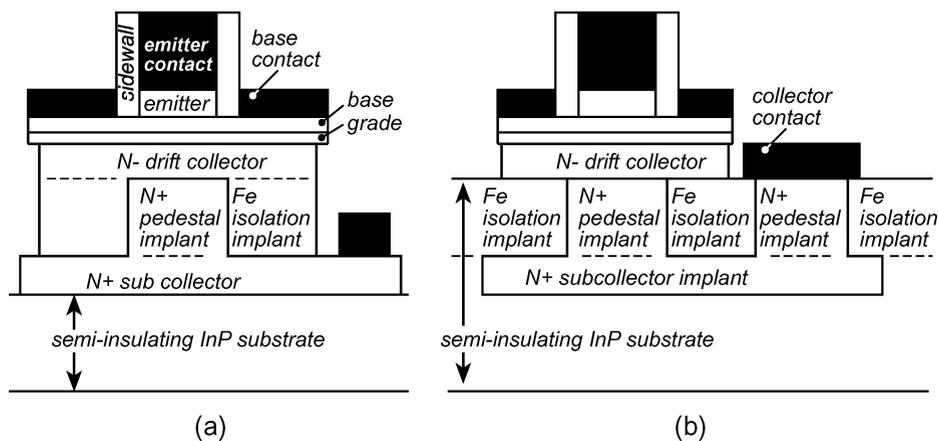

Fig. 16. *SEM top view and cross-section of the base-emitter junction of a plated-base-contact transistor having 500-nm emitter junction width.*
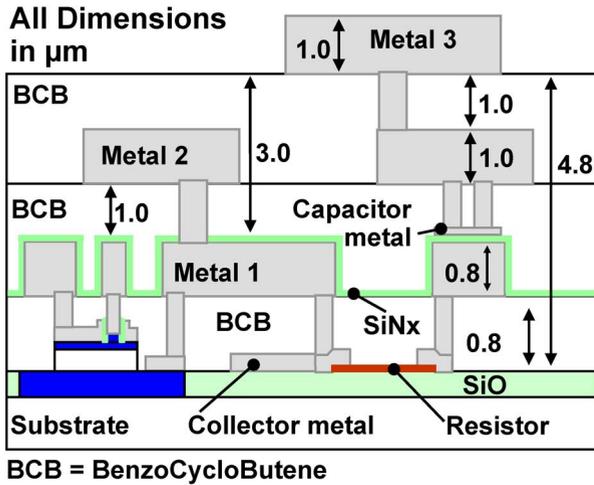
All Dimensions in μm

BCB = BenzoCycloButene

feature size [15]. The epitaxial layer structure has an InAs emitter cap layer to provide low-resistance emitter ohmic contacts. The emitter is InP. In some process runs, abrupt InP/InGaAs emitter-base junctions are employed; in others the junction is graded using an InAlAs/InGaAs short-period superlattice. The base is InGaAs whose p-type carbon doping is graded from $8 \cdot 10^{19}$ to $4 \cdot 10^{19}$ cm^{-3} across the base thickness. The base-collector grade consists of a thin InGaAs setback followed by a short-period InAlAs/InGaAs chirped superlattice. The grade is terminated by a delta-doped layer, which generates a field neutralizing the quasi-field associated with that of the grade. The collector and subcollector are InP. For the $W_E = 500$ nm scaling generation, the base and emitter layers are approximately 40 nm thick while the collector is 150 nm thick.

At $V_{cb} = 0$ V, this structure can support $J_e = 3$ mA/μm^2 current densities before onset of the Kirk effect, increasing to ~ 6 mA/μm^2 at $V_{cb} = 0.5$ V. Devices fail thermally at ~ 18 mW/μm^2; this provides adequate margin for analog circuits. The dc and RF characteristics of the 500-nm generation plated-base-contact technology are described as follows: the device shows a current gain $\beta = 40$ and $V_{be,on} = 0.8$ V at $J_e = 3$ mA/μm^2, while the emitter ideality factor is 1.5. The common-emitter breakdown voltage is $V_{br,ceo} = 5$ V at $J_e = 0.1$ mA/μm^2, while the devices show a peak 330-GHz f_τ and 420-GHz f_{max} at $J_e = 6$ mA/μm^2 and $V_{cb} = 0.5$ V.

G. Collector Pedestal Technology

From the device layout (Fig. 16) of either a plated base contact or a mesa HBT, it is apparent that the base contact post area contributes substantially to the collector-base junction area and consequently limits digital logic speed. The fractional contribution of the base pad to the collector capacitance is particularly strong in low-power circuits wherein L_E is small and the base pad capacitance can dominate C_{cb}. Adapting from similar process modules employed in Si/SiGe HBT fabrication, we have developed a collector pedestal implant process for reduced base pad capacitance [49], [50]. In the implementation reported by Urteaga *et al.* [51] [Fig. 18(a)], the pedestal is formed by molecular beam epitaxy (MBE) growth of the N+ InP subcollector and pedestal precursor layers. N+ silicon and Fe isolation implants convert the pedestal precursor layer into separate N+ pedestal regions and Fe-doped semi-insulating current-block regions. The remainder of the semiconductor layer structure is then formed in a second MBE growth and devices completed using the plated-base-contact process flow described above.

The Fe isolation regions strongly reduce the collector-base capacitance in the base contact pad region and provide a moderate reduction in the component of C_{cb}

(a)   (b)
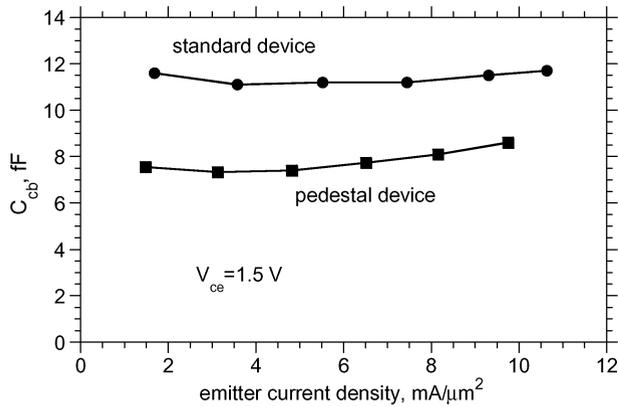
Fig. 19. *Collector capacitance of plated-base-contact devices with and without collector pedestal implants. The collector is 120 nm thick, the emitter junction is 400 nm × 5 μm, and the pedestal mask dimensions are 500 nm × 5 μm. The pedestal device exhibits 365-GHz f_τ and 390-GHz f_{max}.*

associated with the base contact area; the implanted region is ~100 nm wider than the emitter. As extracted from microwave S-parameter measurements, the measured C_{cb} is reduced ~35%.

We extract the collector-base capacitance from small signal measurements as a function of collector current density, as shown in Fig. 19. The pedestal implant permits reduced collector base capacitance in processes wherein relatively wide base contacts are employed because of IC yield considerations. A more aggressive version of implant process reported by Parthasarathy *et al.* [52] uses multiple implants to define both the pedestal and the subcollector and requires only a single MBE growth [Fig. 18(b)]. Significant improvement in collector-base leakage I_{cbo} and breakdown $V_{br,ceo}$ was also observed, this being attributed to the buried collector-base junction and the consequent reduction in surface currents.

H. Plated-Base-Contact Technology: Circuit Results

As with the metal-sidewall process, a number of ICs have been demonstrated in the plated-base-contact pro-

cess. Regarding IC yield, arrays of 1000 parallel HBTs having 0.5 × 3 μm² emitter junction exhibit 71.4% yield. As of March 2007, efforts are still in progress both to characterize and to improve the yield in fabrication of large-scale circuits containing 1000–10 000 transistors.

Master–slave latches configured as static digital frequency dividers provide a measurement of the feasible clock speed of the technology in digital and mixed-signal ICs. CML dividers using the pedestal implant consume 103 mW per latch and operate at 128 GHz [51], while ECL dividers operate at 134 GHz and consume 143 mW per latch. Low-power versions of the dividers operate up to 70-GHz clock but consume only 40 mW per latch.

Eight-bit DDS ICs have also been fabricated in the plated-base-contact process. The DDS circuits employ approximately 5000 transistors. Operation has been verified up to 16-GHz clock frequency with a spurious-free dynamic range exceeding 35 dB. The DDS uses a 12-bit frequency control word, an 8-bit DAC, and a mapper block that is the primary limitation in output resolution. The total power consumed is 17 W. Other ICs demonstrated in the plated-base-contact process include 44-GHz VCOs, 30-GHz track-and-hold circuits, and 8-bit DACs clocking at 20 GHz.

I. Technology Comparison and Summary

The metal-sidewall and plated-base process technologies avoid the yield difficulties associated with emitter-base liftoff processes. It is worthwhile to examine their performance relative to that of a less-manufacturable mesa device [14] of the same scaling generation (Table 2). The three devices use extremely similar epitaxial layer structure designs; this is largely stipulated by considerations of minimum gate delay at a 500-nm feature size. Breakdown voltage and maximum current density are most strongly determined by the layer structure design and are consequently similar in the three devices. Measured circuit performance, as determined by digital clock rate, is similar for the three devices. The slight superiority in f_{max} of the mesa device [14] is primarily a result of low base contact resistivity.

Table 2 Summary of Reported Devices

Parameter	Gen. 1 Roadmap	Mesa Device	Metal Sidewall	Plated Base
Emitter Width	500 nm	500 nm	500 nm	500 nm
Base Thickness	300Å	300Å	300Å	300Å
Collector Thickness	1500 Å	1500 Å	1500 Å	1500 Å
Current Density	4.5 mA/μm²	4-6 mA/μm²	3-5 mA/μm²	4-6 mA/μm²
$V_{BR,CEO}$	4.9 V	5.1 V	4.5 V	5 V
f_τ	370 GHz	391 GHz	330 GHz	330 GHz
f_{max}	490 GHz	505 GHz	320 GHz	420 GHz
Static divider speed	150 GHz	142 GHz	152 GHz	134 GHz

IV. CONCLUSION

Electron device bandwidths are increased by use of improved materials and by scaling. Unlike MOS transistors, which now face serious scaling limits associated with charge carrier tunneling through gate insulators, bipolar device scaling is constrained by achievable contact resistivities, by breakdown voltage, and by device and IC thermal resistances. Of these limits, IC thermal resistance appears the most serious limit. Digital clock rates of 400–500 GHz appear feasible, and innovations in thermal design may further extend the feasible frequency range. For millimeter-wave and submillimeter-wave amplification, smaller ICs are required and thermal limits are less severe; present experimental thermal, contact resistance, and breakdown data suggest that transistor amplifiers to at least 1-THz signal frequency are feasible.

High-yield fabrication processes are crucial to enable such applications. Although liftoff processes and surface passivation considerations have historically limited IC yield, recently emerged dielectric sidewall processes enable high-yield metal contact deposition, self-alignment of contacts and junctions, and well passivated, reliable, and low-leakage devices. ∎

REFERENCES

[1] A. F. J. Levi, B. Jalali, R. N. Nottenburg, and A. Y. Cho, "Vertical scaling in heterojunction bipolar transistors with nonequilibrium base transport," *Appl. Phys. Lett.*, vol. 60, pp. 460–462, 1992.

[2] Y. Matsuoka, S. Yamahata, K. Kurishima, and H. Ito, "Ultrahigh-speed InP/InGaAs double-heterostructure bipolar transistors and analysis of their operation," *Jpn. J. Appl. Phys.*, vol. 35, pp. 5646–5654, 1996.

[3] W. E. Stanchina, J. F. Jensen, R. H. Walden, M. Hafizi, H.-C. Sun, T. Liu, G. Raghavan, K. E. Elliott, M. Kardos, A. E. Schmitz, Y. K. Brown, M. E. Montes, and M. Yung, "An InP-based HBT fab for high-speed digital, analog, mixed-signal, and optoelectronic ICs," in *Tech. Dig. IEEE GaAs IC Symp.*, 1995, pp. 31–34.

[4] H. Nakajima, T. Ishibashi, E. Sano, M. Ida, S. Yamahata, and Y. Ishii, "InP-based high-speed electronics," in *IEDM Tech. Dig.*, 1999, pp. 771–774.

[5] A. Gutierrez-Aitken, E. Kaneshiro, B. Tang, J. Notthoff, P. Chin, D. Streit, and A. Oki, "69 GHz frequency divider with a cantilevered base InP DHBT," in *IEDM Tech. Dig.*, 1999, pp. 779–782.

[6] M. J. W. Rodwell, M. Urteaga, T. Mathew, D. Scott, D. Mensa, Q. Lee, J. Guthrie, Y. Betser, S. C. Martin, R. P. Smith, S. Jaganathan, S. Krishnan, S. I. Long, R. Pullela, B. Agarwal, U. Bhattacharya, L. Samoska, and M. Dahlström, "Submicron scaling of HBTs," *IEEE Trans. Electron Devices*, vol. 48, pp. 2606–2624, Nov. 2001.

[7] M. Littlejohn, K. W. Kim, and H. Tian, "High-field transport in InGaAs and related heterostructures," in *Properties of Lattice-Matched and Strained Indium Gallium Arsenide*, P. Bhattacharya, Ed. London, U.K.: INSPEC, 1993.

[8] I. Vurgaftmana, J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys," *J. Appl. Phys.*, vol. 89, no. 11, pp. 5815–5875, Jun. 1, 2001.

[9] H. Kroemer, "Heterostructure bipolar transistors and integrated circuits," *Proc. IEEE*, vol. 70, pp. 13–25, Jan. 1982.

[10] T. Ishibashi, "Nonequilibrium electron transport in HBTs," *IEEE Trans. Electron Devices*, vol. 48, pp. 2595–2605, Nov. 2001.

[11] E. P. O'Reilly, "Band structure of InP: Overview," in *Properties of Indium Phosphide*. London, U.K.: INSPEC, 1991.

[12] J. Zolper, "Challenges and opportunities for InP HBT mixed signal circuit technology," in *Proc. IEEE Int. Conf. Indium Phosphide Related Mater.*, Santa Barbara, CA, May 2003, pp. 8–11.

[13] G. He, J. Howard, M. Le, P. Partyka, B. Li, G. Kim, R. Hess, R. Bryie, R. Lee, S. Rustomji, J. Pepper, M. Kail, M. Helix, R. Elder, D. Jansen, N. Harff, J. Prairie, E. Daniel, and B. Gilbert, "Self-aligned InP DHBT with ft and fmax over 300 GHz in a new manufacturable technology," *IEEE Electron Device Lett.*, vol. 25, pp. 520–522, Aug. 2004.

[14] Z. Griffith, Y. Dong, D. Scott, Y. Wei, N. Parthasarathy, M. Dahlström, C. Kadow, V. Paidi, M. J. W. Rodwell, M. Urteaga, R. Pierson, P. Rowell, B. Brar, S. Lee, N. Nguyen, and C. Nguyen, "Transistor and circuit design for 100–200 GHz ICs," *IEEE J. Solid-State Circuits*, vol. 40, pp. 2061–2069, Oct. 2005.

[15] M. Urteaga, P. Rowell, R. Pierson, B. Brar, M. Dahlström, Z. Griffith, M. J. W. Rodwell, S. Lee, N. Nguyen, and C. Nguyen, "Deep submicron InP DHBT technology with electroplated emitter and base contacts," in *Proc. 2004 IEEE Device Res. Conf.*, Notre Dame, IL.

[16] T. Hussain, Y. Royter, D. Hitko, M. Montes, M. Madhav, I. Milosavljevic, R. Rajavel, S. Thomas, M. Antcliffe, A. Arthur, Y. Boegeman, and M. Sokolich, "First demonstration of sub-0.25 μm-width emitter InP-DHBTs with > 400 GHz ft and > 400 GHz fmax," in *Tech. Dig. 2004 IEEE Int. Electron Devices Meeting*, Dec. 13–15, 2004.

[17] D. Sawdai, P. C. Chang, V. Gambin, X. Zeng, J. Wang, M. Barsky, B. Chan, B. Oyama, A. Gutierrez-Aitken, and A. Oki, "Vertical scaling of planarized InP/InGaAs heterojunction bipolar transistors with $f_t > 350$ GHz and $f_{max} > 500$ GHz," in *2005 IEEE/OSA Int. Conf. Indium Phosphide Related Mater.*, Glasgow, U.K., May 8–12, 2005.

[18] W. Hafez, W. Snodgrass, and M. Feng, "12.5 nm base pseudomorphic heterojunction bipolar transistors achieving $f_T = 710$ GHz and $f_{MAX} = 340$ GHz," *Appl. Phys. Lett.*, vol. 87, p. 252109, 2005.

[19] M. Le, F. Stroili, R. Elder, and J. Feng, "InP HBT electronics for high performance and mixed signal electronic warfare applications," in *Proc. Government Microcircuit Applicat. Circuit Technol. Con.*, Mar. 2006.

[20] J. Hallin, T. Kjellberg, and T. Swahn, "A 165-Gb/s 4 : 1 multiplexer in InP DHBT technology," *IEEE J. Solid-State Circuits*, vol. 41, pp. 2209–2214, Oct. 2006.

[21] P. Asbeck, F. Chang, K.-C. Wang, G. Sullivan, and D. Cheung, "GaAs based heterojunction bipolar transistors for very high performance electronic circuits," *Proc. IEEE*, vol. 81, pp. 1709–1726, Dec. 1993.

[22] T. Ishibashi, "Influence of electron velocity overshoot on collector transit times of HBTs," *IEEE Trans. Electron Devices*, vol. 37, pp. 2103–2105, Sep. 1990.

[23] M. J. W. Rodwell, M. Urteaga, Y. Betser, D. Scott, M. Dahlström, S. Lee, S. Krishnan, T. Mathew, S. Jaganathan, Y. Wei, D. Mensa, J. Guthrie, R. Pullela, Q. Lee, B. Agarwal, U. Bhattacharya, and S. Long, "Scaling of InGaAs/InAlAs HBTs for high speed mixed-signal and mm-wave Ics," *Int. J. High Speed Electron. Syst.*, vol. 11, no. 1, pp. 159–215, 2001.

[24] E. Sano, Y. Matsuoka, and T. Ishibashi, "Device figure-of-merits for high-speed digital

ICs and baseband amplifiers," *IEICE Trans. Electron.*, vol. E78-C, no. 9, pp. 1182–1188, 1995.

[25] M. Vaidyanathan and D. L. Pulfrey, "Extrapolated f of heterojunction bipolar transistors," *IEEE Trans. Electron Devices*, vol. 46, pp. 301–309, Feb. 1999.

[26] R. H. Dennard, F. H. Gaensslen, L. Kuhn, and H. N. Yu, "Design of micron MOS switching devices," in *IEDM Tech. Dig.*, 1972, pp. 168–170.

[27] D. D. Tang and P. M. Solomon, "Bipolar transistor design for optimized power-delay logic circuits," *IEEE J. Solid-State Circuits*, vol. 14, pp. 679–684, Aug. 1979.

[28] T. H. Ning, D. D. Tang, and P. M. Solomon, "Scaling properties of bipolar devices," in *Proc. 1980 IEEE Int. Electron Devices Meeting*, vol. 26, pp. 61–64.

[29] W. B. Joyce and R. W. Dixon, *J. Appl. Phys.*, vol. 46, p. 855, 1975.

[30] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*. Oxford, U.K.: Clarendon, 1959, p. 265.

[31] M. G. Alerstein, "Thermal stability of emitter ballasted HBT's," *IEEE Trans. Electron Devices*, vol. 45, pp. 1653–1655, Aug. 1998.

[32] N. Nenadović, L. K. Nanver, and J. W. Slotboom, "Electrothermal limitations on the current density of high-frequency bipolar transistors," *IEEE Trans. Electron Devices*, vol. 51, pp. 2175–2180, Dec. 2004.

[33] E. F. Chor, D. Zhang, H. Gong, W. K. Chong, and S. Y. Ong, "Electrical characterization, metallurgical investigation, and thermal stability studies of (Pd, Ti, Au)-based ohmic contacts," *J. Appl. Phys.*, vol. 87, no. 5, pp. 2437–2444, Mar. 1, 2000, AIP.

[34] U. Singisetti, A. M. Crook, S. R. Bank, E. Lind, J. D. Zimmerman, M. A. Wistey, A. C. Gossard, and M. J. M. Rodwell, "Ultra-low resistance ohmic contacts to InGaAs/InP," in *Proc. 2007 IEEE Device Res. Conf.*

[35] J. S. Kofol, B. J. F. Lin, M. Mierzwinski, A. Kim, A. Armstrong, and R. Van Tuyl, "A backside via process for thermal resistance improvement demonstrated using GaAs HBTs," in *Tech. Dig. 1992 Gallium Arsenide Integr. Circuit (GaAs IC) Symp.*, Oct. 4–7, 1992, pp. 267–270.

[36] K. S. Champling and G. Eisenstein, "Cutoff frequency of submillimeter Schottky-barrier diodes," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-26, pp. 31–34, Jan. 1978.

[37] Z. Griffith, E. Lind, and M. J. W. Rodwell, "Sub-300 nm InGaAs/InP type-I DHBTs with a 150 nm collector, 30 nm base demonstrating 755 GHz fmax and 416 GHz ft," in *Proc. 2007 IEEE Indium Phosphide Related Mater. Conf.*, Matsue, Japan.

[38] Z. Griffith, E. Lind, M. Rodwell, X. Fang, D. Loubychev, Y. Wu, J. Fastenau, and A. Liu, "60 nm collector InGaAs/InP type-I DHBTs demonstrating 660 GHz ft, BVceo= 2.5 V, and BVcbo= 2.7 V," in *Proc. 2006 IEEE Compound Semiconduct. IC Symp.*, San Antonio, TX.

[39] Y. Miyamoto, A. G. Dental, J. M. M. Rios, and S. Chandrasekhar, "GaInAs/InP DHBT incorporating thick extrinsic base and selectivelyregrown emitter," *Electron. Lett.*, vol. 31, no. 17, pp. 1510–1511, Aug. 17, 1995.

[40] Y. Wei, D. Scott, Y. Dong, A. C. Gossard, and M. J. W. Rodwell, "280 GHz ft InP DHBT with 1.2 μm base-emitter junction area in MBE regrown-emitter technology," in *Proc. Device Res. Conf.*, Notre Dame, IN, Jun. 2004, pp. 237–238.

[41] C. Kadow, A. C. Gossard, and M. J. W. Rodwell, "Regrown-emitter InP HBTs," presented at the 22nd North Amer. Conf. Molecular Beam Epitaxy Conf., Alberta, ON, Canada, Oct. 2004.

[42] M. Khater, J.-S. Rieh, T. Adam, A. Chinthakindi, J. Johnson, R. Krishnasamy, M. Meghelli, F. Pagette, D. Sanderson, C. Schnabel, K. T. Schonenberg, P. Smith, K. Stein, A. Stricker, S.-J. Jeng, D. Ahlgren, and G. Freeman, "SiGe HBT technology with f = f = 350 = 300 GHz and gate delay below 3.3 ps," in *Proc. IEEE Int. Electron Device Meeting*, Dec. 2004, pp. 247–250.

[43] T. Oka, K. Hirata, K. Ouchi, H. Uchiyama, K. Mochizuki, and T. Nakamura, "Small-scaled InGaP/GaAs HBT's with WSi/Ti base electrode and buried SiO2," *IEEE Trans. Electron Devices*, vol. 45, pp. 2276–2282, 1998.

[44] V. E. Houtsma, J. Frackoviak, R. F. Kopf, R. R. Reyes, W. Sung, A. Tate, J. Tong, Y. Yang, N. G. Weimann, and Y. K. Chen, "Investigation of stressing InP/InGaAs DHBTs under high current density," in *Proc. 2003 GaAs Mantech*.

[45] K. Feng *et al.*, "Reliability of commercial InP HBTs under high current density lifetests," in *2002 GaAs IC Symposium Proceedings*.

[46] R. Yamabi *et al.*, "Reliability of InGaAs/InP HBTs with InP passivation structure," in *Proc. Indium Phosphide Related Mater. Conf.*, 2003, pp. 122–125.

[47] K. Kiziloglu, S. Thomas, III, F. Williams, Jr., and B. M. Paine, "Reliability and failure criteria for AlInAs/GaInAs/InP HBTs," in *Proc. Indium Phosphide Related Mater. Conf.*, 2000, pp. 294–297.

[48] T. S. Low, M. W. Dvorak, M. Farhoud, R. E. Yeats, M. Iwamoto, G. K. Essilfie, T. Engel, B. Keppeler, J. S. C. Chang, J. Hadley, G. Patterson, F. Kellert, N. Moll, S. R. Bahl, C. P. Hutchinson, E. Ehlers, M. E. Adamski, M. K. Culver, D. C. D'Avanzo, and T. Shirley, "GaAsSb DHBT IC technology for RF and microwave instrumentation," in *Proc. IEEE Compound Semicond. Integr. Circuit Symp. Tech. Dig.*, 2005, pp. 69–72.

[49] Y. Dong, Z. Griffith, M. Dahlstrom, and M. J. W. Rodwell, " *Cbc* reduction in InP heterojunction bipolar transistors with selectively implanted collector pedestal," in *Proc. 2004 Device Res. Conf.*, Notre Dame, IN.

[50] J. C. Li, M. Chen, D. A. Hitko, C. H. Fields, S. Binqiang, R. Rajesh, P. M. Asbeck, and M. S. Marko, "A submicromenter 252 GHz f_T and 283 GHz fmax InP DHBT with reduced CBC using selectively implanted buried subcollector (SIBS)," *IEEE Electron Device Lett.*, vol. 26, pp. 136–138, Mar. 2005.

[51] M. Urteaga, K. Shinohara, R. Pierson, P. Rowell, B. Brar, Z. Griffith, N. Parthasarathy, and M. J. W. Rodwell, "InP DHBT IC technology with implanted collector-pedestal and electroplated device contacts," in *Proc. 2006 IEEE Compound Semiconduct. IC Symp.*, San Antonio, TX.

[52] N. Parthasarathy, Z. Griffith, C. Kadow, U. Singisetti, M. J. W. Rodwell, X.-M. Fang, D. Loubychev, Y. Wu, J. M. Fastenau, and A. W. K. Liu, "Collector-pedestal InGaAs/InP DHBTs fabricated in a single-growth, triple implant process," *IEEE Electron Device Lett.*, vol. 27, pp. 313–316, May 2006.

ABOUT THE AUTHORS

Mark J. W. Rodwell (Fellow, IEEE) received the Ph.D. degree from Stanford University, Stanford, CA, in 1988.

He is a Professor and Director of the Nanofabrication Laboratory, the NSF Nanofabrication Infrastructure Network, and the SRC Nonclassical CMOS Research Center with the University of California, Santa Barbara. He was with AT&T Bell Laboratories during 1982–1984. His research focuses on high-bandwidth InP bipolar transistors, compound semiconductor field-effect transistors for VLSI applications, and millimeter-wave integrated circuit design in both silicon VLSI and III–V processes.

Prof. Rodwell received the 1997 IEEE Microwave Prize for his work on GaAs Schottky-diode ICs for millimeter-wave instrumentation.

Minh Le (Member, IEEE) received the S.B. and S.M. degrees in electrical engineering and physics from the Massachusetts Institute of Technology, Cambridge, in 1995 and 1997, respectively.

He is currently a Manager with Evergreen Solar, Marlboro, MA. From 2000 to 2006, he was the InP Foundry Manager with Vitesse Semiconductor, Camarillo, CA, and managed customer and government programs for the development of InP HBT circuits and technology. From 1997 to 2000, he worked on the development of GaAs MESFET circuit technology.

Berinder Brar (Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the University of California, Santa Barbara (UCSB) in 1991 and 1995, respectively.

At UCSB, he studied the InAs/AlSb/GaSb material system for high-speed electronic and optoelectronic applications. In 1995, he joined the Nanoelectronics branch, Central Research Labs, Texas Instruments, to work on InP-based resonant tunneling devices and field-effect transistors for high-speed mixed-signal applications. Under the Ultra program sponsored by the Defense Advanced Research Projects Agency, he also conducted research on ultrathin crystalline layers grown on silicon substrates to build resonant tunneling structures for silicon-based quantum devices. He joined Teledyne Scientific and Imaging in 1999 and managed the Advanced III–V Devices and Materials department for five years, developing GaAs, InP, and antimonide-based compound semiconductors for low-power mixed-signal and low-noise monolithic microwave integrated circuit applications. He is presently Executive Director of the Electronics Division, Teledyne Scientific and Imaging, Thousand Oaks, CA. He is author or coauthor of more than 50 papers in technical journals.